ISARC 2014

Proceedings of the 31st International Symposium on Automation and Robotics in Construction and Mining

9-11 July 2014
Sydney, NSW, Australia

Edited by:
Quang Ha, Xuesong Shen, and Ali Akbarnezhad
SPONSORS
Welcome to the 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC2014) in Sydney, Australia from July 9th to 11th, at the Aerial Function Centre, the University of Technology, Sydney (UTS). It is our great honour and pleasure to welcome you to ISARC2014.

The ISARC series have been organised by members of the International Association for Automation and Robotics in Construction (IAARC) to address the needs and concerns of a global community in all fields of construction, including civil and building engineering, machine automation, robotics applications to construction, mining automation, infrastructure networks, construction and environmental sustainability, Information Technology innovations, planning, logistics, etc. IAARC welcomes participation from other industrial sectors and governments.

This year’s ISARC theme is Automation, Construction and Environment. Apart from addressing latest advances in automation and robotic technologies for construction, building and mining, ISARC2014 has a specific focus on efficiency, productivity, quality, and reliability attributes of the construction/mining automation process and its interactions with the environment.

ISARC2014 has received a total of 230 submissions from 33 countries, all of which have been peer-reviewed by international experts, Track Chairs and the Program Committee. With 136 papers being included for presentations, the ISARC2014 has an acceptance rate less than 60%. The Technical Program features five Keynotes, including a Tucker-Hasegawa speech, 5 invited papers, 30 parallel sessions, a workshop and 2 Lab tours.

We would like to take this opportunity to gratefully acknowledge ISARC2014 sponsors, namely, The New South Wales Government under its Research Attraction and Acceleration Program, The Australasian Joint Research Centre on Building Information Modelling at Curtin University, The Institute for Infrastructure Engineering at The University of Western Sydney, The Faculty of Design Architecture & Building, Faculty of Engineering & IT at UTS, and IAARC.

We would like to extend our great gratitude to all the contributors who submitted abstracts and papers to ISARC2014. Your excellent work ensures the quality and impact of this symposium! Many thanks also go to all members of the Organising Committee, IAARC Board of Directors, Program Committee, Track Chairs, reviewers, volunteers and everyone who contributed in one way or another to ISARC2014. Your continued efforts are the foundation of the success for this symposium; it has been truly a great pleasure to work with all of you!

In addition to the Technical Program, ISARC2014 has a number of social programs, including Welcome Reception, Farewell Lunch, and Gala Dinner accompanied by a special concert, led by Guest Performers from The Australian National University and Bridgewater State University, USA, on the interrelation of science and technology with nature to illustrate the symposium theme via music.

We do hope that ISARC2014 provides substantial opportunities for collegial networking and friendship development. This is the first time the symposium takes place in the Southern hemisphere, and particularly, at UTS, located in the heart of the spectacular Sydney, which has been ranked among the top cities over the world to live in and to visit.

We wish you a rewarding symposium and a pleasant time here in Sydney, Australia!

Quang Ha
General Chair, ISARC2014
Sydney, Winter 2014
CONFERENCE COMMITTEES

General Chair
Quang Ha

Conference Co-Chairs
Perry Forsythe
Bijan Samali
Xiangyu Wang

Local Organising Committee
Merched Azzi
Leonhard Bernold
Peter Brady
Gamini Dissanayake
Hiep Duc
Ngai-Ming Kwok
Gu Fang
Jianchun Li
Dikai Liu
Ali Akbarnezhad
Le-Tuyen Nguyen
Matt Riley
David Rye
Xuesong (Johnson) Shen

International Committee
Par Ahman  
Yukio Hassegawa
Bursu Akinci  
Arturas Kaklauskas
Tatsuo Arai  
Hyoungkwan Kim
Carlos Balaguer  
Soonwook Kwon
Thomas Bock  
Tai Sik Lee
Ioannis Brilakis  
Alan Lytle
Carlos Caldas  
Ger Maas
Alessandro Carbonari  
Fiorenzo Malaguti
Min-Yuan Cheng  
Osama Moselhi
Seokho Chi  
Ronie Navon
Moon Young Cho  
Yehiel Rosenfeld
Fred Delabbio  
Kamel Saidi
Ernesto Gambao  
Miroslaw Skibniewski
Jozef Gasparik  
Jochen Teizer
Frans van Gassel  
Richard Tucker
Quang Ha  
Koshy Varghese
Carl Haas  
Xiangyu Wang
TRACKS AND TRACK CHAIRS

Track Automation and Control
Chairs Jianchun Li (University of Technology, Sydney, Australia)
Ngai-Ming Kwok (University of New South Wales, Australia)

Track Building and Architecture
Chair Ioannis Brilakis (Cambridge University, UK)

Track Construction Management
Chairs Seokho Chi (Seoul National University, Korea Republic)
Manwoo Park (Myongji University, Korea Republic)

Track Energy and Environment
Chair Alessandro Carbonari (Università Politecnica delle Marche, Italy)

Track IT Applications
Chairs Ali Akbarnezhad (University of New South Wales, Australia)
Peter Brady (University of Technology, Sydney, Australia)

Track Mining, Built Infrastructure and Human Factors
Chairs Tim Horberry (The University of Queensland, Australia)
Burcu Akinci (Carnegie Mellon University, USA)
Tri (Anthony) Tran (Nanyang Technological University, Singapore)

Track Robotics and Mechatronics
Chair Hung La (University of Nevada, USA)

Track Sensing and Communication
Chair Xuesong (Johnson) Shen (University of New South Wales, Australia)
### REVIEWERS

<table>
<thead>
<tr>
<th>Alireza Ahmadian</th>
<th>Habib Fathi</th>
<th>Hyoungkwan Kim</th>
<th>Samad M.E. Sepasgozar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changbum Ahn</td>
<td>Jiali Fu</td>
<td>Jaewon Kim</td>
<td>Seyed Mohsen</td>
</tr>
<tr>
<td>Ali Akbarnezhad</td>
<td>John Gambatese</td>
<td>Christian Koch</td>
<td>Shahandashti</td>
</tr>
<tr>
<td>Mohammed Sadiq Altai</td>
<td>Jozef Gašparik</td>
<td>Ngai-Ming Kwok</td>
<td>Xuesong Shen</td>
</tr>
<tr>
<td>Engin Anil</td>
<td>Frans van Gassel</td>
<td>Hung La</td>
<td>Weihua Sheng</td>
</tr>
<tr>
<td>Meghdad Attarzadeh</td>
<td>Christos Georgoulas</td>
<td>Jim La</td>
<td>Haiyan Shi</td>
</tr>
<tr>
<td>Rita Awwad</td>
<td>Alireza Golabchi</td>
<td>Trung Le</td>
<td>Yoon-Jung Shin</td>
</tr>
<tr>
<td>Ehsan Rezazadeh Azar</td>
<td>Jie Gong</td>
<td>Nan Li</td>
<td>Reihaneh Sarraf Shirazi</td>
</tr>
<tr>
<td>Burcin Becerik-Gerber</td>
<td>Ye Gu</td>
<td>Li-Chuan Lien</td>
<td>Mohsin Siddiqui</td>
</tr>
<tr>
<td>Amir Behzadan</td>
<td>Sanghyeok Han</td>
<td>Samburg Lim</td>
<td>JeongWook Son</td>
</tr>
<tr>
<td>Michael Belsky</td>
<td>SangUk Han</td>
<td>Soram Lim</td>
<td>Sungjoon Suk</td>
</tr>
<tr>
<td>Frédéric Bosché</td>
<td>Carol Hon</td>
<td>Fei Liu</td>
<td>Anthony Tran</td>
</tr>
<tr>
<td>Peter Brady</td>
<td>Pranjal Jain</td>
<td>Ming Lu</td>
<td>Yelda Turkan</td>
</tr>
<tr>
<td>Ioannis Brilakis</td>
<td>Farrokh Jazizadeh Karimi</td>
<td></td>
<td>Vahid Vakiloroaya</td>
</tr>
<tr>
<td>Alessandro Carbonari</td>
<td>Youngsoo Jung</td>
<td>Xiaowei Luo</td>
<td>Frans van-Gassel</td>
</tr>
<tr>
<td>Daniel Castro-Lacouture</td>
<td>Alireza S. Kaboli</td>
<td>Ger Maas</td>
<td>Esa Viljamaa</td>
</tr>
<tr>
<td>Greg Chan</td>
<td>Vineet Kamat</td>
<td>Rohana Mahbub</td>
<td>Han-Hsiang Wang</td>
</tr>
<tr>
<td>V Paul C Charlesraj</td>
<td>Julian Kang</td>
<td>Carol Menassa</td>
<td>Johnny Wong</td>
</tr>
<tr>
<td>Pei-Yin Chen</td>
<td>Youngcheol Kang</td>
<td>Sungkon Moon</td>
<td>Paul Bo Xia</td>
</tr>
<tr>
<td>Seokho Chi</td>
<td>Alireza Khalili</td>
<td>Zahra Sadat Moussavi Nadoushani</td>
<td>Jamie Yeung</td>
</tr>
<tr>
<td>Angelo Ciribini</td>
<td>Kourosh Khoshelham</td>
<td>Mohammad Nahangi</td>
<td>Jingang Yi</td>
</tr>
<tr>
<td>Fei Dai</td>
<td>Byoung-Il Kim</td>
<td>Madhav Nepal</td>
<td>Sungmin Yun</td>
</tr>
<tr>
<td>Jianhao Du</td>
<td>Changwan Kim</td>
<td>Luan Nguyen</td>
<td>Payam Zekavat</td>
</tr>
<tr>
<td>Esin Ergen</td>
<td>Changyoon Kim</td>
<td>Thang Nguyen</td>
<td>Guangcong Zhang</td>
</tr>
<tr>
<td>Semiha Ergan</td>
<td>Dae Young Kim</td>
<td>Man-Woo Park</td>
<td>Yizhai Zhang</td>
</tr>
<tr>
<td>Gu Fang</td>
<td>Duyon Kim</td>
<td>Abbas Rashidi</td>
<td>Zhenhua Zhu</td>
</tr>
</tbody>
</table>
CONTENTS

SPONSORS iii

FOREWORD - MESSAGE FROM THE CHAIR iv

CONFERENCE COMMITTEES v

TRACKS AND TRACK CHAIRS vi

REVIEWERS vii

TUCKER-HASEGAWA KEYNOTE 1

Construction Automation Needs and Challenges in Emerging Countries 2

Koshy Varghese

KEYNOTE PAPERS 3

The Challenges and Trends of Building Information Modelling (BIM) for Construction and Resources Sectors 4

Xiangyu Wang and Heap-Yih Chong

Options for Sustainable Energy Planning for Australia 7

Merched Azzi

Ki-NEW-Matics 11

Thomas Bock

Towards Fully Automated Tunnel Inspection: A Survey and Future Trends 19

C. Balaguer, R. Montero, J. G. Victores, S. Martinez, A. Jardón

INVITED PAPERS 34

Economic Model Predictive Control - A Review 35

Tri Tran, Keck Voon Ling and Jan M. Maciejowski

Infrastructure Robotics: Research Challenges and Opportunities 43

Dikai Liu, Gamini Dissanayake, Jaime Valls Miro and Kenneth Waldron

Visual and Acoustic Data Analysis for the Bridge Deck Inspection Robotic System 50

Hung La, Nenad Gucunski, Seong-Hoon Kee and Luan Nguyen

Demand Response Under the Smart Grid Paradigm 58

Michael Negnevitsky
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Future Intelligent Civil Structures: Challenges and Opportunities</td>
<td>72</td>
</tr>
<tr>
<td>Jianchun Li, Yancheng Li, Mohsen Askari and Quang Ha</td>
<td></td>
</tr>
<tr>
<td><strong>AUTOMATION AND CONTROL</strong></td>
<td>80</td>
</tr>
<tr>
<td>Control of An Adaptive Light Shelf Using Multi-Objective Optimization</td>
<td>81</td>
</tr>
<tr>
<td>Benny Raphael</td>
<td></td>
</tr>
<tr>
<td>Automatic Positioning and Alignment for Hole Navigation in Surface Drilling</td>
<td>88</td>
</tr>
<tr>
<td>Tuomo Pirinen, Juha Lassila, Mikko Loimusalo, Juha Pursimo and Sami Hanski</td>
<td></td>
</tr>
<tr>
<td>A Study on the Thermal Crack Control of Large Turbine Foundation using Automated Curing System</td>
<td>95</td>
</tr>
<tr>
<td>Ju-Hyung Ha, Ok-Pin Na, Chang-Keun Lim and Yun-Gu Cho</td>
<td></td>
</tr>
<tr>
<td>Integrated Approach to Machine Guidance and Operations Monitoring in Tunnel Construction</td>
<td>103</td>
</tr>
<tr>
<td>Xuesong Shen, Ming Lu, Sheng Mao and Xiaodong Wu</td>
<td></td>
</tr>
<tr>
<td>Dynamic Risk Assessment in Construction Projects Using Bayesian Networks</td>
<td>110</td>
</tr>
<tr>
<td>Limao Zhang, Xianguo Wu, Miroslaw J. Skibniewski and Jingbing Zhong</td>
<td></td>
</tr>
<tr>
<td>Nonlinear Flatness-Based Controller for Permanent Magnet-Excited Synchronous Motor</td>
<td>120</td>
</tr>
<tr>
<td>Pham Tam Thanh and Nguyen Dinh That</td>
<td></td>
</tr>
<tr>
<td>Modified Discrete Event Simulation Algorithm for Control of Automated Construction Operations</td>
<td>126</td>
</tr>
<tr>
<td>Joseph Louis and Phillip Dunston</td>
<td></td>
</tr>
<tr>
<td>Performance Test for Rapid Surface Modeling of Dynamic Construction Equipment from Laser Scanner Data</td>
<td>134</td>
</tr>
<tr>
<td>Chao Wang and Yong Cho</td>
<td></td>
</tr>
<tr>
<td>Space Technologies of Life Support Systems for the Metropolitan Cities</td>
<td>142</td>
</tr>
<tr>
<td>Natalia Buzalo, Alexej Bulgakov and Thomas Bock</td>
<td></td>
</tr>
<tr>
<td>Offshore Container Crane Systems with Robust Optimal Sliding Mode Control</td>
<td>149</td>
</tr>
<tr>
<td>R. M. T. Raja Ismail, Nguyen D. That and Quang Ha</td>
<td></td>
</tr>
<tr>
<td>Force Control of Cleaning Tool System for Building Wall Maintenance Robot on Built-in Guide Rail</td>
<td>157</td>
</tr>
<tr>
<td>Chang-Yeob Shin, Sung-Min Moon, Jun-Ho Kwon, Jaemyung Huh and Daehie Hong</td>
<td></td>
</tr>
<tr>
<td>Towards Autonomous Robotic In-Situ Assembly on Unstructured Construction Sites Using Monocular Vision</td>
<td>163</td>
</tr>
<tr>
<td>Chen Feng, Yong Xiao, Aaron Willette, Wesley McGee and Vineet Kamat</td>
<td></td>
</tr>
<tr>
<td>Automated Dynamic Management of Road Construction Sites</td>
<td>171</td>
</tr>
<tr>
<td>Rauno Heikkilä, Esa Viljamaa, Annemari Kaaranka, Tomi Makkonen and Irina Peltoma</td>
<td></td>
</tr>
<tr>
<td>A New Hysteretic Model for Magnetorheological Elastomer Base Isolator and Parameter Identification Based on Modified Artificial Fish Swarm Algorithm</td>
<td>176</td>
</tr>
<tr>
<td>Yang Yu, Yancheng Li and Jianchun Li</td>
<td></td>
</tr>
</tbody>
</table>
## Building and Architecture

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of Lightweight BIM Shape Format Structure to Represent Large Volume Geometry Objects Using GIS with Facility Management</td>
<td>184</td>
</tr>
<tr>
<td>Tae Wook Kang and Chang Hee Hong</td>
<td></td>
</tr>
<tr>
<td>The Adoption of Building Information Modeling in the Design Organization: An Empirical Study of Architects in Korean Design Firms</td>
<td>185</td>
</tr>
<tr>
<td>Hyojoon Son, Sungwook Lee, Nahyae Hwang and Changwan Kim</td>
<td></td>
</tr>
<tr>
<td>A Demonstration of BIM-enabled Quantitative Circulation Analysis using BERA Language</td>
<td>194</td>
</tr>
<tr>
<td>Hyunsoo Lee, Jisoo Kim, Minkyu Shin, Inhan Kim and Jin-Kook Lee</td>
<td></td>
</tr>
<tr>
<td>Graph-based Representation of Building Circulation with the Most-Remote Point and Virtual Space Objects</td>
<td>202</td>
</tr>
<tr>
<td>Jisoo Kim, Hyunsoo Lee, Minkyu Shin, Jinwon Choi and Jin-Kook Lee</td>
<td></td>
</tr>
<tr>
<td>An Integrated Approach to the Visualization of Indoor Temperature Changes on the Floor Plans</td>
<td>210</td>
</tr>
<tr>
<td>Minkyu Shin, Sang-Jk Lee, Jisoo Kim, Hyunsoo Lee, Gyuyeob Jeon and Jin-Kook Lee</td>
<td>217</td>
</tr>
<tr>
<td>Application of Dijkstra's Algorithm in the Smart Exit Sign</td>
<td>224</td>
</tr>
<tr>
<td>Jehyun Cho, Ghang Lee, Jongsung Won and Eunseo Ryu</td>
<td></td>
</tr>
<tr>
<td>Smart Housing for Reducing Energy Use in Single Households, Emphasizing Effective Energy Management</td>
<td>230</td>
</tr>
<tr>
<td>Sung Jun Park, Mi Jeong Kim and Sung-Yul Kim</td>
<td></td>
</tr>
<tr>
<td>Is There Really A Case for Off-site Manufacturing?</td>
<td>238</td>
</tr>
<tr>
<td>Edward Duc, Perry Forsythe and Kirsten Orr</td>
<td></td>
</tr>
<tr>
<td>Building Arrangement Optimization for Urban Ventilation Potential Using Genetic Algorithm and CFD Simulation</td>
<td>247</td>
</tr>
<tr>
<td>Jongyeon Lim and Ryozo Ooka</td>
<td></td>
</tr>
</tbody>
</table>

## Construction Management

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wood-Frame Wall Panel Sequencing Based on Discrete-Event Simulation and Particle Swarm Optimization</td>
<td>253</td>
</tr>
<tr>
<td>Mohammed Sadiq Aliaf, Mohamed Al-Hussein and Haitao Yu</td>
<td></td>
</tr>
<tr>
<td>Integrated Visualization and Simulation for Lifting Operations of Modules under Congested Environment</td>
<td>254</td>
</tr>
<tr>
<td>Sanghyeok Han, Zhen Lei, Ahmed Bouferguène, Mohamed Al-Hussein and Ulrich Hermann</td>
<td>262</td>
</tr>
<tr>
<td>Post-Simulation Visualization Application for Production Improvement of Modular Construction Manufacturing</td>
<td>270</td>
</tr>
<tr>
<td>Mana Moghadam, Beda Barkokbas and Mohamed Al-Hussein</td>
<td></td>
</tr>
<tr>
<td>Improving Productivity of Yard Trucks in Port Container Terminal Using Computer Simulation</td>
<td>278</td>
</tr>
</tbody>
</table>

---

x
Essmeil Ahmed, Tarek Zayed and Sabah Alkass

BIM-based Integrated Framework for Detailed Cost Estimation and Schedule Planning of Construction Projects
Hexu Liu, Ming Lu and Mohamed Al-Hussein

A Discrete Firefly Algorithm for the Scaffolding Modular Construction in Mega Projects
Jianjun Liu, Lei Hou and Xiangyu Wang

Enhanced Online LS-SVM Using EMD Algorithm for Prices Prediction of Building Materials
Ying-Hao Yu, Hsiao-Che Chien, Pi-Hui Ting, Jung-Yi Jiang and Pei-Yin Chen

An Interactive Progress Monitoring System using Image Processing in Mobile Computing Environment
Hongjo Kim, Kinam Kim, Sungjae Park, Jihoon Kim and Hyungkwan Kim

Constructability and Safety Performance based Design: a Design and Assessment Tool for the Building Process
Pietro Capone, Vito Getuli and Tommaso Giusti

Integrating Building Information Modelling and Firefly Algorithm to Optimize Tower Crane Layout
Jun Wang, Jianjun Liu, Wenchi Shou, Xiangyu Wang and Lei Hou

A Model for Construction Contractor Selection Using Competitive Intelligence (CI)
Mahdi Safa, Arash Shahi, Carl Haas, Majeed Safa, Keith Hipel, Sandra MacGillivray and Dawn Fiander-Mccann

Safety Management for Existing Buildings in Tunnel Construction
Limao Zhang, Xianguo Wu, Miroslaw J. Skibniewski and Hongyu Chen

Developing Technology-assisted Multi-disciplinary Learning Strategies
Shalini Gandhi, Shankar Sankaran, Michael Er, Kirsten Orr and Hadi Khabbaz

Safety Practices in the Lebanese Construction Market: Contractors’ Perspective
Rita Awwad, Melanie Jabbour and Omar El Souki

5D-BIM: A Case Study of An Implementation Strategy in the Construction Industry
Anoop Sattineni and Jennifer A. Macdonald

Simulation of On-Shore Wind Farm Construction Process in Lebanon
Emile Zankoul and Hiam Khoury

Developing An Ontology-Based Representation Framework for Establishing Cost Analysis Knowledge Base for Construction Work Items
Han-Hsiang Wang, Shao-Wei Weng, Abdoul Aziz Gansonre and Wei-Chih Wang

Estimating the Costs, Energy Use and Carbon Emissions of Concrete Recycling Using Building Information Modelling
Ali Akbarnezhad and Zahra Moussavi Nadoushani

An Integrated 5D Tool for Quantification of Construction Process Emissions and Accident Identification
Johnny Kwok Wai Wong, Heng Li, Greg Chan, Haoran Wang, Ting Huang and Eric Luo

xi
Employing Ant Colony for the Optimal Reduction of Project Risk Severity
Nael Zabel, Maged Georgy and Moheeb Ibrahim 398

Incorporating Uncertainty into Project Schedule Crashing: An Algorithm
Ehab Subhy, Maged Georgy and Moheeb Ibrahim 404

Automation for Mobile Crane Motion Planning in Industrial Projects
Zhen Lei, Sanghyeok Han, Ahmed Bouferguène, Mohamed Al-Hussein and Ulrich Hermann 410

A Study on Significance of System Dynamics Approach in Understanding Adoption of Information Technology in Building Construction Projects
Vinay Mathews, Koshi Varghese and Ashlin Mahalingam 418

Learning Effect of Interior Finishing and Building Services Works for Multi-Dwelling Complex Project
Naoto Mine, Abdullaete Olanrewaju and Masaki Takeuchi 426

A Knowledge-Based Framework for Quantity Takeoff and Cost Estimation in the AEC Industry Using BIM
Shiva Aram, Charles Eastman and Rafael Sacks 434

The Power of Technology in Bridge Construction Project Management
Fahimeh Zaeri and James Rotimi 443

A BIM Based Construction Site Layout Planning Framework Considering Actual Travel Paths
Jack Cheng and Srinath Kumar 450

Contrivances to Assist Forest Machine Operator on Forest Road with Steep Slope
Katsutoshi Saibara, Shigeomi Nishigaki, Fujio Matsuda and Shinichi Kubota 458

Importance of Planning for the Transport Stage in Procurement of Construction Materials
Alireza Ahmadian Fard Fini, Ali Akbarnezhad, Taha Hossein Rashidi and Steven Travis Waller 466

A Formwork Layout Model based on Genetic Algorithm
Dongmin Lee, Hyunsu Lim, Taehoon Kim, Hunhee Cho and Kyung-In Kang 474

The Case for BIM Uptake among Small Construction Contracting Businesses
Perry Forsythe 480

ENERGY AND ENVIRONMENT 488

Energy-Efficient Air-Cooled DX Air-Conditioning Systems with Liquid Pressure Amplification
Vahid Vakiloroaya and Quang Ha 489

A Closed-loop System of Construction and Demolition Waste Recycling
Jane Brennan, Grace Ding, Robert Wonschik and Kirk Vessalas 499

Development and First Testing of a Framework for Predictive Energy Control of Underground Stations

xii
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Specific Trajectory Profile Selection for Energy Efficient Servo Drive Movements</td>
<td>514</td>
</tr>
<tr>
<td>Christian Hansen, Kai Eggers, Jens Kotliarski and Tobias Ormaier</td>
<td></td>
</tr>
<tr>
<td>Implications of Legal Frameworks on Construction and Demolition Waste Recycling - A Comparative Study of the German and Australian Systems</td>
<td>523</td>
</tr>
<tr>
<td>Robert Wonschik, Jane Brennan, Grace Ding, Andrea Heilmann and Kirk Vessalas</td>
<td></td>
</tr>
<tr>
<td>Predicting Energy Usage Using Historical Data and Linear Models</td>
<td>531</td>
</tr>
<tr>
<td>Majeed Safa, Jeremy Allen and Mahdi Safa</td>
<td></td>
</tr>
<tr>
<td>Labor Management in Masonry Construction: A Sustainable Approach</td>
<td>536</td>
</tr>
<tr>
<td>Laura Florez and Daniel Castro-Lacouture</td>
<td></td>
</tr>
<tr>
<td>Modelling October 2013 Bushfire Pollution Episode in New South Wales, Australia</td>
<td>544</td>
</tr>
<tr>
<td>Hiep Nguyen Duc, Sean Watt, David Salter and Toan Trieu</td>
<td></td>
</tr>
<tr>
<td>Towards Measuring the Impact of Personal Control on Energy Use through the Use of Immersive Virtual Environments</td>
<td>549</td>
</tr>
<tr>
<td>Arsalan Heydarian, Joao Carneiro, David Gerber and Burcin Becerik-Gerber</td>
<td></td>
</tr>
<tr>
<td>Automatic Power Managing and Monitoring System Applying for Underground Mines in Vinacomin</td>
<td>557</td>
</tr>
<tr>
<td>Tuan Anh Vu and The Nam Vu</td>
<td></td>
</tr>
<tr>
<td>Integrating Energy Simulation in Energy Saving Analysis of Taiwan’s Green Hospital Buildings</td>
<td>561</td>
</tr>
<tr>
<td>Po-Han Chen and Meng-Shen Kan</td>
<td></td>
</tr>
<tr>
<td>Measurement Scheme and Automatic Prediction for Ground Vibration Induced by High-Speed Rail on Embankments</td>
<td>568</td>
</tr>
<tr>
<td>Yit-Jin Chen, Yi-Jiun Shen and Chi-Jim Chen</td>
<td></td>
</tr>
<tr>
<td>Daniel Castro-Lacouture, Steven Jige Quan and Perry Pei-Ju Yang</td>
<td></td>
</tr>
<tr>
<td>Prediction of NOx Vehicular Emissions using On-Board Measurement and Chassis Dynamometer Testing</td>
<td>584</td>
</tr>
<tr>
<td>Seth Daniel Oduro, Santanu Metia, Hiep Duc, Guang Hong and Quang Ha</td>
<td></td>
</tr>
<tr>
<td>IT APPLICATIONS</td>
<td>592</td>
</tr>
<tr>
<td>Application of Building Information Modeling in Designing Fire Evacuation – A Case Study</td>
<td>593</td>
</tr>
<tr>
<td>Kun-Chi Wang, Shih-Yu Shih, Wen-Shuo Chan, Wei-Chih Wang, Shih-Hsu Wang, Abdoul-Aziz Gansonre, Jang-Jeng Liu, Ming-Tsung Lee, Yuan-Yuan Cheng and Ming-Feng Yeh</td>
<td></td>
</tr>
<tr>
<td>General Information Model and Its Application for Rock Excavation in Underground Work Sites</td>
<td>602</td>
</tr>
<tr>
<td>Esa Viljamaa, Irina Peltomaa and Annemari Kaaranka</td>
<td></td>
</tr>
<tr>
<td>Performance and Impacts of Web-based Project Management Systems in Construction</td>
<td>610</td>
</tr>
</tbody>
</table>
Projects

Hemanta Doloi

Integration of Augmented Reality and Indoor Positioning Technologies for On-site Viewing of BIM Information
Hung-Ming Chen and Ting-Yu Chang

Towards 3-D Shape Restructuring for Rapid Prototyping of Joining Interface System
Seong-Ki Lee, Christos Georgoulas and Thomas Bock

Reactive Adaptation of Construction Schedules by Applying Simulation-based Optimization
Kamil Szczesny and Markus König

Detection of Internal Defects in As-Built Pipelines for Structural Health Monitoring: A Sensor Fusion Approach Using Infrared Thermography Combined with 3D Laser-Scanned Data
Hyojoo Son, Changmin Kim and Changwan Kim

Roadmap to Guide Construction Safety Research Commercialization
Aaron Costin, Jakub Felkl, Olga Golovina and Jochen Teizer

Exploring Local Feature Descriptors for Construction Site Video Stabilization
Jung Yeol Kim and Carlos Caldas

A Novel Inference Model for Post-Earthquake Bridge Safety and Failure Probabilities Prediction - A Case Study in Taiwan
Min-Yuan Cheng, Yu-Wei Wu, Yung-Fang Chiu, Yu-Chen Ou and Chien-Kuo Chiu

BIM based Schedule Control for Precast Concrete Supply Chain
Jaakko Nissilä, Rauno Heikkilä, Ilkka Romo, Mikko Malaska and Timo Aho

Information Modelling based Tunnel Design and Construction Process
Rauno Heikkilä, Annemari Kaaranka and Tomi Makkonen

Rapid 3D Modeling of an Existing Building using Photos
Yang Liu and Julian Kang

Using Benders Decomposition for Solving Ready Mixed Concrete Dispatching Problems
Mojtaba Maghrebi, Vivek Periaraj, S. Travis Waller and Claude Sammut

Mining Rules for Satellite Imagery Using Evolutionary Classification Tree
Li-Chuan Lien, Yan-Ni Liu, Min-Yuan Cheng and I-Cheng Yeh

Roadwork Site 3D Virtual Visualization Using Open Source Game Engine and Open Information Transfer
Tomi Makkonen, Rauno Heikkilä, Annemari Kaaranka and Kalervo Nevala

The Implementation of BIM In a Large European Construction Company
Ger Maas

Visualization of As-built Progress Data Using Construction Site Photographs: Two Case Studies
Hossein Jadidi, Mehdi Ravanshadnia and Mujtaba Hossein Alipour

A Computational Framework for Estimating the Carbon Footprint of Construction
A Mixed-integer Nonlinear Programming Model for Minimising Construction Site Noise Levels through Site Layout Optimisation
Ahmed Hammad, David Rev and Ali Akbarnezhad

MINING, BUILT INFRASTRUCTURE AND HUMAN FACTORS

A New LMS Algorithm and Its Application to Improve Quality of Broadcast-telephone System Used in Underground Coal Mines
Nguyen The Truyen and Nguyen The Vinh

Optimization and Evaluation of Automatic Rigging Path Guidance for Tele- Operated Construction Crane
Hung-Lin Chi, Shih-Chung Kang, Shang-Hsien Hsieh and Xiangyu Wang

Visualization Requirements of Engineers for Risk Assessment of Embankment Dams
Varun Kasireddy, Semiha Ergan and Burcu Akinci

Ineffective Rock Breaking and Its Impacts on Pick Failures
Yong Sun and Xingsheng Li

ROBOTICS AND MECHATRONICS

Control of Hovering Altitude of a Quadrotor with Shifted Centre of Gravity for Inspection of High-rise Structures
Alexey Bulgakov, Sergey Emelianov, Thomas Bock and Daher Sayfeddine

Bulldozer as a Mechatronics System with the Intelligent Control
Alexey Bulgakov, Thomas Bock and Georgy Tokmakov

Generation the 3D Model Building by Using the Quadcopter
Torsten Bertram, Thomas Bock, Alexey Bulgakov and Alexey Evgenov

Performance Evaluation of Mecanum Wheeled Omni-directional Mobile Robot
Baeksuk Chu

Cooperative Control of a Single-user Multi-robot Teleoperated System for Maintenance in Offshore Plants
Sung-hoon Eom, Seungyeol Lee, Deajin Kim, Donghun Shin and Jeon Il Moon

Home Environment Interaction via Service Robots and the Leap Motion Controller
Christos Georgoulas, Ahmad Raza, Jorg Guttler, Thomas Linner and Thomas Bock

SENSING AND COMMUNICATION

Experimental Study of Wireless Sensor Networks for Indoor Construction Operations
Magdy Ibrahim and Osama Moselhi

Controlling Slab Flatness Automatically using Laser Scanning and BIM
Frédéric Bosché and Emeline Guenet
Chip-based Real-time Gesture Tracking for Construction Robot’s Guidance
Ying-Hao Yu, Chun-Hsien Yeh, Tsu-Tian Lee, Pei-Yin Chen and Yeu-Horng Shiau

Comparative Experimental Evaluation of Dust Sensors for Environmental Monitoring on Construction Sites
Alessandro Carbonari, Gabriele Fava and Berardo Naticchia

Extended Range Guidance System for the Teleoperation of Microtunnelling Machines
Alberto Jardón, Santiago Martínez, Juan Gonzalez Victores and Carlos Balagué

A Preliminary Investigation into Automated Identification of Structural Steel Without A Priori Knowledge
Jamie Yeung, Mohammad Nahangi, Scott Walbridge and Carl Haas

Near-Miss Accident Detection for Ironworkers Using Inertial Measurement Unit Sensors
Sepideh S. Aria, Kanghyeok Yang, Changbum R. Ahn and Mehmct C. Vuran

Localization inside Tunnels Using Machine Vision
Yo-Ming Hsieh and Yu-Chin Liao

Performance Assessment of Wireless Data Capture in Construction
Payam Rahnamaviezkavat, Sungkon Moon and Leonhard Bernold

Implementation of As-Built Information Modelling Using Mobile and Terrestrial Lidar Systems
Samad M.E. Sepasgozar, Samsung Lim, Sara Shirowzhan and Yong Min Kim

Three Dimensional Spatial Metrics for Compactness Assessment of Urban Forms
Sara Shirowzhan and Samsung Lim

Human Detection using Gradient Maps and Golden Ratio
Feng Su and Gu Fang

Autocorrelation Statistics-Based Algorithms for Automatic Ground and Non-ground Classification of Lidar Data
Sara Shirowzhan and Samsung Lim

POSTER PAPERS

A Framework for Supporting Planning and Development of Infrastructure Projects from a Societal Perspective
Hemanta Doloi

Lessons Learnt from the Vernacular Architecture of Bedouins in Siwa Oasis, Egypt
Riham Ahmed

Efficient Method of An Optimum Construction Company Supplier Selection Supported by Software
Jozef Gašparík and Peter Bažík

SOH Estimation of Lithium-ion Batteries for Electric Vehicles
Datong Qin, Jingying Huang and Wei Sun
Robust Method for Detecting the HRI Device using RANSAC
Myeongsu Gil, Min-Sung Kang, Yongseok Lee, Seunghoon Lee, Sangho Kim and Changsoo Han

Knowledge-based Building Information Modeling (K-BIM) for Facilities Management
V. Paul C. Charlesraj

Algorithm for Economic Assessment of Infrastructure Adaptation to Climate Change
Hoyoung Jeong, Hyounkyu Lee, Hongjo Kim and Hyoungkwan Kim

Development of the BIM based Process for Dredging Works
Rauno Heikkilä, Tapio Leinonenb, Heikki Paukkeric and H. Virtanen

The Applicability of a Geomagnetic Field based Positioning Technique with Mobile Phone to Underground Tunnels
Tomi Makkonen, Rauno Heikkiläa and Annemari Kaaranaka

Geotechnical Monitoring for Safe Excavation of Large Rock Cavern: A Case Study
A. Mandal, C. Kumar, Altaf Usmani and A. Nanda

Decision Support Model with Life Cycle Assessment for Building in Design Phase
Chia-Chen Wei, Chia-Chi Hsiang and Tzu-Chi Shan

A Study on Vertical Zoning Algorithm of Real-Time Construction Lift Control for High-Rise Building
Joong-Hwan Shin, Soon-Wook Kwon and Dong-Hyun Kim

Introduction of Human-Robot Cooperation Technology at Construction Sites
Seungyeol Lee and Jeon Il Moon

Study on the Optimal Digging Range for Intelligent Excavation
Jeonghwan Kim, Joonhyun Jang, Leeho Lee and Jongwon Seo

Information Modelling on Mechanized Earthworks
Shigeomi Nishigaki, Katsutoshi Saibara and Shigeo Kitahara

AUTHOR INDEX

xvii
AUTOMATION, CONSTRUCTION AND ENVIRONMENT

TUCKER-HASEGAWA KEYNOTE
Construction Automation Needs and Challenges in Emerging Countries

Koshy Varghese

Department of Civil Engineering, Indian Institute of Technology, Madras, India
E-mail: koshy@iitm.ac.in

Abstract -

The need for automation has conventionally been driven by the shortage and rising costs of skilled workers. There has been a perception that in emerging countries such as India, the priority for automation is low due to the availability of skilled workers. However, the massive scale of infrastructure development planned in these nations for the next few years cannot be achieved without appropriate automation in all phases of a construction project. There are several cases where companies invest in automation technologies but don’t get the benefit due to inappropriateness of the technology to local conditions. There are also cases where the automation strategy and implementation is carefully planned and results in immense benefits. This talk will present the needs, opportunities and challenges of construction automation in emerging nations requiring. Cases of implementing both process automation as well as field automation technologies will be presented and the lessons learned from these discussed.
AUTOMATION, CONSTRUCTION AND ENVIRONMENT

KEYNOTE PAPERS
The Challenges and Trends of Building Information Modelling (BIM) for Construction and Resources Sectors

Xiangyu Wang and Heap-Yih Chong

Australasian Joint Research Centre for Building Information Modelling, School of Built Environment, Curtin University, Australia
E-mail: Xiangyu.wang@curtin.edu.au, heapyih.chong@curtin.edu.au

Abstract -
Productivity is a comprehensive problem in many countries. Particularly in construction and resources sectors, low productivity will potentially discourage future investments as building assets is becoming more and more unaffordable and unsustainable. In recent years, Building Information Modelling (BIM) is becoming much more active than ever. In order to fully realize where BIM is believed to be, there has to be a series of mechanisms in place to properly adopt BIM for construction site daily activities, beyond the design and engineering phases. BIM must go beyond itself by being integrated with other technologies towards dynamic planning. The paper also gives an insightful summary of where we are and where we should be, backboned by the BIM cases from construction and resources sectors.

Keywords -
Productivity; Integration; Building Information Modelling (BIM); Challenges; Trends.

1 Introduction

Building Information Modelling (BIM) has been proved as a practical and beneficial technology in numerous projects across the globe. While some parties are promoting the use of BIM in their projects, the governments have regulated certain policies for mandating the use of BIM in their countries. For example, the recent initiative in United Kingdom plans to level-up the mandate for BIM Level Two adoption by 2016. It is about making the construction process quicker and more efficient by digitally modelling a building collaboratively in a virtual environment [1]. The initiative reflects a relatively high level of maturity of BIM adoption in certain countries.

In reality, BIM is regarded to be rather established in building industry, and currently, the technology has moved into other sectors, particularly for infrastructure projects. The infrastructure projects involve complicated processes. The project productivity is always one of the main agendas. BIM is always facing a new challenge to create a productive and collaborative working environment along the entire life cycle process. The paper reckons BIM should be integrated with other technologies or managerial approaches throughout the life cycle. The paper articulates three stages where BIM are integrated to enhance the project productivity, namely, pre-construction phase, construction phase and maintenance phase.

Finally, the paper shares insights into the challenges and the way forward for BIM, particularly for its integration with other technologies in contributing to project performance and productivity in various stages of project life cycle.

2 BIM Integration for Project Performance and Productivity

Design alternatives are coordinated and articulated using BIM in the pre-construction stage. The focus is on an integrated supply chain management, which is a rather infant new research area in BIM applications. Few research attempts have been demonstrated by using BIM in this perspective. Yet, a full integration of BIM with other technologies has yet to be realized. Geographic Information System (GIS) and Global Positioning System (GPS) are the technologies that could improve the supply chain management. The need for real-time tracking information and on-time delivery is obvious, which could cope with the long-lead? supply chain in infrastructure projects. Besides, the integration should extend itself to the construction stage where all of the deviations and changes could be managed and updated in the 4D and 5D BIM. Meanwhile, Radio Frequency Identification technology (RFID) has also been integrated in the construction site for tracking purpose. An automated system of integrating BIM, GIS, GPS and RFID will be the future for creating a smart construction site [4]. As a result, this full integration between the technologies and project phases would
promote waste minimisation and improve the overall project performance.

BIM has been a powerful technology in the construction stage due to its visualisation ability and information traceability. The scope of BIM in construction stage requires an improved coordination in the current practices, particularly for construction processes and engineering management system. Augmented Reality (AR) technology works and integrates well with BIM [5]. A recent study demonstrated a successful integration of BIM and AR for Liquefied Natural Gas projects (LNG) [6]. The integration makes BIM more effective in identifying the interdependence and complex tasks in the construction process. Onsite activities will become more manageable as all the required information of the activities and components of the models can be visualised and retrieved in the real environments [6].

In the maintenance stage, an integration BIM and asset management has been demonstrated in both building [7, 8], and oil and gas industries [9]. There should be more investigations about the incorporation of BIM into asset management. It is about increasing asset value after construction via an effective coordination and maintenance of the physical asset environment of the project. The integration is able to shorten the learning curve in facilitating the services or components of the project. On top of that, the development of BIM model should be integrated with asset management at the earliest stage as possible in the design stage of the project life cycle. It will ensure the full benefits of BIM in asset management to accommodate the complicated human needs in the physical environments.

The BIM integrations would benefit each phase of the project life cycle. Yet, the management of the BIM integrations is another important step in closing the loop of the challenges and trends in BIM in the future. Different types of management and contract administration are required for different levels of BIM adoption and maturity. Some managerial approaches like Lean concepts [10] and Integrated Project Delivery [11] should be adopted in the overall management and contract administration for BIM. Moreover, some tools of information technology could be considered in streamlining the management process, such as database system or data warehousing [10, 11], cloud computing [12], etc. Thus, the project is expected to secure a better results and performance in the project.

3 Discussions and Conclusions

Australian economy is in transition for a better sustainability. Construction and resources sectors play a critical role in boosting economic growth. Productivity is imperative for the economic stability. Nevertheless, the full benefits of BIM have yet to be achieved. To unleash its true potential, it is recommended that intensive research and development should be carried out to bridge the gap between the academic evidence and industrial practice. Project Echo [4, 15] is one of the practical examples, where it serves as a collaborative research platform for the practitioners and academic researchers. It is envisaged that BIM will become either a culture or a technological breakthrough in transforming the conventional construction practices in the near future.

References


Abstract -
Access to sustainable, affordable and secure energy is one of the major Australian strategic priorities to maintain and improve the health of Australians, sustain economic growth, and to mitigate climate change. Australia is investing in clean, efficient, reliable energy systems to secure a prosperous and, environmentally sustainable future. In addition, exploring the options to ensure energy security by diversification of energy sources is an important aspect for securing stability in meeting and delivering the future energy requirements of different industry sectors.

This paper discusses options to manage the production of electricity in Australia using available Australian resources while maintaining international competitiveness.

Keywords -
Future energy; sustainable energy; distributed generation; centralised power generation; Post combustion Capture; ISARC 2014

1 Introduction
Ensuring the availability of clean, abundant and affordable energy will play a key role in developing economic prosperity and enhanced environmental quality. Coal is a key source of energy for power generation. Anticipated growth in energy demand is likely to extend the use of coal for electricity generation. However, moving towards a low-carbon economy requires industry to use coal more efficiently and to reduce its environmental footprint.

It is now well accepted that emissions from fossil fuels burning should be limited because of their contribution to local and regional pollution in addition to their contribution to global warming. Stringent environmental regulations and the proposed future limits on greenhouse gas emissions from fossil-fuelled power plants have played a key role in initiating new research opportunities to overcome these difficulties.

Currently, efforts are being made by major research organisations to develop and demonstrate the availability of ultraclean, efficient, affordable fossil fuel energy technology that can be used to meet future requirements for energy production and use. If technological success is achieved it will be likely for coal and natural gas to remain major energy sources for power generation and will continue to meet the growing energy demand because of their abundant supply and relatively low cost when compared to other energy sources.

Carbon Capture and Storage (CCS) has been developed as a ready option to curb greenhouse gas emissions from industrial sources. The technology is based on (i) cleaning the flue gas stream then separating and capturing CO₂, (ii) compression and transport of CO₂, (iii) underground storage of CO₂. The wide deployment of CCS will reduce CO₂ emissions from industrial sources while ensuring that coal will remain as a relatively cheap option to power a sustainable economy.

Post combustion capture (PCC) of CO₂ using amine solvent is a readily available technology to capture CO₂ from coal-fired power plants. The process would be retrofitted downstream of the cleaned flue gas. The aqueous amine solvent will undergo a reversible chemical reaction with CO₂ and will selectively capture CO₂ from the flue gas stream. Under elevated temperature the CO₂ will be released and the lean solvent will be recycled back to absorb more CO₂ from the flue gas stream of the power plant. The flue gas exiting the power plant to enter the PCC includes different compounds such as sulphur oxides (SOₓ), nitrogen oxides (NOₓ), oxygen (O₂), particulates and other compounds. Some of these compounds interact with the solvent to produce different by-products which may affect the solvent performance. There is potential for selected by-products to leave the plant and undergo additional chemical reactions in the atmosphere. The full understanding of these chemical transformations is needed for plant operating approval where emissions from the plant should not exceed defined limits.

Over the last seven years CSIRO has been engaged in improving knowledge associated with the operation of PCC plants, development of new solvents to resist degradation, elucidating the chemistry of amines in the process and their fate in the atmosphere, and acquiring information to carry out environmental assessment for future PCC deployment in Australia and internationally.
2 Low-Carbon Electricity

There is an urgent need to explore practical options that can help Australia to use available energy resources for electricity production as we move forward to a reliable and sustainable energy future. In Australia and globally, electricity generation from fossil fuelled plants has been subject to stringent environmental regulations because of concerns about their effects on air, water quality and their contribution to greenhouse gas emissions. Centralised large fossil fuels power plants and decentralised energy generation from renewable sources will present the new mix of electricity generation in Australia. The new generation mix represents an attempt to shift from conventional centralised fossil fuels generation to more efficient emission-free or near zero emissions from a variety of energy sources.

How well can we plan and implement sustainable energy systems in Australia?

For centralised power plants, it is crucial to assess the use of high efficiency centralised large fossil-fuelled power plants equipped with emissions reduction technologies and potentially carbon capture systems. A successful sustainable centralised power generation system will be measured by meeting the following targets: increasing efficiency, significantly reducing emissions while keeping costs low for consumers.

How to deploy sustainable decentralised power generation in Australian urban areas?

The most decentralised electricity generation will be produced from a variety of renewable sources including gas, solar, biomass and natural gas and winds. In contrast to conventional generation, decentralised generation consists of smaller scale generation sources interconnected to a utility’s distribution grid. It is expected that the new energy mix system will exhibit a tremendous complexity manifested by its diversity, dynamic behaviour, space and power output. The anticipated energy system would have to ensure the security of supply under different operating modes and be able to demonstrate the safe and environmentally sound operation of the selected system.

What Environmental Regulations and policies are needed?

Thermal power plants are currently the major source of electricity in Australia. Incorporating a large fraction of electricity generation from other sources such as concentrating solar power systems, photovoltaic and others will require changes to many existing regulatory aspects and policies to take account of the new energy sources.

2.1 Decentralised Power Generation

Decentralised energy generation has the potential to supply part of the needed electricity using multiple stationary power generators installed close to end-users. The approach is referred to as distributed generation (DG), which shifts the design and layout of power generation and distribution systems to reduce pollution by increasing efficiency and promoting the inclusion of renewable energy sources. The DG model operates on more efficient smaller-scale than traditional systems.

The DG configuration does not require transmission, as sources are connected directly to the lower-voltage distribution network. Advanced technological achievements in addition to more stringent environmental regulations have resulted in great interest for the deployment of distributed generation systems. The integration of different decentralised energy generation sources with existing centralised power plants requires extensive management to secure the electricity supply to customers.

The distributed generators technologies that are likely to be implemented in major Australian cities include but are not limited to natural gas fired combustion turbines, natural gas fired reciprocating internal combustion engines, natural gas fired micro-turbines, solar photovoltaics and wind turbines. The mix may also include diesel and petroleum distillate fuelled units that can be used as back-up generators.

Economic, regulatory and environmental requirements will play major roles in the selection of the mix of DG technologies to be installed in a region. It is expected that some candidate generator technologies will emit different or more pollutants than central power plants. Many of the proposed technologies that might be adopted within urban areas would result in displaced emissions. It is also anticipated that all electricity generators may experience some degradation in efficiency performance and consequent degradation in terms of pollutant emissions performance. A systematic methodology for the development of realistic distributed generation deployment scenarios is needed. This methodology will take into consideration all characteristics and factors needed to minimise the local concentration profiles of pollutants expected from operating these generators with the least environmental and health impacts.

To assess the environmental benefits of DG in an airshed it is important to consider the mix of technologies employed, total percentage of DG in the region, emission profiles, and the use of an appropriate dispersion model in order to quantify how the deployment of the new energy mix can result in improved air quality by increasing efficiency and
reducing pollutant emissions.

A study carried out by Azzi et. al. 2012 [1], has investigated the air quality impacts of a scenario in which the energy generation paradigm moves from large centralised energy generation (such as coal-fired power station) to the wide spread deployment of small power units which are located close to the point of use. The study considers a scenario in which DG is deployed across the Greater Sydney Region (GSR, includes Newcastle to the north and Wollongong to the South).

It was found that emissions of NO\textsubscript{x}, VOC and CO associated with the DG deployment caused a small increase in domain-wide mortality (changing by less than 0.01 %). The DG deployment contributed an additional 1% to the existing pollution burden in the GSR. When ranking pollutants by degree of impact, nitrogen dioxide was found to be the limiting pollutant, most likely because the pathway between emitted NO\textsubscript{X} and NO\textsubscript{2} involves fewer steps than the generation of ozone or secondary PM\textsubscript{2.5}.

This study demonstrates how decision makers and stakeholders in NSW can plan for the future penetration of combustion based distributed generation systems using modelling tools and thus investigate optimal approaches to meet the State’s growing electricity demands while minimizing health risks.

2.2 Centralised Power Generation

Coal remains the largest provider of electricity generation. Over 75% of Australia’s electricity is still generated by coal-fired power plants. The traditional Australian electrical power system model consists of centrally-located power plants, transmission lines and distribution networks. While renewable energy sources have started to make inroads in the production of electricity in Australia, centralised coal-fired power plants typically located in remote areas will continue to be the major sources for electricity generation. These power plants are known for their low efficiency, high emissions and large land disturbance. In addition, emissions of different pollutants with known environmental and health concerns have been associated with the operation of these plants. The major challenges to the continued use of coal are related to its environmental impact affected by emissions from coal combustion.

Emissions from fossil fuels combustion depend on the fuels used and the type of power plant that is using the fuels. Pollutants generated can include NO\textsubscript{x}, SO\textsubscript{x}, PM, and CO\textsubscript{2}. There are other compounds expected to be produced resulting from incomplete combustion of fuels such as volatile organic compounds, unburned carbon and trace elements. It is recognised that natural gas is the least polluter of all fossil fuels resulting mainly in emissions of NO\textsubscript{x}, CO\textsubscript{2} and very small amounts of other pollutants.

Over the last two decades, regulatory agencies have forced utility companies to meet more stringent emissions standards and improve efficiency. Many of the proposed solutions and policies to reduce emissions from these plants were achieved by technological improvements. The installation of pollution control and prevention equipments in order to reduce the amount of pollutants released to the atmosphere and meet environmental guidelines has helped the industry to operate under regulated limits of emissions.

The average efficiency of coal-fired power plants varies from under 30% to 45% (LHV,net). These differences arise from diverse factors including the age of plant, steam conditions and coal quality. Over 50% of operating plants use subcritical technology. Pulverized coal combustion in supercritical, ultra-supercritical steam cycle, and Integrated Coal gasification Combined gas turbine-steam Cycle (IGCC) technologies are proven efficient technologies which produce fewer GHG emissions per unit of power. These plants are being used in Asia, Europe and in the US. They provide more energy efficient systems for power generation from coal where more electricity is generated from the same amount of coal used with less emission. Different pollution control technologies are installed on modern coal plants to reduce emissions of atmospheric pollutants.

Reducing CO\textsubscript{2} emissions from coal-fired power plants would have a significant impact on climate change. Carbon capture and sequestration is a potential option to keep fossil fuels in the electricity mix. Incorporating carbon capture technologies into coal-fired power plants will reduce greenhouse gas (GHG) emissions from coal and also other pollutants emissions. Chemical absorption with amines is presently the only commercially available technology for industrial application. The CO\textsubscript{2} is first captured from the exhaust gas stream in an absorption tower. The absorbed CO\textsubscript{2} is then heated with steam to strip the CO\textsubscript{2} from the amine solution. The regenerated amine is sent back to the absorber. The recovered CO\textsubscript{2} is cooled, dried, and then compressed to a supercritical fluid. At this stage the fluid can be piped to sequestration.

In a study carried out by CSIRO for the Australian National Low Emissions Coal Research and Development (ANLEC R&D) [2] it was found that Post Combustion Capture (PCC) technology to remove CO\textsubscript{2} from the flue gas of coal-fired power stations using an amine-based solvent has the potential to also reduce the emissions of nitrogen dioxides (NO\textsubscript{x}) particulate matter (PM), and sulfur dioxide (SO\textsubscript{x}), and hence lead to improved air quality. It is important to note that the retro-fitting of a PCC plant has the potential to impact
on both the traditional air pollutants in addition to providing a source of new pollutants based on the emission and degradation products of monoethanolamine (MEA).

New generation sources have been introduced and pollution mitigation devices installed, but air quality impacts from electricity production still exist. Because of this, traditional systems have remained largely unchanged, with the exception of simple technological upgrades.

2.3 Environmental Regulations and policies

Concerns about the environmental impacts of fossil fuel production, processing and utilisation have pushed regulators to introduce regulations to reduce smog formation, air pollution, acid rain and other air pollution compounds. More recently, the reduction of GHG emissions has been the main focus because of its relationship with climate change concern.

These regulations are not the only environmental regulatory challenges facing the electric utility industry. Other new stringent regulations related to coal ash management and disposal and water discharge are being promulgated. Compliance with these new regulations will require new control devices and investment. Coal power utilities will be confronted unprecedented costs and management challenges to meet these new regulations.

The installation of different pollution control devices to meet existing and future environmental guidelines requires a cumulative assessment to avoid any counterproductive effects of these devices.

Decreasing reliance on fossil fuels by incorporating different energy sources from renewables and from natural gas will require different environmental regulations and policies to that designed for the existing thermal power plants. Future regulatory considerations for future energy mix technologies may include the future cost of carbon and increased control requirements for emissions.

References


Abstract –

Conventional construction reached its limits. Numerous construction projects, such as Cologne subway site causing a billion euro collapse of a library containing cultural heritage assets, the x times cost overruns of Elbphilharmony Hamburg and the x tens of thousands quality defects of Berlin Brandenburg Airport indicate the necessity of an overdue transformation of the whole sector. Automation and robotics in construction could achieve what other manufacturing and service industries have already successfully demonstrated. This article shows how robot oriented design, engineering, management and technologies for the construction and the building sector in a form of a five volume construction robotics handbook series would elevate the industry onto a proper and dignified level of quality and performance, which matches the level of the rest of the manufacturing and non-manufacturing sectors.

Keywords –


1 Introduction

Welfare and culture in any society can be achieved through sufficient personal income. One nation must be efficient and achieve further productivity and economic growth. Where there are no natural resources to be exploited and sold, high income can only be accomplished by sophisticated technology. Affordable and efficient socio-economical and socio-technical processes are further required by the demographic change. Half of the total investment is allocated in the built environment, infrastructures and facilities, signifying the strategic importance of the construction sector.

Productivity in the construction industry has been declining for decades world-wide. The construction industry has one of the lowest capital stocks compared to other industries, as well as low capital intensity. Furthermore, construction defects, improper working conditions and low interest in the construction field for younger generations, as well as the tremendous consumption of raw materials and energy by the construction process and building products, state challenges for which conventional construction and the architecture industry currently do not have solutions for it. In high wage countries, the natural ageing of societies will continuously aggravate the situation by reducing our human capital, as well as the ability for change and economic growth. Börsch-Supan, a German macroeconomist, predicts a solution for augmenting productivity and economic wealth, therefore, predominantly in supplementing human capital by capital intensity, non-linear advances in machine technology and productivity (Börsch-Supan et al, 2009).

General manufacturing under the notion of “Industry 4.0” (see for example Jopp, 2013) or “Cognitive Factory” (see for example Zäh et al., 2009) as hyper-flexible and intensively automated manufacturing systems (also considered already as the 4th industrial revolution) in which highly autonomous, flexible and distributed but still networked automation and robot systems cooperate together in producing in a nearly real-time manner individualized and complex products with incredible productivity could offer in construction industry higher productivity and needed change, that has been stagnating for decades. Innovation in construction industry occurs extremely slowly. The reason for that, from one side is caused by the characteristics of the products and their complexity, long life-cycle, diversity, dimensions, materiality and fixed site nature. Additionally, it has its roots in the low R&D budgets and the reluctance to new strategies and technologies. Opposed to the marginal improvements in conventional construction, since the 1970s scientists, R&D departments and innovative companies supported by universities, associations and governmental institutions, pursued on a consistent manner, a new set of technologies and processes which will change the whole course and idea of construction in a fundamental way and which can be summarized under the term “Automated Construction”. (Figure 1, Table 1) Represent the idea as a whole.
In contrast to conventional construction, Automated Construction is capital intensive and machine centred while, being potentially limitless in performance and capable of real-time manufacturing. As Automated Construction necessitates a complementary and also disruptive change in the whole industry (products, processes, organisation, management, stakeholders, business models, etc.) in order to be able to fully unfold its potential, it can be considered as a rather complex type of innovation or change.

Changes of such complexity takes time, sometimes decades. However, now after nearly 40 years of technical development and experimenting in the field, the result is increased activity within companies, research institutes, associations and governmental institutes. This indicates that this new trend gains in acknowledgement, the adoption of future technologies and finally being able to head towards a growth phase.

Figure 1. Automation and robot technology becomes ubiquitous and step by step pervades our life on earth. In our volume series the focus is set to, but not limited, to Automated/Robotic Construction.
<table>
<thead>
<tr>
<th>Thematic Field</th>
<th>Systems and Approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated/Robotic Infrastructure Production</td>
<td>1 Automated road construction</td>
</tr>
<tr>
<td></td>
<td>2 Automated tunnelling (i.e. by TBMs)</td>
</tr>
<tr>
<td></td>
<td>3 Automated bridge construction</td>
</tr>
<tr>
<td></td>
<td>4 Automated con- and deconstruction of dams, power plants, etc.</td>
</tr>
<tr>
<td></td>
<td>5 Automated mining</td>
</tr>
<tr>
<td></td>
<td>6 Automated container port</td>
</tr>
<tr>
<td>Automated/Robotic Transportation Systems</td>
<td>7 Autonomous cars</td>
</tr>
<tr>
<td></td>
<td>8 Autonomous public transport (U-Bahn, Train, etc.)</td>
</tr>
<tr>
<td></td>
<td>9 Autonomous air travel</td>
</tr>
<tr>
<td></td>
<td>10 Automated logistics</td>
</tr>
<tr>
<td></td>
<td>11 Advanced micro-mobility (i.e. Cyberdyne’s HAL or Toyota’s i-Real)</td>
</tr>
<tr>
<td>Automated/Robotic Construction</td>
<td>12 Automated construction of vertically oriented buildings</td>
</tr>
<tr>
<td></td>
<td>13 Automated construction of horizontally oriented buildings</td>
</tr>
<tr>
<td></td>
<td>14 Housing production</td>
</tr>
<tr>
<td></td>
<td>15 Novel construction Markets accessible through automated/robotic construction:</td>
</tr>
<tr>
<td></td>
<td>16 Automated Building Servicing and Maintenance</td>
</tr>
<tr>
<td></td>
<td>17 Automated De-construction and re-customization</td>
</tr>
<tr>
<td>Automated/Robotic Environments</td>
<td>18 Home &amp; Office Automation</td>
</tr>
<tr>
<td></td>
<td>19 Assistance Technologies and Human-Ambient-Technologies</td>
</tr>
<tr>
<td></td>
<td>20 Networked Production Facilities and Supply Networks</td>
</tr>
<tr>
<td></td>
<td>21 Intelligent Energy Generation and Distribution</td>
</tr>
<tr>
<td></td>
<td>22 Service and household robotics</td>
</tr>
<tr>
<td>Automated/Robotic Farming and Food Production</td>
<td>23 Computer Aided/Robotic Farming</td>
</tr>
<tr>
<td></td>
<td>24 Robotic Milking Stanchions</td>
</tr>
<tr>
<td></td>
<td>25 Automated Food Production Facilities</td>
</tr>
<tr>
<td></td>
<td>26 Customized Food</td>
</tr>
<tr>
<td>Automated Robotic Town Management</td>
<td>27 Smart Grids</td>
</tr>
<tr>
<td></td>
<td>28 Automated/Robotic Traffic Control</td>
</tr>
<tr>
<td></td>
<td>29 Automated/Robotic Infrastructure Inspection and Maintenance</td>
</tr>
<tr>
<td></td>
<td>30 Automated Supply Management (water, gas, goods, food, etc.)</td>
</tr>
<tr>
<td>General Manufacturing Industry</td>
<td>31 Digital/Cognitive Factories</td>
</tr>
<tr>
<td></td>
<td>32 Mass Customization</td>
</tr>
<tr>
<td></td>
<td>33 Mini Factories, Cloud Manufacturing</td>
</tr>
<tr>
<td></td>
<td>34 Cellular Logistics</td>
</tr>
</tbody>
</table>
2 Transformation towards Advanced Construction and Building Technology: Robot Technology Becomes Ubiquitous

“The end of the information age will coincide with the beginning of the robot age. However, we will not soon see a world in which humans and androids walk the streets together, like in movies or cartoons; instead, information technology and robotics will gradually fuse so that people will likely only notice when robot technology is already in use in various locations.” Ishiguro (2012)

Robotic systems are an advanced classification of machines characterized by their capabilities, such as re-programmability, autonomy, flexibility and situational awareness. By observing advances within robot technology, it is feasible to predict that robot technology will experience a similar development as did the personal computer during the nineties. Experts and masterminds, as for example Bill Gates, announce the era of robotics and estimate that robotics of our everyday life. The South Korean government recently announced that it supports heavily the development of robot technology having the goal to establish at least one robot system in each household. In 1961 Joe Englberger already wondered, whether relating robotic technologies to only product manufacturing applications makes any sense. “The biggest market will be service robots” (Englberger, 1989) asserted Englberger, who started the industrial robotics era, when his firm (Unimation) delivered GM's first robot.

Currently robots are becoming more user friendly, less expensive, task adaptable, smaller, more widely distributed and seamlessly integrated into work processes and devices. Individuals today are able to acquire modular kits of open source robot hardware, and interface robot systems directly to a computer via USB devices. The modern options make transitions between self-contained devices, such as the classical 6-DOF industrial robots, much more simplistic. Currently these types of machines, automation systems and robotic technology, are advancing in such a way that the differences are becoming seamless. There are examples of Robots which have become no longer visible. The complexity of Robots consists of a network of interconnected, distributed and sometimes invisible sensor and actuator systems (including mobile phones or appliances). This means that individual devices functioning as machines can cooperate as a network to manipulate or achieve goals (by manipulating energy flows as smart grids do) autonomously as a robot system.

With the continuous evolution within the field of Robotic Research, new technical capabilities (modularity, lightweight concepts, wearable robot technology, and social robot technology) have been explored and combined with existing manipulation oriented automation, and robot technology.

Over time, the ability of robot systems has grown, allowing them to work more and more in comparably unstructured environments, to those in which human beings operate. This evolution leads to the fact that robot technology, apart from the classical manufacturing industries, can now be introduced and be

---

Figure 2. ACBT Covers following thematic fields which are outlined in 5 complementary Volumes
deployed in numerous fields, such as aircraft production, farming, the construction industry and health care sector. New capabilities not only extend the application area of robotics, but also allow robot systems to be used in factory-like environments. This provides robot systems the ability to become more dexterous and adaptive, sequentially building a basis for the production of customized and individual products. Faster product life cycles, and the increasing demand for flexible production systems able to produce customized products, has also lead to refined modularization, plug-and play capabilities of subsystems, open source approaches, concepts for self-configuring and self-organizing robot systems and cellular robotic logistics.

In short it can be said that automation and robot technology becomes ubiquitous, and step by step pervades our life on earth lists thematic fields into which automation and robotics is currently advancing as well as approaches and systems within those fields. It shows that nearly all field of professional and private life are subject to the pervasion of robot technology. It can be assumed that with full establishment of automation and robotic in some thematic fields will follow in order to exploit complementarities and synergies. (Figure 2).

3 New Design and Management and Movement by Robot Oriented Design (ROD)

New design, innovation and management methodologies that are keys for the realization and implementation of the advanced concepts and technologies are explained by Robot oriented design in Volume 1, industrialized customization in Volume 2, Single Task Construction Robots in Volume 3, and automated on-site factories in Volume 4 and ambient integrated robotics in Volume 5. Robot Oriented Design and Management enable the efficient deployment of advanced construction and building technology. It is concerned with the co-adaptation of construction products, processes, organisation and management and automated or robotic technology, so that the use of such technology becomes applicable, simpler or more efficient. It is also concerned with technology and innovation management methodologies and the generation of life-cycle oriented views related to the use of advanced technologies in construction and building context. The concept of ROD was first introduced in 1988, in Japan by Bock (1988) and served later as the basis for automated construction and other robot-based construction sites around the world. (Figure 3) It was developed for improving the construction sector and adjusting conventional construction processes and component design to the needs of the novel tools.

In Volume 1, technologies relevant to understanding approaches outlined in Volumes 2, 3, 4 and 5 and the evaluation of related technical, organizational and economical concepts and parameters are introduced. Volume 1 also contains the core glossary, which provides the explanation of concepts, terms and acronyms relevant for all 5 volumes.

4 Ki-NEW-Matics in Industrialized Customization – Automation and Robot Technology in Component, Module and Building Prefabrication

For industrialized customization, concepts, technologies and developments in the field of Building Component Manufacturing (BCM) based on concrete, brickwork, wood, and steel as building materials and Large-scale Prefabrication (LSP) holding the potential to deliver complex components and products are introduced and discussed. BCM refers to the transformation of parts and low-level components into higher level components by highly mechanized, automated or robot supported industrial settings.

The definitions of components are interpreted differently by different industries and even by individual companies. However, the definitions of components share a common element, that they are a
more or less a complex combination of individual pre-existing parts and/or lower level components. Pure BCM can be distinguished from the transformation of raw material, into parts (production of bricks or simple concrete blocks).

Also, component manufacturing can be distinguished from the manufacturing of highly complex modules or units. BCM is clearly distinguished from the manufacturing of high-level building blocks, such as building modules (prefabricated bath modules or assistance modules which can also be referred to as building subsystems) and building units (such as the LSP of fully finished three-dimensional building sections, like Sekisui Heim, Toyota Home, Panahome, Misawa Hybrid). For highly automated LSP, according to the OEM model, component manufacturers represent Tier-1 or Tier-2 suppliers. Tier-1 suppliers would deliver components directly to companies such as Sekisui Heim, whereas, Tier 2 suppliers would supply the suppliers of the bath or assistance units. For Automated/Robotic On-site Factories, component manufacturers again represent Tier-1 or Tier-2 suppliers. BCM and LSP industry can reduce on-site complexity and build the supply backbone in an OEM-like industry structure, which can be considered also as a prerequisite for the successful implementation of Automated/Robotic On-site Factories. (Figure 4)

5 Ki-NEW-Matics in Single Task Construction Robots (STCRs)

After the first experiments in large-scale industrialized, automated and robotized pre-fabrication of system houses were conducted successfully in Japan, and the first products (e.g. Sekisui M1) also proved successful in the market, the main contractor Shimizu, 1975 in Tokyo, set up a research group for construction robots.

The goal was now no longer the mere shifting of complexity into a SE as in LSP, but the development and deployment of systems, which were able to be used locally on the construction site to create structures and buildings. The focus initially was set on simple systems in the form of STCRs that can execute a single, specific construction task in repetitive manner. The fact that STCRs operated task specific made them initially highly flexible (they could be used along with conventional work processes and did not necessitate that the whole site is structured and automated), but also represented a major weakness. As they were, in most cases, not integrated with upstream and downstream processes, which demanded safety measurements and hindered parallel execution of work tasks by human workers in the area where they were operated, productivity gains were often equalized. Above all, the set-up of the robots on-site (equipment, programming) was time consuming and demanded new skills. In addition, the relocation of the systems on-site was, in many cases, complex and time consuming.
The evaluation of the first generations of developed and deployed STCRs and the identification of the above mentioned problems led step-by-step from 1985 onwards to the first concepts for integrated automated/robotics sites. Concepts for Integrated automated construction sites integrated STCRs and other elementary technology as subsystems into and SE set up on the construction site. The development of STCRs, elementary technology, and a concept for structuring on-site environments by ROD was analysed and supported by Bock (Figure 5) from 1984 to 1989 (Bock, 1988). As the development of STCRs, elementary technology and alternatives created the basis for, and paved the way for the realization of integrated automated/robotic construction sites from the 1990s onwards. Furthermore, as the development of STCRs, parallel to or as subsystems of integrated automated construction sites continued up to today, 140 STCRs have been identified, analysed and categorized. In this volume, the conceptual and technological reorientation towards integrated automated construction sites initiated by WASCOR Group, WASeda Construction Robot Group and joined researchers of all major Japanese construction firms and equipment is shown.

6 Ki-NEW-Matics in Automated/robotic On-Site Factories for Construction and Deconstruction

The approach of setting up SEs on the construction site in the form of Automated/Robotic On-site Factories can be considered as a logic taking further of BCM, LSP (Volume 2) and STCRs (Volume 3) technology. In Volume 4, all worldwide conducted approaches following the above set direction to an on-site factory approach are outlined. 30 different systems are identified, resulting in an application of Automated/Robotic On-site Factory technology about 60 times. The outline is split into a more technical part and a part which focuses on parameters related productivity, efficiency and economic performance. All systems were outlined systematically and based on the same frameworks and 13 categories were set up (ten categories for construction and three categories for deconstruction). One of the main ideas for setting up automated on-site factories was to integrate stand alone or STCR technology into structured on-site environments to networked machine systems and to improve organization, integration and material flow on the construction site, apart from the possibility to off-site manufacture components. Automated On-Site factories show for which building typologies Automated/Robotic On-site Factories are an applicable approach, and how and to extent those systems can be made technologically flexible in order to be able manufacture a variety of different buildings (products) on the basis of industrialized, automated and flow-line like stable factory processes. Also, it is shown how, in contrast to the STCR approach, the approach of setting up Automated/Robotic On-site Factories is capable of achieving a performance multiplication, which usually accompanies the switch from arts and crafts based manufacturing, to machine based manufacturing.

7 Ki-NEW-Matics Example 5: Ambient Integrated Robotics - Automation and Robot technology in building maintenance, assistance and service

For a long time the focus of innovation in the area of automation and robotics in construction has been laid on the industrialized and rationalized off-site or on site construction as well as on related processes, building systems, management and logistic tools and high-tech construction and prefabrication robotic equipment.
Still today, the complexity of buildings continues to rise rapidly due to new paradigms as Ubiquitous Computing, the demand for energy efficiency and emerging assistance technologies. Buildings become integrated with a multitude of new sub-systems and extend their performance to areas which have formerly not been accounted as being part of construction and building industry, yet, which now gradually merge with our built environment. With the integration of Microsystems Technology into buildings, and due to the tendency towards more and more user integration (Piller, 2006), buildings become not only more intelligent, but they can be much more personalized to the inhabitants needs and could further serve as platforms for a multitude of continuous and commercial services. These changes could have a tremendous impact on the whole value chain and are likely to transform building structures, construction technologies and business models. (Figure 7).

8 Conclusion

The five examples of Ki-NEW-Matics are based on the Greek meaning of motion and movement. The existing problems and defects of the construction sector can be successfully addressed by introducing advanced construction and building technology based on automation, robotics and services. This new movement will lead to an innovation leap.

References

[4] Linner, T. (2013) ‘Automated and Robotic Construction: Integrated Automated Construction Sites’ [Dr.-Ing. Dissertation], Technische Universität München – in particular chapters 1, 3 and 4 in this Volume (Volume 3) are based on chapters 4 the thesis; furthermore the classification system for the STCRs has been adopted from the thesis and expanded and detailed in this volume (Volume 3).
Towards Fully Automated Tunnel Inspection: A Survey and Future Trends

C. Balaguer, R. Montero, J. G. Victores, S. Martínez, A. Jardón
RoboticsLab, University Carlos III of Madrid, Spain
E-mail: balaguer@ing.uc3m.es, romonter@pa.uc3m.es, jcgvicto@ing.uc3m.es, scasa@ing.uc3m.es, ajardon@ing.uc3m.es

Abstract –
Tunnels environments are characterized by dust, humidity, and absence of natural light. Artificial and natural impacts, change in load criteria, or the simple effect of ageing, make tunnels require inspection and maintenance. These operations are commonly performed by human workers taking time and expertise without guarantee quality control. Robotic tunnel inspection and maintenance (RTIM) introduces high productivity, quality and repetitiveness. This paper describes the current trends in the subject, and introduces new technologies such as scenario modeling, robotic platforms, image and ultrasound sensors, control algorithms and decision making strategies. Additionally, the result of several recent and ongoing projects will be presented.

Keywords - Robotics, Automation, Inspection, Maintenance, Tunnels, IAARC

1 Introduction

One of the greatest challenges engineers face is the inspection, assessment, maintenance and safe operation of the existing civil infrastructure. This includes large-scale constructs such as tunnels, bridges, roads and pipelines. In the case of tunnels (water supply, metro, railway, road, etc.), they have increased in both total length and number, and will continue to do so. Furthermore, some tunnels still in service were completed over 50 years ago, with the existing construction and materials technology.

Only in Japan in 2006, the number of active tunnels was up to 9000 [1], with tunnels such as the Seikan Tunnel which is 54 km long and partially below the seabed [2]. Figure 1 shows the evolution of Japanese tunnels in terms of number and length until 2006.

Tunnels progressively deteriorate due to ageing, environmental factors, increased loading, change in use, damages caused by human/natural factors, inadequate or poor maintenance, and deferred repairs. Unfortunately, several incidents related to the structural condition of tunnels have taken place, such as the Big Dig ceiling collapse in 2006 in Boston [3], or the Sasago Tunnel collapse in 2012 in Tokyo [4].

Figure 1. Changes in the number and total length of road tunnels in Japan (2006) [1]

These examples highlight the need of automated, cost-effective and exhaustive inspection of tunnels that prevents such disasters. In this work, we present current tendencies and future trends within this area.

1.1 Tunnel Defects

The first aspect of inspection that must be defined is the related to the types of defects that may affect tunnels. Identifying these defects is crucial for performing a successful inspection, verifying the state of a tunnel, and performing maintenance if required.
The following list of common defects in tunnels is based upon the *TOMIE Manual* [5], created by the Federal Highway Administration (FHWA).

- Concrete structures: scaling, cracking (traverse, longitudinal, horizontal, vertical, diagonal, pattern, d-cracks, random), spalling, joint spall, pop-outs, efflorescence, staining, delamination, honey-comb, leakage
- Steel structures: corrosion, cracks, buckles and kinks, leakage, protective layer fail
- Masonry structures: masonry units (displaced, cracked, broken, crushed, or missing), mortar, shape, alignment, leakage
- Timber structures: decay, insects, checks/splits, fire damage, hollow area, leakage

The walls of most tunnels are made of concrete, though these walls may contain finishes such as ceramic tiles or metal panels. In most cases, the typical defects found in a tunnel are cracks, spalling and efflorescence/leakage [6]. Examples of this type of defects are displayed in Figure 2. If the walls are covered by a finish, the condition of such walls is generally defined by the deficiencies of the finish on the wall surface. An analysis of the causes of these common defects in tunnels can be found in the work by C. C. Xia et al. [7].

**Figure 2.** Examples of spalling (left) and crack with efflorescence (right) [5]

### 1.2 Tunnel Inspection Methods

The purpose of inspection is to check if a structure that has been functional for years is still safe or not. Furthermore, it is desirable to do this without creating any negative effect on the structure or component, and this is why the non-destructive inspection (NDI) methods [5][8] are far more commonly used than destructive methods. As said before, the most common structural material in tunnels is concrete, thus the following inspection processes are usually applied to concrete tunnels. NDI methods in structures can be divided in visual, strength-based, sonic and ultrasonic, magnetic, electrical, thermography, radar, radiography, and endoscopy methods.

#### 1.2.1 Visual Methods

Visual testing is probably the most important of all non-destructive tests. It can often provide valuable information to the well trained eye. Visual features may be related to workmanship, structural serviceability, and material deterioration, and it is particularly important for the engineer to be able to differentiate between the various signs of distress which may be encountered. Information can be gathered from visual inspection to give a preliminary indication of the condition of a structure and allow the formulation of a subsequent testing program.

#### 1.2.2 Strength Based Methods

Rebound and penetration tests measure the surface hardness of materials and provides an estimation of surface compressive strength, uniformity and quality of the structure. Examples include the Schmidt Hammer [8][9] (rebound), the Windsor Probe [8][10] (penetrating), Flat Jack Testing [11][12] (applied to masonry), or methods without contact [13].

#### 1.2.3 Sonic and Ultrasonic Methods

In sonic methods, also known as impact-echo tests, hammer blows create impulses, and the time of travel of these sonic pulses is measured with pickups placed on the wall [14][15][16], as can be seen in Figure 3. The time of travel is related to the modulus of elasticity and, hence, the strength. Sometimes chain drags, sounding rods or standard hammers are used for detecting delamination on horizontal surfaces if the inspector has experience in detecting hollow sounds.

**Figure 3.** Inspection using an impact hammer [17]
Ultrasonic devices are normally used by measuring the velocity in the material of a pulse generated by a piezoelectric transducer [18][19][20]. The pulse velocity depends on the composition and maturity of the structural material and its elastic properties. The relationship to strength depends on several other properties and is best determined experimentally [21].

1.2.4 Magnetic Methods

Magnetic methods are used to determine the position of reinforcements and are not techniques for detecting defects or deterioration directly, but the fact that inadequate cover is often associated with corrosion-induced deterioration indicates that a method for locating the reinforcing bars can be important in corrosion control. Examples of these methods are the Magnetic Flux Leakage method [22][23] or the Magnetic Field Disturbance method [5].

1.2.5 Electrical Methods

Electrical methods for inspection of tunnel components include resistance and potential measurements [24][25][26][27]. Electrical resistance has been used for measuring the permeability of deck seal coats and involves measuring the resistance between the reinforcing steel and surface, while electrical potential differences are caused by corrosion of reinforcement.

1.2.6 Thermography Methods

Infrared thermography measures the thermal radiation emitted by the tunnel’s walls. Infrared registration techniques allow visual presentation of the temperature distribution on the surface [28][29][30]. The temperature on the surface represents the thermal flow through the surface, which in turn is influenced by the mechanical and/or hydraulic discontinuities of the structure. Consequently, thermal discontinuities on a surface reflects abnormalities within the underlying structure.

1.2.7 Radar Methods

Radar methods have been widely used to detect defects in tunnels and other structures, and the most used is the Ground-Penetrating Radar (GPR) [31][32][33][34]. GPR is the electromagnetic analogue of sonic and ultrasonic pulse echo methods. It is based on the propagation of electromagnetic energy through materials of different dielectric constants. The greater the difference between dielectric constants at an interface between two materials, the greater the amount of electromagnetic energy reflected at the interface.

1.2.8 Radiography Methods

X-rays, gamma radiation or neutron rays can penetrate structural materials and therefore can be used on inspection purposes [35][36][37]. The amount of radiation absorbed by the material is dependent upon the density and thickness. This radiation can be detected and recorded on either film or sensitized paper, viewed on a fluorescent screen, such as a television screen, or detected and monitored by electronic sensing equipment. With this method, limitations are imposed by accessibility to both sides of the object, long exposure times, and safety precautions required to protect both the operators and public.

1.2.9 Endoscopy Methods

Endoscopes or videoscopes consist of rigid or flexible viewing tubes that can be inserted into pre-drilled boreholes of an element under investigation to examine its condition [38][39]. Light can be provided by glass fibers from an external source. In the rigid tubes, viewing is provided through reflecting prisms and, in the flexible tubes, a fiber optics system is used. New models consist of an additional CCD chip to improve the images. These scopes allow close examination of parts of the structure which could not be otherwise viewed. Although this is a viewing instrument, some destruction of material is necessary for its proper use.

2 A Survey on Robotic Inspection

Even with the great variety of inspection methods presented in Section 1, presently structural tunnel inspection is predominantly performed through scheduled, periodic, tunnel-wide visual observations by inspectors who identify structural defects and rate these defects. This process is slow, labor intensive and subjective (depending on the experience and fatigue of the inspector), working in an unpleasant environment due to dust, absence of natural light, uncomfortable conditions or even toxic substances such as lead and asbestos. These working conditions are a main motivation behind the development of robotic systems.

2.1 Robotic Tunnel Inspection Systems

The use of robotics systems in the construction field
had been a common research area, and several studies review the advantages in the use of robotic platforms for construction [40][41] and underground construction [42] purposes. Robotic systems can complete the inspection process with objective results and high efficiency. They also improve safety by performing inspection in dangerous environments instead of the inspectors. Therefore, manual and (human) visual inspection are being replaced with more precise methods using mechanical, electronic and robotic systems and processing data provided by cameras, laser, sonar, etc.

In the case of the system in the Figure 4, a small mobile robot is equipped with a CCD camera [43][44][45]. The robot stays at a constant distance of the wall using a differential-drive wheel configuration, and a set of photos are taken. The camera is mounted on an anti-vibration device to stabilize the images. The robot goes through all the tunnel performing the inspection, but the data is processed after all the images are collected. The inspection consists in the detection of cracks via computer vision algorithms.

![Figure 4. Robotic platform with camera used in tunnel inspections [43]](image)

A similar robot can be found in the work by F. Yao et al. [46][47]. In this case, the mobile robot is equipped with 21 ultrasonic sensors and 6 video cameras. These sensors are mounted on the same plane and with a semi-ring shape. The inspection consists in the scan of the tunnel lining to search for deformations. The experimental results show that this system can detect the deformed inner-walls at the division of 14 mm when the robot moves at 20 mm/s.

Figure 5 shows a system built with an industrial manipulator robot [48]. The system consists of an eight-ton truck used as a base machine, tunnel cross section measuring systems, Electronic Distance Measuring (EDM) instruments employed to measure impact locations, an impact unit with five hammers that generates impact sounds and its equipped on the robotic arm, a lifter that raises the robot up to ceiling level, and finally a computer unit that controls all these components.

![Figure 5. A robotic tunnel inspection system that uses the impact sound method [48]](image)

The system uses an impact acoustics method for the inspection procedure, which impacts the concrete wall with hydraulic hammers, converts the impact sounds into electric signals, and then analyses them. The system is capable of finding exfoliation and cavities in a concrete lining. In order to maintain stable attitude, the truck has been equipped with outriggers on the non-motorized wheels. Three people conduct the impact sound diagnosis: a foreman, an operator and a driver. The machine is operated from the touch panel of the computer that is situated at the operator console.

Another example, seen in Figure 6, uses two lasers to perform a hammer-like inspection to detect inner defects in concrete structures like transportation tunnels [49]. The system is mounted on a motor vehicle and the technique is based on the initiation and detection of standing Lamb waves (or natural vibration) in the concrete layer between surface and inner defects. The concept consists in one laser used like a hammer to impact the surface and another one used to take the measurements. The system can detect various types of inner defects like voids, cracks and honeycombs. The accuracy of defect location is about 1 to 3 cm and the detection depth up to 5 cm.
Also of interest is the Tunnelings project [50]. The tunnel inspection system developed by Euroconsult and Pavemetrics, shown in Figure 7, is based on cameras and laser sensors that allow scanning a tunnel’s wall linings at speeds up to 30 km/h. The software of the system also allows the data from two different inspection runs to be rapidly compared, and structural changes and wall lining defects to be assessed.

The measurement sensors for the condition survey are installed on a truck capable of running on rails and on flat terrain. The vehicle comprises all the systems necessary for safe road and rail travel (lane occupation indicator, speed governor, electric power supply for all systems, signaling equipment etc.). It can hold up to 6 laser cameras, as. Each pair of laser-camera units inspects a 2 m wide section with an accuracy of 1 mm. Using the six cameras, tunnels with a 9 m diameter can be inspected at the system’s maximum resolution.

The system developed by N. Sano et al. [51] consists in a crack detecting vehicle equipped with laser sensors and CCD cameras. The vehicle is driven through the tunnel by an operator and the cameras take pictures of the tunnel walls. The isolated images taken by the cameras are merged together into a surface map of the tunnel. After the map is obtained, a dedicated vision software detects cracks in it.

The system shown in Figure 8 checks for voids behind lining by drilling holes with a mechanized crane [52]. It performs high speed drills of 33 mm diameter by a combination of rotation and striking the lining concrete of a tunnel surface in order to investigate the thickness of a lining and the height of a rear cavity with high accuracy.

This example and others (such as methods based on a mechanized hammering tester installed on a crane, Figure 9) were described by Hideto Mashimo and Toshiaki Ishimura in [1], where they define the status of road tunnel inspection and maintenance in Japan in 2006.

More examples of systems working on Japan tunnels can be found in the work by Toshihiro Asakura and Yoshiyuki Kojima [53], which shows maintenance technology and typical deformation cases of Japanese railway tunnels, along with some methods of inspection.
and diagnosis and three case studies (Tsukayama Tunnel, Fukuoka Tunnel and Rebunhama Tunnel). The inspection methods examples includes Hammer testing on the lining (performed by an operator in this case), crack measurements in tunnel lining using line-sensor cameras mounted in a vehicle in rails, and investigation of the surface of the tunnel lining using an infrared camera and CCD cameras.

The mechanized cleaning truck shown in Figure 10 is an example of a tunnel maintenance system [54][55], which is important to prevent further damage. It has been designed by engineers of Colas, Switzerland, in collaboration with operators of road networks. It was set up in seven months in 2012. This tunnel cleaning system consists in a standard commercial truck equipped with eight mechanical arms with different types of brushes. The arms and the brushes have hydraulic actuators which provide movement and water flow for the cleaning process. The mechanical arms can be positioned remotely to adapt to different tunnel geometries.

![Figure 10. Tunnel cleaning system [55]](image)

This is achieved with a communication briefcase-like system controlled by an operator near the truck. A second operator is needed to drive the truck at a speed of 2 km/h while the tunnel walls are being cleaned with the brushes. The system can operate in tunnels of a 7.66 maximum height. Only one half of the tunnel section is covered each time, without blocking the traffic on the free lanes.

Another commercial example is the IRIS Hyrail built by Penetradar [56], shown in the Figure 11. The system is based on a GPR sensor mounted in a telescopic piece in the front of a Hyrail vehicle (e.g. a vehicle able to go on road and on rails). The GPR positioning device can be rotated to cover the sides and top of the tunnel walls and the motorized boom can be retracted to avoid obstructions. Penetradar provides specialized software to manage data collection, data processing and display of GPR data.

![Figure 11. IRIS Hyrail system inspecting a tunnel. Note the capability to be mount on rails [56]](image)

When tunnels that need to be inspected have reduced dimensions, such as underground tunnels used to deploy power cables, the use of robotic platforms is more than appropriate. In this scenario, small tele-operated mobile robots can make inspections providing visual and concentration data of some poisonous gases, like the system by Fu Zhuang et al. [57] shown in Figure 12. This tele-operated robot (420mm long, 320mm wide and 300mm high) can operate in 1 meter wide tunnels, move at a rate of 24 m/min, and has 2 hours of autonomy. Its sensor system includes a pan-tilt-zoom camera, inclinometer, gyroscopes, gas sensors (CO, CH4, CO2, and O2), thermometer, IR distance sensors, and ultrasonic sensors.

![Figure 12. Cable tunnel inspection robot [57]](image)
In other cases, cables are not inside small tunnels but along a greater one placed on the walls. Taking this into account, Songyi Dian et al. [58] designed a robot based on a shrimp-rover vehicle [59] with six wheels able to go over tunnel power cables while making the inspection. Unfortunately this work is only theoretical and the robot does not exist physically. Other examples of tunnel cables inspection robots can be found in B. Jiang et al. [60][61] and Claudio Mello et al. [62].

Another type of small tunnel is ventilation tunnels. R. Minichan et al. [63] designed three different mobile robots to inspect the ventilation tunnels of the H-Canyon Facility in 2003, 2009 and 2011. Due to the toxic environment of the tunnels, only a robot could perform the inspection process. The control of the robots is made remotely and the system was connected through a long tether to the control station. The inspection consisted in a visual assessment with the images provided by the robot cameras. Figure 13 shows the three robot models.

Not all tunnels are designed to carry vehicles, people or cables. Water distribution is managed with tunnels too, and different solutions must be used to inspect this kind of structures. In this scenario, alternatives to mobile wheeled robots include Autonomous Underwater Vehicles (AUV) [64] and Remotely Operated Vehicles (ROV) [65][66] which can exploit the use of sonar sensors for the mapping procedure.

Figure 13. Three ventilation tunnel robotic inspection systems [63]

All the systems seen in this section have their own sensors, and publish the data obtained to perform the assessment of the structure. However, an alternative strategy involves the use of sensors embedded in the structures to be inspected, such as strain gauges, which are usually more precise and reliable. Brian Esser et al. [67] implemented this method and developed a robot capable of remotely powering and collecting data from a network of embedded sensing nodes, and providing remote data access via the Internet. The system uses Addressable Sensing Modules (i.e. ASM’s) to sample data from a wide variety of sensors (e.g. peak displacement, peak strain, corrosion, temperature, inclination, etc.). This kind of system is useful in long tunnels where a wired sensor network is difficult to implement, or in tunnels with complicated access.

2.2 Other Related Systems

There is also a great number of systems that have been developed for inspection purposes that may be applicable to the tunnel environment. These systems were designed for the inspection of bridges, pipes, or pavement among others. The similarities in geometry, materials, defects and inspection procedures of these systems leads to similar technological solutions able to be used for the tunnel inspection procedure.

Two examples of this kind of robots are the ones designed by Carnegie Mellon. One is a tele-operated vehicle for mapping of abandoned mines [68] (similar in geometry and dimensions to a tunnel). The other is a similar robot, developed for the inspection of hazardous environments with cameras embedded on an articulated arm [69].

Inspection of pipes is also a relevant area. Some pipes’ dimensions can be up to 3 m, similar to a small ventilation tunnel. These scenarios involve small tele-operated robots [70][71] able to make visual inspections along with mapping of pipes and deformation analysis [72]. Some of these systems can even perform cleaning [73] and maintenance [74][75][76] operations. Commercial solutions like Redzone Robotics [77] also provide pipe inspection robots.

Certain robots designed for bridge inspection have the vehicle-crane configuration, similar to some of the ones used in the tunnel inspection [1][53][48][78][79], but modified to reach zones under the bridges [80][81][82]. The majority of these systems use the same sensors to achieve the inspection (vision, laser, ultrasound, etc.) and some solutions use robotic arms installed on the tip of the crane to perform maintenance operations [83]. On the other hand, robots designed to inspect the superior part of bridges mainly focus on the crack detection of the pavement [84], and are similar to road pavement inspection robots [85][86][87] that have similar sensors and algorithms to the tunnel crack detecting in tunnels; while others mount a variety of sensors to achieve a more complete assessment [88].
Lastly, climbing capabilities of some robots are being used to perform inspections in zones with difficult access [89][90]. Early attempts used legged robots capable of climbing metallic-based structures [91], while the most modern systems use suction [92] or negative pressure [93][94][95][96] devices to attach to structure walls.

3 Main Drawbacks

As seen in the previous section, all of the robotic tunnel inspection systems are tele-operated in some way. This is one of the main disadvantages of all these systems. These systems slightly improve the working conditions of the operators in the tunnel, but to successfully overcome all the problems of the manual inspection procedures, a fully automated tunnel inspection system must to be developed.

The need for one or more human operators of the presented systems sometimes require the workers to be in the same location as the robot, eliminating one of the benefits of the robotic inspection, which is remote operation. Because of this, the operator is exposed to the dangerous tunnel environment, including large isolated areas, low visibility, dust, humidity or even toxic gases.

In some cases, the inspection data gathered by the system is not enough to make a complete assessment of the tunnel, and an additional manual inspection performed by a qualified inspector is required. This causes subjectivity in the inspection results that relies on the inspector judgment and may contain diverse errors.

In other cases, the limitations in the type of communication used (e.g. tether length, wireless area) leads to the same problem described before. In the case of wired tele-operation, the main bottleneck is the length of the cable itself, which limits the operational range of the robot with respect to the control station. Regarding the wireless communication, one problem is the signal intensity decay, and it could depend on the tunnel length, material or complexity. Other problem is the bandwidth that needs to be high if the robot has little autonomy and sends a large amount of data to the tele-operator.

The difficulties mentioned before mainly affect the quality of the inspection and the operators working conditions, but an important aspect of the inspection procedure is the economical impact that is incurred when a tunnel must to be closed for inspection. Leaving the tunnel inoperable reverts in losses for the tunnel owners and users. Because of this, taking into account this issue is desirable to develop systems that can allow the use of the tunnel during the inspection procedure.

One of the solutions to these problems begins with the improvement of the automatic behaviors of the robots. With a fully automated inspection system, the security of the operators is guaranteed, along with the management, quality and objectivity of the tunnel data. Current efforts should be focused on obtaining more autonomous behaviors, that may adapt to different tunnel environments with less operator dependency.

4 Current Efforts in Fully Automated Tunnel Inspection

The latest developments in robotic and automation science allow the current tunnel inspection systems being developed to become more automated than the previously seen tele-operated systems. This confirms that the tendency is to reach a fully automated inspection system that allows a remote inspection with no direct human operation needed. The first part of this section reviews the TunConstruct system, which was partially autonomous, while the second part explains the ROBINsPect system, which aims to perform fully automated tunnel inspection.

4.1 TunConstruct System

The TunConstruct project [78][79] was part of the European Commission 6th Framework Program (FP6), and was conducted by 41 partners from 11 European countries. This project’s objectives involved the development of a robotic system capable of performing inspection and maintenance in concrete tunnels.

The TunConstruct system is shown in Figure 14, and consists in a robotic arm on the tip of a crane, mounted on a vehicle. The system is able of applying a composite material on cracks of a tunnel. Inspection is performed through a user-friendly guided HMI (Human Machine Interaction), where a 2D camera image stream is displayed and operation procedures are requested. Visual servoing based on the depth measure captured by the tool’s laser telemeter, and operation-oriented actuators are coordinated through task-specific control software, allow the process to be automatically performed by the robotic system.
The robot arm chosen for the application was the Mitsubishi PA-10, a 7 DOF manipulator with a 10 kg load capacity and a 1 meter maximum extension range. The global increment of this range is achieved by mounting the Robot on a 5 meter extensible articulated lift platform. The HMI is installed in the wheeled vehicle’s cabinet to which the articulated lift platform is attached. Power for the system can be supplied from an on-board generator, the wheeled vehicle’s motor, or the tunnel’s basic provided services.

Once in the tunnel, the inspection procedure is achieved by means of a camera mounted in the robotic arm. Using the HMI with the images provided, the operator can guide the robot and select the crack in the tunnel surface in which the system is going to apply the repair material. For the maintenance process, the robot uses a specially designed tool to prepare the selected surface and apply the resin and the composite material on it. The superficial preparation is accomplished by blowing compressed air to clean the zone. The mentioned tool is composed by two complementary systems: a material application system composed by mechanical subsystems and actuators, and a vision and security system composed by the camera, the laser distance sensor, and security micro-switches. The resin is attached to the tool in a cartridge, while the composite (Fiber Reinforced Polymer) is loaded in the form of a roll.

The system was tested successfully first in laboratory conditions and then in real, non-controlled environments in tunnels in León, Spain. Figure 15 shows one of the mentioned tests in real tunnels.

4.2 ROBINSPECT: Towards a Fully Automated Robotic Inspection System

ROBINSPECT (ROBotic System with Intelligent Vision and Control for Tunnel Structural INSPECTion and Evaluation) [97] is a project co-funded by the European Commission, under its 7th Framework Program (FP7). The project begun on October 2013 and will finalize in 2016. This project comprises the design of an autonomous robotic system capable of performing intelligent inspection and assessment of a tunnel in one pass. Figure 16 depicts a schematic representation of the ROBINSPECT system. Its similarity with previous projects of the field is in line with a “not reinventing the wheel” philosophy, and instead focusing on automation, intelligence, and benchmarking objectives.
ROBINSPECT is driven by the tunnel inspection industry and adapts and integrates recent research results in intelligent control in robotics, and computer vision tailored with semi-supervised and active continuous learning and sensing, in an integrated robotic system for automatically scanning intrados for potential defects on concrete surfaces and detecting and measuring radial deformation in cross-sections, distances between parallel cracks, and cracks and open joints that impact tunnel stability, with millimeter accuracies. This will allow, in one pass, both the inspection and structural assessment of tunnels.

The initial dataset on tunnel defects will be provided from case studies (including London Underground, an interesting case study as the network incorporates the world’s first underground railway) to be used not only for transfer learning but also for the evaluation of structural models. The robotic system will be evaluated and benchmarked at the research infrastructure of tunnels of VSH, at three road tunnels of the Egnatia Motorway, and sections of the railway tunnel of London Post Office.

The ROBINSPECT extended mobile robotic system will be a wheeled robotic system that will be able to extend an automated crane to the lengths commonly found in tunnels (4 to 7 meter range) sustaining a robot manipulator while being automated through the use of robotic controllers. This robotic system will be composed by three subsystems: a mobile robot, an automated crane arm, and an industrial-quality robot manipulator. Systems with similar configurations are the TunConstruct system [78][79], and other existing systems of the construction industry [98]. Basic 1D (such as laser, infrared and ultrasound proximity and distance sensors), 2D (such as vision camera and SICK sensors) and 3D (as in time of flight or similar technology) sensors will be incorporated to the subsystems to enable collision avoidance. Apart from that, a special new ultrasonic sensor will be equipped to measure cracks width and depth. The mechanical concept design of the robotic system will be based on an existing industrial platform, so to ensure the successful operation under real tunnel conditions, although localization and navigation tests will be performed initially under simulated environments.

At the software level, Component Based Software Engineering (CBSE) techniques will be applied. Specifically, a set of low-level “device drivers” for each of the subsystems will be developed to allow the component’s control to be integrated into the developments of the following tasks of the project. Currently, several robotic software architectures (YARP [99], ROS [100], OROCOS [101], etc.) for implementing CBSE exist and are interoperable. The dynamic and kinematic requirements of the robotic platform needed to reach the measurement area will be designed. Special attention will be placed in keeping the vehicle stability as well as developing the platform modular enough to allow both road and railway navigation. Controllability will be improved as much as possible to achieve an accurate path following in tunnels.

A global controller for the system in the presented scenario will be developed for two main reasons. First, the additional length requirement means that any deviation of the control output at any of the stages of the low-level will be multiplied at the end-effector of the robotic manipulator (sensing tip). Secondly, the three different subsystems (the mobile robot, the automated crane arm, and the robot manipulator) must fulfill a set of required behaviors conjunctly: only a global controller can assure coherent and optimized trajectories. Moreover, the input of this global controller will come from the vision and sensory systems as data of three very different natures: an online stream of updated 3D model data of the tunnel environment coming from the 1D/2D/3D sensors, associated uncertainties (both intrinsic to the nature of the sensors as information regarding the confidence at each given instant), and additional semantic information regarding the state of the system and the required action/behavior. An intelligent controller will be developed as the global controller. It will update its prior belief model of the environment continuously by using the 3D model stream as input while taking into account the uncertainties as confidence values of the given data. The semantic information will be treated as conditional clauses for generating trajectories that comply with the general system’s requisites. The feedback will be used for the global controller to auto-tune its parameters. The system’s safety for the robotic system and environment will additionally be assured by the local controllers developed for each of the as the global intelligent controller is set at the high level to send references to these (and not directly on the actuators).

5 Future Trends

The previous sections have presented an extensive review on tunnel inspection methods and robotic systems with tele-operated and more modern
approaches focusing on the structural inspection problem. A discussion is open in this section, in the context of presenting solutions based on forthcoming improvements in the robotics area and related fields, in order to have a more complete vision of the future trends of technology and new approaches.

In terms of complex and unstructured environments, the great majority of large inspection systems cannot work properly, or have difficulties in doing so because of the use of wheeled platforms. One possible solution could be the use of legged robots with insect-like legs, such as the proof of concept robotic harvest system developed by John Deere [102] or the quadruped robot built by Boston Dynamics [103], to go through rough and uneven terrain and avoid unexpected obstacles easily. Another technology applicable to the future tunnel inspection process that avoids the mobility limitations are the Unmanned Aerial Vehicles (UAV) developments. There are some works that start to use this approach [104][105]. One advantage is that these robots may be produced at low cost, and also used simultaneously as they do not interrupt traffic. A complete swarm of robots with reduced dimensions may be used to perform the inspection process faster and exhaustively, concurrently gathering large amounts of data. Figure 17 depicts a hypothetical tunnel inspection scenario using legged robots and UAVs.

A more advanced notion of legged robots would be the use of humanoids and anthropomorphic robots. Advances within the field of robot walking gait generation and solving stability issues are required for this step, but it is favored in turn with the advantages of the capabilities of these robots, which have the potential of using the same vehicles and tools that human workers currently employ, without exposing humans to the associated risks.

On the other side, reducing robot dimensions with advances in the field of nanotechnology could lead to a different type of inspections in the future, using nano-robots, a.k.a. nano-bots. These nano-mechanisms could penetrate tunnel walls through cracks or small fissures and perform inspection and structural assessments of the materials from the inside, or search for invisible cracks in the extrados. Tests that are currently destructive could be performed through the use of non-destructive nano-bot mechanisms.

In terms of including preventive measures within the design phase, another good idea is to take into account the inspection and maintenance processes at the time of the structure construction. A set of rails may be planned and located along the tunnels, ready to a rail-robot to be attached. In this way, the robot (e.g. a robotic arm with sensors at the tip) may travel around the structure and inspect it with different sensors. Additionally, the robot could store different tools and sensors, and only pick up the appropriate ones when required for a given tunnel structural assessment.

Predictive efforts should benefit from advances in sharing and storing big data and information theory. Tunnel data can be saved and retrieved, and this process should become globally accessible through integrated services. Data may be used to compare different inspections of the same tunnel or similar tunnels around the world. Information from nearby and similar structures could be used to infer an idea of the land movements and predict how this will affect to the structures.

6 Conclusions

The vast majority of the presented tunnel inspection methods presented is currently performed manually by human workers in tunnel environments. Moreover, most part of the presented robotic inspection systems is actually tele-operated, and none of the presented currently operational systems is fully autonomous. One or more human operators must move the robotic system, control where to go, and command what to do. In addition, in most cases the data is processed after completing the inspection procedure, causing a delay between the time the inspection is completed and the results are available. This makes the tunnel unsecure for longer times if critical defects exists.

1 Original background photograph by Scott Beale
This situation causes the efforts to be focused on developing systems that aim to be fully autonomous, such as the ROBINSPECT system. Future systems must be capable of both inspection and maintenance with minimal human intervention, and perhaps with no supervision at all. Much research is ongoing in able to design better systems, capable of performing accurate and cost-effective inspection, maintenance and assessment of civil structures that reverts in more safety in environments and less budget spent in reparations.

Acknowledgements

This work was supported by RoboCity2030-II-CM project (S2009/DPI-1559) funded by Programas de Actividades I+D in Comunidad de Madrid and co-funded by Structural Funds of the EU, ARCADIA project DPI2010-21047-C02-01 funded by CICYT project grant, and ROBINSPECT project number 611145 co-funded by the European Commission under 7th Framework program. The authors would also like to acknowledge the work of the other partners involved in the TunConstruct project: DRAGADOS, SIKA, and M. Lorenzana.

References


[85] Fugro Roadware: Fugro Roadware. no date.


[77] Redzone Robotics: Redzone Robotics - Water Infrastructure and Sewer Inspection Robots and Software. no date.


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

INVITED PAPERS
Economic Model Predictive Control - A Review

Tri Tran\textsuperscript{a}, K-V. Ling\textsuperscript{a}, and Jan M. Maciejowski\textsuperscript{b}

\textsuperscript{a}School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 
\textsuperscript{b}University of Cambridge, Department of Engineering, Cambridge CB2 1PZ, United Kingdom.

E-mail: tcTran@ieee.org, ekvLing@ntu.edu.sg, jmM@eng.cam.ac.uk

Abstract -
Leveraging the advanced estimation and control algorithms for power systems have always been associated with the renewable energy sources, rational power generation, consumer stimulus, emission reduction, as well as economically viable objectives. The model predictive control (MPC) strategies, that employ an economic-related cost function for real-time control, has lately proved a numerically efficient approach to managing the portfolio of energy usage in various residential and industrial projects. They are designated as economic MPCs, whose main endeavour is to cope with regularly changing energy prices. Unlike the traditional MPCs, economic MPCs optimize the process operations in a time-varying fashion, rather than maintain the process variables around a few desired steady states. The process may thus totally operate in the transient state with economic MPCs. This paper provides a rigorous review on the developed and progressive economic MPCs, as a contribution to the field while it is still in its infancy. In the second part of the paper, we briefly show the potential of applying the quadratic dissipativity constraint, previously introduced, to the closed-loop stability problem of an economic MPC problem.

Keywords -
Economic MPC; Quadratic dissipativity constraint

Acknowledgement: This publication is made possible by the Singapore National Research Foundation under its Campus for Research Excellence And Technological Enterprise (CREATE) programme, and was supported by Cambridge Centre for Carbon Reduction in Chemical Technology (C4T) and Singapore-Cambridge CREATE Research Centre (CARES).

1 Introduction

There are different methods in optimization and real-time control, direct or hierarchical, to deal with economic problems in the process and other industries. The articles \cite{21, 31, 10} provides a clear overview of practical approaches to such economic problems in the process industry. The economic MPC has been found effective among these approaches, in which the real-time optimization (RTO) layer is not required for computing targets to the lower layer MPC as usually be the case in the process industry. Figure 1 depicts the differences between the typical hierarchical RTO plus MPC structure and the economic MPC. The objective function of EMPC consists of the economic objective and the MPC target tracking objective functions. According to Rawlings et Al. \cite{31}, an EMPC “directly and dynamically optimizes the economic operating cost of the process, doing so without reference to any steady state”. EMPC has been successfully implemented for HVAC energy saving applications, and started flourishing in power systems, as has been found in recent works, including those from \cite{17, 18, 15, 27, 1, 26, 30, 33, 5}. For chemical process systems, the review in \cite{10} has also pointed out that “while steady-state operation is typically adopted in chemical process industries, steady-state operation may not necessarily be the economically best operation strategy”. The review in this paper outlines concisely the main features of the developed economic MPC that will be useful for interested researchers and industrial practitioners.

Rawlings et Al. have first described the so-called unreachable set-point in MPC implementations in their paper \cite{32}, and shown that the set-point tracking MPC exhibit some advantages over the traditional target-tracking MPC \cite{24, 23, 22}, when the set-points are not reachable. It is recommended that, the approach “should also prove useful in applications where optimization of a systems economic performance is a more desirable goal than simple target tracking” \cite{32}. A strictly convex cost function $L(x, u)$ is considered in this paper. For a traditional MPC that has a reachable steady-state target, the optimal $(x^*, u^*)$ is tracked and the value of $L(x^*, u^*)$ is zero. In a set-point tracking EMPC, the cost function has the form of $L(x - x_{SP}, u - u_{SP})$, where $(x_{SP}, u_{SP})$ is usually different than $(x^*, u^*)$. An apparent difficulty around the pre-existing MPC theory for unreachable set-points is that the controller cost function was established as a Lyapunov function for the closed-loop system, while the cost function for the unreachable set-point problem may not
The optimally steady state operation herein means the stage cost at steady state is the upper bound of the average stage costs. A dissipation inequality with an adequately chosen supply function have been used in this work to extend the sufficient condition for the asymptotic stability of the steady state. The authors have chosen the storage and supply functions such that the dissipativity is also sufficient for the optimal steady-state operation, while the strict dissipativity assures the closed-loop stability. The distance from the current stage cost to the steady state one plays an important role in the supply function, but not in the Lyapunov function as in [6]. The rotated cost function is formed in a slightly different manner than those in [6]. Both equality terminal constraint and inequality terminal constraint with terminal penalty have been addressed in [31]. As a complement to the results in [4], the sufficiency of dissipativity for the optimal steady-state operation has lately been analyzed in [29].

Another approach to guarantee the closed-loop stability for EMPC is to enforce a stability constraint to the optimization. In this enforcing stability approach, the stage cost is modified directly. A rather critical view on the modified cost function is as follows: The stability enforcing problem is described as the one that lies somewhere between the two extrema, when the original stage cost is added with an appropriately chosen positive scalar function $\alpha(\ldots)$. One extreme represents the original economic stage cost, when choosing $\alpha(\ldots) \equiv 0$, which may leave the optimal steady state unstable. The standard tracking problem is viewed as the other extreme, assigning $\alpha(x,u)$ with the usual tracking objective while canceling the economic one. The asymptotic stability of the optimal steady state is obtained while sacrificing the economic objectives.

Periodic terminal constraint and average constraints presented in [4] are the two main extensions for applications having the pre-computed optimal periodic solutions and for the unsteady closed-loop processes. The average constraint is determined by having the average of a predefined auxiliary output variable $y = h(x,u)$ belong to a time-varying convex set. The periodic terminal constraint, obtained from a dynamic state feedback strategy, is introduced here for achieving a periodic optimal solution described in another work [3]. The notion of sub-optimally operate off steady state is also introduced herein. When the average constraint is satisfied, it is said that the “control system optimally operates at steady state on averagely constrained solutions”. This statement also implies that the system sub-optimally operates off the steady state. The Lyapunov function for the periodic EMPC has recently been introduced in [42].

Along the line of enforcing stability, a self-tuning terminal weight version of economic MPC has been devel-
oped in [29]. This departed from the work in [11] which “shown that if the terminal weight is large enough, then the cost of the predicted terminal steady-state will be arbitrarily close to the cost of the best reachable steady-state.” [29]. The main aim of this work is to eliminate the condition on having a sufficiently large terminal weight to keep up with the control performance, as well as to maintain the economic MPC objective. The optimal steady state cost is assumed zero herein. Under mild assumptions, the authors have proved that there exists an upper bound for the average stage costs of the closed-loop system, then determined an updating rule, which allows for a non-monotonic stage cost such that it minimizes that upper bound.

The research in smart grid and energy saving has started flourishing from economic MPCs, as has been found in recent works including those from [1, 15, 17, 18, 26, 30]. Unlike the rigorous developments in the previous three papers, these application-oriented papers simply prolong the predictive horizon for feasibility or applies an infinite-horizon approximation, while emphasizing the effectiveness of the economic MPC approach in power systems and energy saving problems. The existence of such prolonged predictive horizons for recursive feasibility under mild assumptions has lately been proved in [14]. The authors in [29] called the conditions developed in [14] as controllability.

Apart from the above, the work in [20] interpreted the economic problem as the one having cyclic steady state, which arose from the regularly changing energy price. The solution is to normalize the system around such cyclic steady state. A transformed system having a fixed steady state is obtained as a result of the normalization. The MPC objective is then constructed with this transformed model. Similarly to the work in [32], the chosen Lyapunov function is also the difference between the current stage cost and the steady state stage cost. The stability proof is supported by a Lipschitz continuous property, plus an assumption on the weak controllability of the cyclic steady state. Different strategies for the infinite-horizon EMPCs have also been proposed in [40, 19, 27, 26, 41, 30]. However, these are only the preliminary results or EMPC alike formulations, thus will not be given a review herein.

The chemical process control field has envisaged the importance of economic MPC. Recent research works in [16, 9, 43] have shown strong interests and beneficial results. These works extend the previously developed stabilizing method of Lyapunov-based MPC [28] to EMPC. The authors have criticized the method in [6] that, “it is difficult, in general, to characterize, a priori, the set of initial conditions starting from where feasibility and closed-loop stability of the proposed MPC scheme [6] are guaranteed.” Nevertheless, the proposed scheme in this work consists of two artificially operational stages, which somehow resembles the idea of stability margins. The disadvantage of this proposed approach could possibly be at the difficulty to define such stability margins while not jeopardizing the optimality. Moreover, the periodic steady states and unreachable set-points have not been addressed formally. We are, therefore, not pursuing a detailed review for the Lyapunov-based EMPC herein. Interested readers can find a detailed review in [10] instead.

In the second part of this paper, we will show the potential of applying the quadratic dissipativity constraint, previously developed in [35, 38, 37, 39], to the economic MPC problem that is able to avoid the disadvantages of the above approaches. In this quadratic dissipativity constraint approach, two constraints on the initial control vector, one as a stability constraint, one as a recursive-feasibility constraint, for model predictive control are derived for implementation. Recursive feasibility and input-to-power-and-state stability are simultaneously achieved as a result of imposing these two additional constraints into the MPC optimization. The MPC stage cost is not employed as a Lyapunov function in our approach, while the quadratic dissipativity constraint allows both non-monotonic decreasing storage function (of the dissipativity inequality) and economic MPC stage cost. Developments for the dynamically optimal operating point or rotated steady state with the quadratic dissipativity constraint is underway.

This paper is organized as follows. Notation, system model and economic MPC problem formulation are outlined in Section 2. Section 3 is reserved for summarizing the results of the auxiliary stage cost for stability purpose, the proposed Lyapunov function and supply rate for dissipativity, terminal constraints and terminal cost, and asymptotic average performance in the works of Rawlings et Al. The potential of quadratic dissipativity constraint for economic MPC is briefly discussed in Section 4. Section 5 concludes this paper.

2 Preliminaries

2.1 Notation

Capital and lower case alphabet letters denote matrices and column vectors, respectively. \( (\cdot)^T \) denotes the transpose operation. \( \| u \| \) is the \( \ell_2 \)–norm of vector \( u \). \( \| M \| \) is the induced 2-norm of matrix \( M \). In the discrete time domain, the time index is denoted by \( k \), \( k \in \mathbb{Z} \). \( x^+ \) denotes \( x(k + 1) \) for conciseness. In symmetric block matrices, * is used as an ellipsis for terms that are induced by symmetry. Bold typefaces are used to denote decision variables.
2.2 System Model and Economic MPC Problem

The general system model of the form

$$\Sigma : x^+ = f(x, u),$$

(1)

without any disturbances, and the variable constraints of $u \in U \subset \mathbb{R}^m$ and $x \in \mathcal{X} \subset \mathbb{R}^n$ and the state transition map $f : \mathcal{X} \times U \rightarrow \mathcal{X}$ are usually considered in the developed economic MPCs. $f$ is normally Lipschitz continuous.

The economic stage cost, which is not the target tracking one, is represented by $\ell(x(k), u(k))$. $\ell(\cdot)$ is ususally convex and continuous for linear systems. The optimisation problem is not normally convex for nonlinear systems. Similarly to the target tracking MPC problem, the optimization problem of economic MPC will be to minimize

$$V_N(x, u) = \sum_{k=0}^{N-1} \ell(x(k), u(k))$$

(2)

subject to the model (1), as well as point-wise constraints of $(x(k), u(k)) \in \mathcal{Z}, \ k = 0, 1, \ldots, N - 1$, for some compact and time-invariant set $\mathcal{Z} \subset \mathcal{X} \times \mathcal{U}$, the terminal constraint $x(N) = x_s, x(0) = x$, where the decision vector $u := [u(0), u(1), \ldots, u(N-1)]$ is the MPC computed control sequence at each time step. This finite horizon optimization problem is solved recursively. Only the first element of the optimal control sequence $u$ is applied to control the plant. The state measurement or estimate is fed back as the new initial state for solving the problem in the next time step. The admissible set $\mathcal{Z}_N$ is defined next. $\mathcal{Z}_N$ is a set of $(x, u)$ pairs satisfying the constraints

$$\mathcal{Z}_N := \{ (x, u) \mid \exists x(1), \ldots, x(N) : x^+ = f(x, u),
\forall k = 1, 2, \ldots, N-1, x(N) = x_s, x(0) = x \}.$$

(3)

The set of admissible states $\mathcal{X}_N$ as the projection of $\mathcal{Z}_N$ onto $\mathcal{X}$ is defined as

$$\mathcal{X}_N := \{ x \in \mathcal{X} \mid \text{such that } (x, u) \in \mathcal{Z}_N \}.$$

The control sequence $u$ is called feasible with the initial state $x$ if $(x, u) \in \mathcal{Z}_N$. The so-called optimal steady state is defined as the pair $(x_s, u_s)$ fulfilling such feasible condition, i.e.

$$\ell(x_s, u_s) := \min_{x, u} \{ \ell(x, u) \mid (x, u) \in \mathcal{Z}, x^+ = f(x, u) \}.\quad (3)$$

The assumptions involved in this problem will be the continuity of $f$ and $\ell$; the admissible set $\mathcal{X}_N$ contains $x_s$ in its interior; and the existence of finite state/input gains [31].

Unlike the traditional target tracking MPC, the optimal steady state stage cost $\ell(x_s, u_s)$ is not necessarily smaller than $\ell(x, u)$ in the economic MPC. In economic MPC, it is desirable to achieve the following asymptotic average limit:

$$\lim_{T \rightarrow +\infty} \sup_{k=1}^{T} \frac{\sum_{k=0}^{T} \ell(x(k), u(k))}{T+1} \leq \ell(x_s, u_s).$$

The formal definition for asymptotic average is given in Section 3.1. The authors in [31] have proved that: “There exists at least one admissible control sequence that steers the state to $x_s$ at time $N$ without leaving $\mathcal{X}_N$ and the closed-loop system has an asymptotic average performance that is at least as good as the best admissible steady state”. While asymptotic-average economic performance is guarantee, this problem (2) does not assure the closed-loop stability. The next section addresses the proposed solutions to the stability problem of economic MPC. One may expect that, the optimality will be, to some extends, compromised due to these stability measures.

3 Augmented Stage Cost, Strict Dissipativity and Stability

Details in the chosen Lyapunov function and the supply rate, average performance, strong duality and relaxation with dissipativity, as well as strict dissipativity and stability in the works of Rawlings et Al. [6, 4, 31] are presented in this section.

In traditional stability-guarantee MPCs, the optimal cost of $V_N(x)$, denoted as $V_N^*(x)$, is employed as a Lyapunov function for the closed-loop system. This optimal cost is monotonically decreasing along solutions of the closed-loop system, i.e. $V_N^*(x^+) \leq V_N^*(x)$. In economic MPC, however, that is not necessarily the case, even when the system is stable. “More fundamentally, for general nonlinear systems and cost functionals, it is not even guaranteed that $x_s$ is an equilibrium point of the closed loop system. Since there exists $(x, u)$ such that $\ell(x, u) < \ell(x_s, u_s)$, it may be the case that the optimal trajectory from $x_s$ at time 0 to $x_s$ at time $N$ is different than $x(k) = x_s$ for all $k = 0, 1, \ldots, N$” [31].

3.1 Storage function and supply rate

The auxiliary rotated stage cost $L(x, u)$ is used for achieving the monotonic decreasing instead of the actual stage cost in (2) of the economic MPC problem. The supply rate is chosen as

$$s(x, u) := \ell(x, u) - \ell(x_s, u_s).\quad (3)$$

And the rotated stage cost is of the form

$$L(x, u) := \ell(x, u) + V_P - V_P^+,\quad (4)$$

with a chosen storage function $V_P(x, P), P$ is the multiplier.
For any vector \( v(k) \), the asymptotic average is defined as follows [4]:

\[
\text{Av}[v] = \{ v \in \mathbb{R}^n \mid \exists \kappa_n \to +\infty : \lim_{n \to +\infty} \inf \frac{\sum_{k=\kappa_n}^{\kappa_n+1} |v(k)|}{\kappa_n+1} = \bar{v} \}.
\]

In [6], it assumed the strong duality of the steady state problem, known via

\[
\min_{(x,u) \in \mathbb{Z}} \ell(x, u) + V_p - V_{P^+} \geq \ell(x_s, u_s),
\]

for achieving stability. And it is proved in [4] that the dissipativity w.r.t the above chosen storage function and an extended supply rate is a relaxation of strong duality, while strong duality is sufficient for dissipativity. According to [31] and [4], the dissipativity is sufficient for the optimal operation at steady state, defined as \( \text{Av}[\ell(x, u)] \subseteq [\ell(x_s, u_s), +\infty) \), but not for stability. The closed-loop stability which is concluded by the strict dissipativity is proved in [4].

### 3.2 Closed-loop stability

There is a variety of formulations for achieving the closed-loop stability herein.

#### 3.2.1 Terminal constraint

For the above economic problem (2) with equality terminal constraint, \( x_s \) is an asymptotically stable equilibrium point of the closed-loop system with region of attraction \( \mathcal{X}_N \) if it is strictly dissipative with respect to the supply rate (3). The proof starts with creating an auxiliary augmented problem with rotated stage cost, then shown the feasible sets, \( \mathcal{X}_N \), coincide with those of original problem, and so does the optimising result. By having the rotated cost-to-go as the Lyapunov function, the stability is achieved by the strict dissipativity accordingly.

#### 3.2.2 Terminal cost and terminal region

The authors have also proved in [4] that the closed-loop stability is achievable by using an adequately chosen terminal cost with an inequality terminal constraint (instead of an equality one). The proof also used an augmented rotated stage cost, considered the rotated cost-to-go as the Lyapunov function, and employed the strict dissipativity property.

#### 3.2.3 Enforcing stability with modified stage cost

The stability is also achievable by directly modifying the stage cost, which is considered as a tuning activity in [31]. For

\[
h(x, u) = V_p - V_{P^+} + \rho(x) - \ell(x, u) + \ell(x_s, u_s),
\]

in which \( \rho(x) \leq L(x, u) - L(x_s, u_s) \), the modified stage cost of the form \( \ell_m(x, u) := \ell(x, u) + \alpha(x, u) \), with \( \alpha(x, u) \geq h(x, u) \), will help achieve the asymptotically stable equilibrium point of \( x_s \) with region of attraction \( \mathcal{X}_N \). The chosen value of \( \alpha(x, u) \) is given in Theorem 4 in [31]. The last paragraph of Section VI in [4] explains the modification to the stage cost in economic MPCs as an intermediary value lies between the original economic stage cost and the traditional target tracking MPC stage cost, adequately chosen.

### 3.2.4 Extensions to unsteady closed-loop processes

Periodic terminal constraint is considered for systems having pre-computed optimal periodic solutions, while average constraints are introduced to apply to state and input averages of unsteady closed-loop processes. The periodic terminal constraint, obtained from a dynamic state feedback strategy, is introduced in order for achieving a periodic optimal solution described in a previous work [3]. Reasoning for the average constraints is clearly given in [4]. The authors stated that, “shifting the focus from convergence to average performance leads naturally to the consideration of constraints on average values of variables (typically inputs and states), besides point-wise in time hard bounds as discussed in the previous sections and customary in MPC” [4].

According to [3], when

\[
\text{Av}[\ell(x, u)] \subseteq [\ell(x_s, u_s), +\infty),
\]

the system is said optimally operated at steady state. And if, in addition, one or both of the conditions of

1. \( \text{Av}[\ell(x, u)] \subseteq (\ell(x_s, u_s), +\infty) \),

2. \( \lim_{k \to \infty} \inf \|x(k) - x_s\| = 0 \)

hold, it is said sub-optimally operated off steady state.

The economic MPC problem with average constraint is to obtain the following:

\[
\text{Av}[\ell(x, u)] \subseteq \{ -\infty, \ell(x_s, u_s) \}, \quad (5)
\]

\((x, u) \in \mathbb{Z} \forall k \geq 0, \) and \( \text{Av}[y] \subseteq \mathbb{Y} \),

where \( y = h(x, u) \in \mathbb{Y} \) is an auxiliary output. \( \mathbb{Y} \), the chosen average constraint set, is defined as \( \mathbb{Y} \subset \mathbb{R}^p \) and required to contain \( h(x_s, u_s) \), be closed and convex. On the ground of (5), the economic MPC with average constraint is thus sub-optimally operated off steady state.

In the next section, we show that the previously developed quadratic dissipativity constraint [35, 38, 37, 39] can be conveniently apply to the economic MPC problem.
4 Quadratic Dissipativity Constraint for Economic MPC

The quadratic dissipativity constraint for economic MPC is presented in this section. In this proposed stabilisation approach, the storage function and supply rate of the dissipation inequality are independent to the stage cost of economic MPC.

Define a parameterised quadratic supply rate for \( \Sigma (1) \), as follows:

\[
\xi(u, x | Q, S, R) := u^T R u + 2 x^T S u + x^T Q x, \tag{6}
\]

where \( Q, S, R \) are multiplier matrices with symmetric \( Q, R \). The input-state pair \((u, x)\) of \( \Sigma \) is said to satisfy the quadratic dissipativity constraint (QDC) if there exists a function \( \alpha \) of class \( \mathcal{KL} \), and \( k_0 \geq 0, k_0 \in \mathbb{R} \), such that

\[
|\xi(k)| \leq \alpha(|\xi(1)|, k - 1) \quad \forall k \geq k_0. \tag{7}
\]

For implementation, it is of interest to consider \( \xi(k) \) in its half plane. The following inequalities are used instead:

For \( \xi(1) > 0 \), \( 0 \leq \xi(k) \leq \beta \xi(k-1) \quad \forall k \geq 1, 0 < \beta < 1. \tag{8} \)

For \( \xi(1) < 0 \), \( 0 \geq \xi(k) \geq \beta \xi(k-1) \quad \forall k \geq 1, 0 < \beta < 1. \tag{9} \)

The convex quadratic constraint w.r.t. \( u \) (10) below is equivalent to (8) when \( R > 0 \):

\[
u^T R u + 2 S u + \psi \leq 0, \tag{10}
\]

where \( \psi = x^T Q x - \delta(k-1), \delta(k-1) = \beta \xi(k-1), \beta < 1. \)

The constraint (10) will be imposed on the economic MPC optimisation (the original one) as an enforcing stability constraint. For the closed-loop stabilisability of \( \Sigma \), the input \( u \) is required to be bounded by the above quadratic dissipativity constraint (10), in association with the quadratic dissipativity of the open-loop \( \Sigma (1) \) w.r.t. the supply rate \( \xi(\ldots) \).

\( \Sigma (1) \) is said to be quadratically dissipative w.r.t the quadratic supply rate \( \xi(u, x) \), if there exists a nonnegative storage function \( V(x) \) such that for all \( x(k) \) and all \( k \in \mathbb{Z}^+ \), the following dissipation inequality is satisfied irrespectively of the initial value of the state \( x(0) \):

\[
V(x(k+1)) - \sigma V(x(k)) \leq \xi(u, x), 0 < \sigma < 1. \tag{11}
\]

\( V(x) = x^T P x, P > 0 \) is considered herein.

**Proposition 1**: \( \Sigma : x^+ = Ax + Bu \) is quadratically dissipative w.r.t the quadratic supply rate \( \xi(u, x) \) if the following LM1 is feasible in \( P, Q, S, R \), see, e.g. [12]:

\[
\begin{bmatrix}
P & PA & PB \\
* & P + Q & S \\
* & * & R
\end{bmatrix} > 0, \; P > 0. \tag{12}
\]

The stabilisability condition for unconstrained system \( \Sigma \) is then stated in below proposition.

**Proposition 2**: Let \( \xi(0) > 0 \), and \( 0 < \sigma < 1 \). Consider \( \Sigma \) without control and state constraints. Suppose that the following optimisation is feasible:

\[
\begin{aligned}
\min_{P, Q, S, R} & \quad x_0^T Q x_0 \\
\text{subject to} & \quad (12), \quad Q < 0, \quad R > 0;
\end{aligned}
\]

Then any \( u(k) \) feasible to (10), employing the resulting multiplier matrices \( Q, S, R \), stabilise \( \Sigma \).

The stabilisability theorem for nonlinear input-affine system \( \Sigma (1) \) with an extended supply rate has been developed in [38]. On the ground of this stabilisability condition, the economic MPC in association with the stability constraint (10), i.e. the MPC optimisation (2) is added with a new inequality constraint of (10), will assure the stability of the closed-loop system. The feasibility of (10) for constrained problems has been presented in previous works, see, e.g. [34, 36]. Since the MPC stage cost is not employed as a Lyapunov function in our approach, the quadratic dissipativity constraint allows both non-monotonic decreasing storage function \( V(x) \) and economic MPC stage cost \( \ell(x, u) \).

From the dissipativity perspective, the supply rate is different to those in [31]. The distance stage cost is not associated with the supply rate in this QDC approach, but the classical quadratic function w.r.t input and output. Its use is justified by the fact that the inclusion of the product \( x^T u \) in quadratic supply rates is perceived as less conservative than the small-gain type supply rate, see, e.g. [13]. Nevertheless, the average performance and constraint feasibility will need to be addressed thoroughly in future developments. Developments for the dynamically optimal operating point or rotated steady state with the quadratic dissipativity constraint is underway.

5 Conclusion

A review on the economic MPC (EMPC) was given in this paper. EMPCs have been found effective in various energy-efficient applications. The unreachable setpoints and periodic steady states, as well as modified stage cost for stability purpose and average performance are the essentials of these emerging EMPCs. The potential of applying the quadratic dissipativity constraint to the economic MPC has also been shown in the last section. From this review, we can, indeed, conclude that the economic MPC has merely started progressing both theoretically and practically.
References


Infrastructure Robotics: Research Challenges and Opportunities

D.K. Liu, G. Dissanayake, J. Valls Miro, K.J. Waldron,
Centre for Autonomous Systems, University of Technology Sydney, Australia
E-mail: {dikai.liu, gamini.dissanayake, jaime.vallsmiro, kenneth.waldron}@uts.edu.au

Abstract -
Infrastructure robotics is about research on and development of methodologies that enable robotic systems to be used in civil infrastructure inspection, maintenance and rehabilitation. This paper briefly discusses the current research challenges and opportunities in infrastructure robotics, and presents a review of the research activities and projects in this field at the Centre for Autonomous Systems, University of Technology Sydney.

Keywords -
Infrastructure robotics; autonomous robots; inspection; maintenance; civil infrastructure;

1 Introduction

Civil infrastructure in sectors such as transport (e.g. bridges, roads and tunnels), energy (e.g. power transmission line, transmission tower, power plant, gas/oil pipeline and offshore structure) and water (e.g. pipeline and dam) is progressively deteriorating, due to ageing, environmental factors, increased loading, damages caused by human/natural factors, and inadequate/poor maintenance. Appropriate operations of periodic inspection, maintenance and rehabilitation are needed to ensure that the designed life of service of civil infrastructure can be achieved or extended. However, because of the associated employee health and safety issues, these operations are very expensive undertakings.

A steel bridge is one of the many examples of civil infrastructure. There are approximately 42,000 steel bridges in Europe, and 210,000 and 270,000 steel bridges, respectively in the USA and in Japan [1]. Steel bridges such as the Sydney Harbour Bridge are very complex structures. Some sections (e.g. truss joints and box girders) in steel bridges cannot be inspected without special equipment because of the difficulty of access. Human inspection requires equipment (e.g. special lifts, scaffolds) and a team to support each inspector, which implies high cost and low productivity. Grit-blasting to remove rust and old protective coatings followed by repainting is the common approach in steel bridge maintenance, but grit-blasting is extremely labour intensive and hazardous.

Therefore, supplementing manual labour in civil infrastructure inspection, maintenance and rehabilitation with robotic aids has potential for significant safety, cost and health impacts.

There is a significant amount of research on infrastructure robotics and development of robotic systems for practical applications. Examples include climbing robots for maintenance and inspection of vertical structures [2]; Shady3D robot for climbing trusses [3], climbing robot that can move in complex 3-D metallic-based structures [1][4]; the SM2 robot designed to walk along the I-beam structure of a space station [5][6]; robot-aided tunnel inspection and maintenance system [7]; and autonomous grit-blasting robot for surface preparation in steel bridge maintenance [9][10]. Robotic systems have also been deployed in various operations in construction [8] and material handling.

Although there are many robots that have been developed for civil infrastructure construction and maintenance, most of these robots are either laboratory prototypes that have not yet made it into the commercial world, or are limited to specific applications. These robots may not have enough intelligence and dexterity for performing various tasks autonomously in complex environments. Developing autonomous robots that can operate automatically for civil infrastructure inspection, maintenance and rehabilitation is challenging because the structural environment is normally complex, compact, unknown (to the robots), hazardous, and in field conditions (Figure 1). There are many research and engineering challenges that need to be addressed. A usual reaction is that, ‘at least it’s a structured

Figure 1. Two example infrastructure environments ((b) from [11])
environment’, but old steel structures have usually not been digitally modelled, and any blueprints that are available are out of date. The bottom line is that the robot has to be able to build its own solid model of the environment, be aware of the conditions of the structures, plan its collision-free motion and perform various tasks automatically.

2 Infrastructure Robotics Research

Infrastructure robotics is about research on and development of methodologies that enable robotic systems to operate autonomously in civil infrastructure inspection, maintenance and rehabilitation. Deploying autonomous robotic systems in such applications has specific research challenges. These challenges include, but are not limited to:

1. Sensing technology and sensor network that can be used in field environments for sensing, inspection, and assessment of surface and structural condition of civil infrastructure;
2. Robot environmental awareness, localisation and mapping of true 3D complex environments such as trusses;
3. Robust and reliable mechanisms for adhesion and grasping that allow robots to stay and move along various structures (e.g. trusses) safely;
4. Efficient algorithms for real-time robot motion planning and collision avoidance in true 3D environments;
5. Methodology (e.g. co-design between end users and researchers) for robot and user interface design;
6. Development of robot teams and human-robot teams for applications in infrastructure construction, inspection and maintenance;
7. Interaction strategies for intuitive physical human-robot interaction;
8. Prototype autonomous robots that have the required payload for conducting various operations such as blasting, cleaning, painting, and object manipulation, and can be used in different structures;
9. Uncertainty handling;
10. Machine learning that allows robots to adapt to various structural and field environments and conduct various tasks.

Besides the research challenges, there are many engineering challenges in developing robots for civil infrastructure inspection, maintenance and rehabilitation. Examples include:

1. Design of lightweight robots that can move and climb in compact and complex structures;
2. High torque to weight ratio actuators that allow robots to have enough payload for performing tasks such as blasting, painting and object handling;
3. Fail-safe robotic system design;
4. Protection of sensors and robots for dust and water proofing;
5. Methodology (e.g. co-design between end users and researchers) for robot and user interface design;
6. Development of robot teams and human-robot teams for applications in infrastructure construction, inspection and maintenance;
7. Interaction strategies for intuitive physical human-robot interaction;
8. Prototype autonomous robots that have the required payload for conducting various operations such as blasting, cleaning, painting, and object manipulation, and can be used in different structures;
9. Uncertainty handling;
10. Machine learning that allows robots to adapt to various structural and field environments and conduct various tasks.

3 Infrastructure Robotics Research Activities in the Centre for Autonomous Systems

The Centre for Autonomous Systems (CAS) at the University of Technology, Sydney (UTS) has been conducting infrastructure robotics research since 2003. Besides conducting fundamental research in two research themes: (1) robots in unknown and complex environments: sensing, localisation, mapping and motion planning; and (2) human-robot interaction: human models and control, CAS has been doing enabling research and development of robotic systems for various applications in collaboration with industry. Infrastructure robotics has been the centre’s flagship research, with $6.5 million worth of externally funded industry projects in infrastructure robotics since 2006.

Currently seven academic staff members, five research fellows, three research engineers and a number of PhD students are involved in these projects.

UTS CAS’s infrastructure robotics research focuses on applications in the following sectors:

- Transport: bridge inspection, condition assessment and maintenance;
- Water: underground pipe condition assessment;
- Energy: transmission tower inspection, condition assessment and maintenance;
- Ports: automation of container handling vehicles;
- Mining: underground mining vehicles;
- Underwater and offshore structure: semi-autonomous underwater robot for submerged structure inspection, condition assessment and maintenance.

3.1 Autonomous Robotics Systems for Steel Bridge Maintenance

This research focused on addressing research and engineering challenges in developing autonomous robots for steel bridge maintenance, i.e. autonomous grit-blasting (Figure 2). Theoretical research and enabling methodologies developed include: (1) Sensing technology for object ranging and surface classification [12][13]; (2) Algorithms that enable robotic manipulator-based exploration in complex 3D environments [14]; (3) Algorithms that can efficiently build 3D maps of complex steel structures [15][16][25]; (4) Algorithms for efficient robot motion and blasting planning [17][18]; (5) Collision detection and avoidance algorithms that make it possible for a large robotic manipulator to move safely in complex 3D environments.
environments, including the 3D force-field approach for collision avoidance [19] and an algorithm for collision detection and distance query [20]; (6) Effective methods for human-robot-environment interaction [21], including an extended hand-movement model [23], effect of view distance and movement scale on haptic-based operation [22], and workspace mapping; and (7) Ease of user-robot interaction, facilitated by an interface designed jointly by the project team and four bridge maintenance professionals [24].

3.2 Bio-logically Inspired Climbing Robots for Autonomous Inspection of Steel Structures

This research focuses on developing a biologically inspired climbing robot for inspection and condition assessment of steel structures where human inspectors cannot reach, or are not allowed access due to the compactness of structures and the OH&S requirement, such as box girder structure. Research work includes analysis of an arthropodal system [31] and ant locomotion [32], exploration methods [26][27], map building, motion planning, collision avoidance, surface inspection, robot design [28], adhesion mechanism design [29][30] and robot control. Two prototype robots have been constructed (Figure 4) with the first field trial conducted in September 2013. Three more field trials are to be conducted in June, September and November 2014. It is expected to deliver one fully functional robot for extensive testing in February 2015.

3.3 Concrete Bridge Box Girder Inspection

This research focuses on demonstrating the feasibility of automated solutions for the structural health inspection of concrete box girders [33]. Novel combinations of perceptual robotic sensing allows for highly-detailed surface condition pictures to be remotely evaluated by human operators while the robot navigates inside the enclosed structure. The proposed technique makes it feasible to safely monitor areas over a sequence of inspections, leading to more effective maintenance procedures. Functionality to automatically identify defects/regions of interest is also being investigated.
3.4 Coordination of Autonomous Vehicles for Automated Container Handling – Automation of Port Infrastructure

This research addressed several research issues associated with the use of a large fleet of autonomous robots/vehicles for automated container handling (Figure 6) in dynamic environments. Methodologies that are able to significantly improve the efficiency of multi-robot systems in dynamic container handling environments have been developed, including simultaneous task allocation and motion coordination methods [34], multi-objective optimisation-based scheduling [35], a crossover approach [36] and an genetic algorithm for job scheduling [37], a comprehensive mathematical model and a job grouping method [39], a parallel scheduling algorithm [38] and task allocation under uncertainty [40].

3.5 Robotic Technology for Condition Assessment of Buried, Pressurised Water Mains

This research focuses on developing methods for advanced analysis of sensor data obtained from condition assessment tools of buried large water mains (Figure 7) through physical and probabilistic modelling. The project is focused on non-destructive inspection techniques (NDT) generally based on electromagnetic and acoustic sensing with the ultimate aim of improving the accuracy and level of confidence in the estimate of the geometry of pipe being inspected. This, in turn, allows the utilities to evaluate the remaining life of this critical assets with more certainty in their estimates [41][42]. This is a project carried out in collaboration with other universities, water utilities in Australia, UK and USA, and the participation of major international providers of NDT inspection technologies.

3.6 Semi-autonomous Robot for Underwater Structure Cleaning and Inspection

This research focuses on developing a semi-autonomous robotic system that will enable safe and cost-effective cleaning and inspection of submerged structures. The system concept being considered is a semi-autonomous robotic system that can explore the environment around bridge pylons, build a surface condition map of the submerged bridge pylons, clean the pylons by using high pressure water blasters, and perform a detailed inspection after the cleaning process. Research work includes design of the robotic system that can be used for various submerged structures; surface condition monitoring; 3D map building; control of robot navigation under uncertainties; manipulation of underwater blasting tools; maintaining robot stability under water current and with the reaction force from the high pressure blasting nozzles; and blasting planning based on marine growth for maximum productivity and minimum damage to the pylon surface.
process; inspection of surface quality; control of the robotic system for blasting surface spots that are missed out in previous blasting operation or where the quality is not good enough; and coordination of multiple grit-blasting robots for cooperative blasting operation (Figure 8).

3.8 Human-interactive Robots for Strength Augmentation of Workers in Infrastructure Maintenance

This research aims to improve our understanding on how robots can be enabled to provide physical assistance to diverse human workers performing labour intensive tasks in infrastructure maintenance, and to develop methodologies that enable the development of assistive robotic systems (Figure 9). The research questions that are being answered include how much assistance is needed for a worker [43][44], how to provide the required assistance[45], what are the strategies for intuitive human-robot interaction and safety of human workers when they are in physical contact with robots.

![Figure 9. An 8-DOF scaled-down proof-of-concept prototype robot](image)

4 Conclusion and Remarks

Advances in sensing, mapping, actuation, locomotion and control have made it possible to build robotic systems that are able to operate in unstructured and cluttered outdoor environments. As a result, a number of complex tasks associated with the maintenance of civil infrastructure can now be carried out by robots. Main focus of the research effort to-date has been on addressing important occupational health and safety issues. Grit-blasting and inspection of complex steel structures are among the applications that have been described in this paper. All the indications are that robotics will play a significant role in maintaining civil infrastructure in the near future.

References


[29] Ward P. and Liu D.K., Design of a high capacity electro permanent magnetic adhesion for climbing


Abstract -
Bridge deck inspection is essential task to monitor the health of the bridges. This paper reports the data collection and analysis for bridge decks based on our novel robotic system which can autonomously and accurately navigate on the bridge. The developed robotic system can lessen the cost and time of the bridge deck data collection and risks of human inspections. The advanced software is developed to allow the robot to collect visual images and conduct nondestructive evaluation (NDE) measurements. The image stitching algorithm to build a whole bridge image from individual images is presented in detail. The impact-echo (IE) and ultrasonic surface waves (USW) data collected by the robot are analyzed to generate the delamination and concrete elastic modulus maps of the deck.

Keywords -
Mobile robotic systems, Bridge deck inspection, Image Stitching, Nondestructive evaluation.

I. Introduction
The condition of bridges is critical for the safety of the traveling public and economic vitality of the country. There are many bridges through the U.S. that are structurally deficient or functionally obsolete. Condition monitoring and timely implementation of maintenance and rehabilitation procedures are needed to reduce future costs associated with bridge management. Application of nondestructive evaluation (NDE) technologies is one of the effective ways to monitor and predict bridge deterioration. A number of NDE technologies are currently used in bridge deck evaluation, including impact-echo (IE), ground penetrating radar (GPR), electrical resistivity (ER), ultrasonic surface waves (USW) testing, visual inspection, etc. [5], [22]. For a comprehensive and accurate condition assessment, data fusion of simultaneous multiple NDE techniques and sensory measurements is desirable. Automated multi-sensor NDE techniques have been proposed to meet the increasing demands for highly-efficient, cost-effective and safety-guaranteed inspection and evaluation [7].

Automated technologies have gained much attention for bridge inspection, maintenance, and rehabilitation. Mobile robotic inspection and maintenance systems are developed for vision based crack detection and maintenance of highways and tunnels [18], [19], [23]. A robotic system for underwater inspection of bridge piers is reported in [3]. An adaptive control algorithm for a bridge-climbing robot is developed [15]. Additionally, robotic systems for steel structured bridges are developed [2], [16], [21]. In one case, a mobile manipulator is used for bridge crack inspection [20]. A bridge inspection system that includes a specially designed car with a robotic mechanism and a control system for automatic crack detection is reported in [11], [12], [17]. Similar systems are reported in [13]–[15] for vision-based automatic crack detection and mapping and [24] to detect cracks on the bridge deck and tunnel. Edge/crack detection algorithms such as Sobel and Laplacian operators are used.

Difference to all of the above mentioned works, our paper focus on the bridge deck data analysis which is collected by our novel robotic system integrated with advanced NDE technologies. The developed data analysis algorithms allows the robot to build the entire bridge deck image and the global mapping of delamination and elastic modulus of the bridge decks. These advanced data analysis algorithms take into account the advantages of the accurate robotic localization and navigation to provide the high-efficient assessments of the bridge decks.

The paper is organized as follows. In the next section, we describe the robotic data collection system and coordinate transformation. In Section III we present the image stitching algorithm and bridge deck viewer/monitoring software. In Section IV, we present
the IE and USW data collection and analysis. Finally, we provide conclusions from the current work and discuss the future work in Section V.

II. The bridge robotic inspection system

A. Data Collection

The robotic system is integrated with several non-destructive evaluation (NDE) sensors such as Ground Penetrating Radar (GPR), acoustic array consisting of Impact Echo (IE) and Ultrasonic Surface Waves (USW), Electrical Resistivity (ER), and high resolution cameras as shown in Fig. 1. The robot autonomously maneuvers on the bridge based on the advanced localization and navigation algorithm reported in our previous works [8]–[10]. The NDE data collection system is run on two Windows operating computers and communicate with the robot Linux operating computer through the serial communication protocols. The NDE software is developed by utilizing Qt development kit and Cpp to enable the robot to collect and monitor the data simultaneously. The software architecture is designed based on multi-thread programming. The software consists of five slave threads and one master thread. The master thread controls the entire user interface. The slave threads are:

- Robot thread which communicates with LinuxWindowsSerial program in the robot computer (Linux / ROS) using RS-232 protocol and sends position information of robot to the user interface;
- Acoustic thread which controls the data acquisition of the acoustic device consisting of IW and USW using USB protocol and logs the time series data;
- GPR thread which communicates with IDS vendor software using TCP/IP protocol. GPR thread is able to start, stop, and receive stream data from the GPR acquisition device;
- Camera thread which uses the Canon SDK protocol to control the camera such as triggering to shoot, changing lighting parameters, and downloading collected images;
- Electrical Resistivity (ER) thread which communicates with Resipod sensor using RS-232 protocol and logs the resipod data.

Overall, the robot thread controls the other threads to trigger and sync the data collection system. During the operation, the robot thread waits for a serial message from robot Linux computer. When the serial message is received, it will be used to command the other NDE thread to perform the data collection. The data flow of the NDE GUI is shown in Fig. 2. The serial message also consists of the robot position and orientation, and number of line inspections and their indices.

B. NDE Coordinate Transformations

This subsection presents coordinate transformations in the robotic system which allows the NDE data analysis and mapping process. Since the relationship between the GPR, Acoustic, ER coordinates and the robot coordinate are fixed, we just present the transformation from camera frame to robot frame which allows the image stitching and crack mapping process to map from the local image coordinate to the world coordinate.

The system involves four coordinate systems as shown in Fig. 3. They are: image coordinate system ($F_I$), camera coordinate system ($F_C$), robot coordinate system ($F_R$) and world coordinate system ($F_W$). To transform the image coordinate system ($F_I$) to the world coordinate system ($F_W$), we need to implement the sequential transformations: $(X_{im}, Y_{im})^{F_I} \rightarrow (X_c, Y_c)^{F_C} \rightarrow (X_r, Y_r)^{F_R} \rightarrow (X_w, Y_w)^{F_W}$.

The intrinsic and the extrinsic matrices are obtained once the calibration is finished. The intrinsic matrix consisting of focal length ($f$), skew value ($s$) and the origin of image coordinate system $(x_{im}(0), y_{im}(0))$ is described in Equ. (1).

$$P = \begin{bmatrix} sf & 0 & x_{im}(0) & 0 \\ 0 & f & y_{im}(0) & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$ (1)

The extrinsic matrix consists of rotation and translation parameters as in Equ. (2).

$$M = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4x4}$$ (2)

here $R$ is a $3 \times 3$ rotation matrix which can be defined by the three Euler angles [6], and $T = [t_x, t_y, t_z]^T$ is the translation between two frames.

We have the following transformation from the image
Steel rebars are detected by the GPR mounted on the robot.

Fig. 2. The GUI for NDE data collection and monitoring of a bridge near Chicago, Illinois, USA.

Fig. 3. Coordinate systems in the robotic bridge deck inspection system.

\[
\begin{bmatrix}
x_{im} \\
y_{im} \\
1
\end{bmatrix} = I_{TC} \times \begin{bmatrix}
x_c \\
y_c \\
z_c
\end{bmatrix},
\]

(3)

here \(I_{TC}\) is the transformation matrix from the image coordinate to the camera coordinate, and \(I_{TC} = PM\).

Now we can find the camera coordinate corresponding to the image coordinate using pseudo-inverse as

\[
C = (I_{TC}^T I_{TC})^{-1} I_{TC}^T I.
\]

(4)

here \(I_{TC}^T\) is the transpose of \(I_{TC}\).

To find the transformation from \((F_C)\) to \((F_R)\), we use the static relationship between these two coordinates.

Namely, the relationship between the camera and robot coordinate systems is fixed because the camera orientation is fixed (see Fig. 3). Therefore the transformation from \((F_C)\) to \((F_R)\) can be obtained by measuring the physical offset distances between the robot center and the camera pose. This transformation can be described as:

\[
\begin{bmatrix}
x_{cr}^T \\
y_{cr}^T
\end{bmatrix} = \begin{bmatrix}
x_c^T \\
y_c^T \end{bmatrix} - \begin{bmatrix}
\Delta x_{cr} \\
\Delta y_{cr}
\end{bmatrix},
\]

(5)

where \(\Delta x_{cr}\) and \(\Delta y_{cr}\) are the offset distances between the camera coordinate and the robot coordinate along \(x\) and \(y\), respectively.

Finally, to find the transformation from \((F_R)\) to \((F_W)\) we use the following relationship:

\[
\begin{bmatrix}
x_w \\
y_w \\
1
\end{bmatrix} = R_{TW} \times \begin{bmatrix}
x_{tran}^T \\
y_{tran}^T \\
1
\end{bmatrix},
\]

(5)

where the transformation matrix \(R_{TW}\) is defined as

\[
R_{TW} = \begin{bmatrix}
\cos(\theta_r) & -\sin(\theta_r) & x_r \\
\sin(\theta_r) & \cos(\theta_r) & y_r \\
0 & 0 & 1
\end{bmatrix},
\]

(6)

where \((x_r, y_r, \theta_r)\) are the position and heading of the robot obtained by the Extended Kalman Filter (EKF) [10].

III. Bridge Deck Image Stitching

A. Bridge Deck Image Stitching

For the ease of bridge deck inspection and monitoring, we combine taken photos into a single large image as shown in Fig. 4. This is a specific case of the general image stitching problem. In image stitching problem,
camera motion is unknown and not constrained and intrinsic camera parameters can change between the given images. In our specific problem of bridge deck surface image stitching, we benefit from constraints we know to exist due to the nature of the problem and the setup of the hardware system. We have two identical cameras that simultaneously take images of different area, appearance-based template matching can give finer estimation of the camera motion. If the robot motion is not accessible or not accurate enough, overlapping area can be searched over the whole image, which is a more time consuming process.

To reduce the tremendous amount of data to be processed, we resort to multi-resolution pyramidal search method [4], where we search for a larger motion range in lower resolution image and reduce the possible motion range for higher resolution image. Due to possible large illumination and reflection changes between different frames, we use image comparison method Normalized Correlation Coefficient (8) that is less illumination independent. In Eq. (8) correlation coefficient for each location \(x, y\) is denoted by \(R(x, y)\), where search image region is \(I\), template image that is searched is \(T'\) and normalized versions of them are \(I'\) and \(T'\) respectively. We compare the grayscale versions of the images to get rid of any white-balance effects in different images.

\[
R(x, y) = \frac{\sum_{x', y'} I'(x', y')I(x+x', y+y')}{\sqrt{\sum_{x', y'} I'(x', y')^2 \sum_{x', y'} I(x+x', y+y')^2}}
\]

here, \(w\) and \(h\) are the width and height of the image \(I\), respectively.

2) Exposure Compensation and Blending

Exposure compensation step obtains the most blending exposures for each image by selecting the suitable brightness ratio of overlapping area between images. When combining existing image and the new arrived image, we are performing an image-blending step to remove the shadows in the image (9). If the new arriving pixel is considerably brighter than the existing pixel in the same location, we replace the pixel with the new one. A threshold value of 0.7 is used for \(th\) to indicate being considerable is brighter than corresponding pixel.

\[
I(x, y) = f(x) \begin{cases} 
I_2(x, y), I_2(x, y) \ast th > I_1(x, y) \\
I_1(x, y), \text{else.}
\end{cases}
\]

(9)

Figure 5 shows the image stitching result. There are 200 images collected by the two front surface cameras.

B. Bridge Deck Monitor

The bridge deck viewer (BDV) software (see Fig. 6) is developed using Java language to support the bridge engineer to monitor the bridge decks in an efficient way. The stitched images are first loaded and then calibrated to map to the bridge coordinate as Fig. 7. The BDV
software can find the crack locations on the surface of bridge in the viewing image and allows to mark them for the next view or any purpose by left mouse click on that locations. The BDV also shows the notification about the position of the cracks. As can be seen in Fig. 8, the flags appear at the crack locations corresponding with coordinates (x, y) on the bridge deck.

Additionally, the BDV software allows to measure the distance of crack on the deck by right mouse click on the starting point and drag the hold right mouse to the last point. A line that connects the starting point and ending point appears to show the length of the crack as shown in Fig. 8.
IV. IE and USW Data Collection and Analysis

This section presents impact-echo (IE) and ultrasonic surface waves (USW) data collection and analysis. The robot is equipped with two acoustic arrays, and each array can produce 8 IE and 6 USW data sets as shown in Fig. 9. These raw data sets are collected by the robot at every two feet (60 cm) on the bridge deck.

A. Impact-Echo (IE) Data Analysis

Impact-Echo (IE) is an elastic-wave based method to identify and characterize delaminations in concrete structures. This method uses the transient vibration response of a plate-like structure subjected to a mechanical impact. The mechanical impact generates body waves (P-waves or longitudinal waves and S-waves or transverse waves), and surface-guided waves (e.g., Lamb and Rayleigh surface waves) that propagate in the plate. The multiple-reflected and mode-converted body waves eventually construct infinite sets of vibration resonance modes within the plate. In practice, the transient time response of the solid structure is commonly measured with a contact sensor (e.g., a displacement sensor or accelerometer) coupled to the surface close to the impact source. The fast Fourier transform (amplitude spectrum) of the measured transient time-signal will show maxima (peaks) at certain frequencies, which represent particular resonance modes as shown in Fig. 10.

There are different ways of interpreting the severity of the delamination in a concrete deck with the IE method. One of the ways used in this study is shown in Fig. 11. A test point is described as solid if the dominant frequency corresponds to the thickness stretch modes (Lamb waves) family. In that case, the frequency of the fundamental thickness stretch mode (the zero-group-velocity frequency of the first symmetric ($S_1$) Lamb mode, or also called the IE frequency ($f_{IE}$)). The frequency can be related to the thickness of a plate $H$ for a known $P$-wave velocity $C_p$ of concrete by

$$H = \frac{\beta_1 C_p}{f_{IE}}$$

where $\beta_1$ is a correction factor that depends on Poisson’s ratio of concrete, ranging from 0.945 to 0.957 for the normal range of concrete. A delaminated point in the deck will theoretically demonstrate a shift in the thickness stretch mode toward higher values because the wave reflections occur at shallower depths. Depending on the extent and continuity of the delamination, the partitioning of the wave energy reflected from the bottom of the deck and the delamination may vary. The initial or incipient delamination, described as occasional separation within the depth of the slab, can be identified through the presence of dominant frequencies associated with the thickness stretch modes from both the bottom of the deck and the delamination. Progressed delamination is characterized by a single peak at a frequency corresponding to the depth of the delamination. Finally, in case of wide or shallow delaminations, the dominant response of the deck to an impact is characterized by a low frequency response of flexural-mode oscillations of the upper delaminated portion of the deck.

B. Ultrasonic Surface Waves (USW) Data Analysis

The ultrasonic surface waves (USW) technique is an offshoot of the spectral analysis of surface waves (SASW) method used to evaluate material properties...

---

**Fig. 9.** Acoustic/seismic array sensor is developed and integrated with the robot to collect IE and USW data.

**Fig. 10.** (a) IE raw data of one channel, C4, in the acoustic array in time domain, and (b) After using Fast Fourier Transform (FFT) in frequency domain.
Fig. 11. Grades for various degrees of deck delamination for IE method.

Fig. 12. Schematic of evaluation of a layer modulus by SASW (USW) method.

(elastic moduli) in the near-surface zone. The SASW uses the phenomenon of surface wave dispersion (i.e., velocity of propagation as a function of frequency and wave length, in layered systems to obtain the information about layer thickness and elastic moduli) as shown in Fig. 12. A SASW test consists of recording the response of the deck, at two receiver locations, to an impact on the surface of the deck. The surface wave velocity can be obtained by measuring the phase difference $\Delta \phi$ between two different sensors (sensor 1 and sensor 2) as follows,

$$C = 2\pi f \frac{d}{\Delta \phi} \quad (11)$$

where $f$ is frequency, $d$ is distance between two sensors. The USW test is identical to the SASW, except that the frequency range of interest is limited to a narrow high-frequency range in which the surface wave penetration depth does not exceed the thickness of the tested object. In cases of relatively homogeneous materials, the velocity of the surface waves does not vary significantly with frequency. The surface wave velocity can be precisely related to the material modulus, or concrete modulus in the case of bridge decks, using either the measured or assumed mass density, or Poisson’s ratio of the material. In the case of a sound and homogenous deck, the velocity of the surface waves will show little variability. An average velocity is used to correlate it to the concrete modulus. Significant variation in the phase velocity will be an indication of the presence of a delamination or other anomaly.

The IE condition map is presented in Fig. 13-top. The test regions classified as serious condition (red color) are interpreted as likely delaminated areas on the concrete deck. The USW condition map is presented in Fig. 13-bottom. The USW map provides the condition assessment and quality of concrete through measuring concrete modulus. The presented modulus plot indicates that zones of very low moduli provide a good match with delaminated zones identified by other methods.

Acknowledgment

This work was supported by the Federal Highway Administration’s Long Term Bridge Performance (LTBP) Program. The authors would like to thank Profs. Basily Basily, Kristin Dana, Jingang Yi and Ali Maher of Rutgers University for their support for the project development. The authors are also grateful to Ronny Lim, Turgay Senlet, Hooman Parvardeh, Kenneth Lee, Prateek Prasanna, Yizhai Zhang, Moiz Ezzy and Fei Liu of Rutgers University for their help during the system development and field testing.

V. Conclusions and Future Work

The bridge deck data collection and analysis has been reported in this paper. Several challenging problems of data collection software, image stitching, IE and USW analysis have been tackled. The image stitching algorithm allowed to generate a very high resolution image of the whole bridge deck, and the bridge viewer software allows to calibrate the stitched image to the bridge coordinate. The delamination and elastic modulus maps were built based on IE and USW data collected by the robot to provide easy evaluation and condition monitoring of bridge decks. Extensive testings and deployments of the proposed system on a number of bridges proved the efficiency of the new approach for bridge deck inspection and evaluation.

In the future work we will include development of a fusion algorithm for the NDE sensor and camera data for a more comprehensive and intuitive bridge deck condition assessment data presentation.
Fig. 13. The impact-echo (IE) and ultrasonic surface waves (USW) condition maps of a bridge deck in Illinois, USA, based on data collected by the developed robot system. The robot covers a half of the bridge with two scans (6ft width/scan) and 100ft along within 25 minutes.

References


Demand Response Under the Smart Grid Paradigm

M. Negnevitsky
School of Engineering and ICT, University of Tasmania, Australia
E-mail: Michael.Negnevitsky@utas.edu.au

Abstract -
Electric power industry is undergoing a profound change. The change is driven by technical, economic and environmental factors. The emerging challenges are particularly significant for distribution grids, where the level of automation or “smartness” is relatively low. With the push for energy conservation, demand-side management (DSM) and demand response (DR) are becoming vital tools under the smart grid paradigm. This paper outlines some experience obtained at the University of Tasmania, Australia in developing DSM and DR systems.

Keywords -
Smart Grid; Distribution System; Demand-side Management; Demand Response.

1 Introduction

Electric power systems are undergoing a profound change. This change is driven by several factors that include technical, economic and environmental factors. We need to deal with an aging infrastructure of power systems and maintain the required level of grid reliability. We need to integrate renewable energy sources, particularly wind and solar, and provide secure power supply to our customers, and at the same time improve operational efficiency. The emerging changes and challenges are particularly significant for distribution grids, where the level of automation or “smartness” is relatively low. Manual and “blind” operations along with old electromechanical relays are to be transformed into a “smart grid”. This transformation is necessary to meet environmental targets, accommodate distributed generation, and support plug-in electric vehicles. In fact, these needs present the power industry with the biggest challenge it has ever faced. On one hand, the transition to the “grid of the future” has to be evolutionary – we still need to supply electricity to our customers to keep the lights on. On the other hand, the challenges associated with the smart grid are significant enough to expect revolutionary changes in power system design and operation.

With the push for energy conservation, demand-side management and demand response are becoming vital tools under the broad smart grid paradigm.

The term “demand-side management” (DSM) was first introduced by Electric Power Research Institute (EPRI) in the 1980s, and since then has been widely used around the world. In fact, DSM is a term that implies many activities such as direct load control, peak shaving, peak shifting, and various load management strategies. Effective load management programs are often referred to as demand response (DR). According to the US Federal Energy Regulatory Commission, DR is defined as:

“Changes in electric usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.”

This paper outlines some experience obtained at the University of Tasmania, Australia in developing DSM and DR systems. Section 2 presents an evaluation tool for DSM of domestic hot water systems in distribution grid, and Section 3 discusses the development and implementation of fast DR in isolated power systems.

2 Demand-side Management Evaluation Tool

Effective implementation of DSM programs delivers operational benefits such as reduced peak demands and relieved overloads, which are essential in a power system with growing penetration of fundamentally intermittent renewable energy sources [1]. Successful DSM programs also provide economic gains such as deferrals of costly network upgrades as well as network security enhancements [1]. Moreover, in a deregulated electricity market, DSM programs offer opportunities for aggregation of demand reduction to support market and network operations of a power system [2]. In addition, consumers receive financial incentives through participation in DSM programs.

There are three different methods to implement DSM in a power distribution network. In indirect load control, consumers manually adjust their consumption in response to incentive programs such as time-of-use
(TOU) tariffs [3]. In autonomous load control, devices autonomously adjust their consumption in response to detected changes in the power system or to commands sent from the control centre. In direct load control (DLC), devices are centrally controlled by the utility operator [4].

Hot water load forms a significant share of the total domestic demand. For example, it accounts for up to 40% of domestic energy consumption in Australia, and around one third in Tasmania [5], [6]. Moreover, domestic hot water systems represent insulated thermal energy storages that continually supply hot water even during periods of power interruption. Hence, they are commonly targeted for DSM programs to reduce peak loads and improve the load factor. Well-designed DSM programs minimize customer discomfort due to cold showers.

This section presents the development of an evaluation tool that assists in designing a DSM program to deliver desired peak load reductions while maintaining satisfactory level of comfort for all customers. The tool estimates the available domestic hot water loads in a controlled area, and determines optimal switching programs. A switching program refers to a direct load control schedule applied to domestic hot water systems (to strategically switch them on and off) in order to achieve a desired load reduction during peak periods.

### 2.1 Structure of the Tool

Main modules of the tool are shown in Figure 1. The modules are grouped in three main functional blocks. The numbered grey circles represent inputs and outputs (I/O).

![Figure 1. The structure of the DSM evaluation tool.](image)

The Input block represents the user interface, which allows the tool user to enter parameters required for simulation (the number of households in the controlled area, the number of Monte Carlo simulations, the desired peak reduction, etc.) as well as to view default parameters and change them if necessary. The Simulation block is the main block of the tool; it contains four modules: the hot water consumption generator, hot water cylinder model, switching program optimizer, and performance calculator. The Output block contains the exporter, which exports the data to an external (Excel) file.

Default parameters and parameters entered by the user via the user input interface are represented by I/O 1. The hot water consumption generator receives I/O 1 and determines hot water consumption profiles for individual households; these profiles are represented by I/O 2. The hot water cylinder model receives I/O 2 and calculates uncontrolled hot water loads and shower temperatures for the households; the results are represented by I/O 3. The user can observe the aggregate uncontrolled load curve of the households in the controlled area, and proceed with the optimization of switching programs. The switching program optimizer receives I/O 3 and produces switching programs based on the user-defined parameters (the desired peak reduction target, control periods etc.). The best switching programs are presented to the user, so that he/she can select the most suitable switching program. The hot water cylinder model then calculates controlled hot water loads (I/O 5) by applying the user-selected switching program (I/O 4) to the hot water consumption profiles (I/O 2). The performance calculator receives I/O 5 and determines key performance indicators such as peak reductions and customer’s comfort. Results in the form of 24-hour load curves are presented to the user (I/O 6), and exported to an external file (I/O 7) via the exporter.

### 2.2 Hot Water Consumption Generator

The first step in the development of the hot water consumption generator was to acquire knowledge of hot water consumption patterns of households in the controlled area. To achieve this objective, a telephone survey was conducted on 1000 randomly selected households across Tasmania. It recorded demographic data (e.g. number of usual residents, combined income etc.) and details of hot water usages (e.g. average number of showers per day, average shower length etc.) of the surveyed households. This survey focused on two peak periods in the Tasmanian power distribution network, i.e. morning and evening peaks from 6am to 10am and from 5pm to 8pm, respectively. Figure 2 and Figure 3 show major results of the survey. Figure 2 suggests a positive correlation between the average number of showers and the family size, in the morning and evening peaks. An unexpected drop in the average number of morning showers in households with six or more residents can be explained by a relatively small
sample size of this household type (just 2.3% of the total surveyed households).

Figure 2. The average number of showers versus family size.

Figure 3. The histogram of average lengths of showers.

As can be seen in Figure 3, the length of a shower can vary from 2 min to 15 min, however, a great majority of showers (about 51%) last from 5 min to 8 min.

To estimate domestic hot water consumption profiles, we also acquired energy metering data of 279 households across Tasmania. These data were obtained from meters dedicated for metering electric water heating alone, and represented water heating energy consumption of individual households recorded in 5-minute intervals. We considered two types of hot water usages: high volume usage that lasts for more than 5 min and low volume usage that lasts for 5 min or less. Based on the modelling, 1 min of hot water usage requires approximately 10 min of heating to restore the temperature set by the thermostat. Thus, a continuous energy consumption (a switched-on condition of the electric water heater) for a period of more than 50 min is regarded as a high volume usage. Using weekday data only, we derived probability distributions of the starting time for showers (Figure 4) and low volume usages (Figure 5).

Figure 4. The probability distribution of the shower starting time.

Figure 5. Probability distribution of the starting time for low volume usages.

Both survey results and energy metering data revealed that domestic hot water consumption depends mostly on the family size. Therefore, all households in a controlled area are divided into four groups according to the family type based on the number of residents in a household. Table 1 shows a typical distribution of families in a controlled area.

We need also specify probabilities of household occupants taking morning showers only, evening showers only, or both. Demographic data [7] and household energy consumption records are used to estimate probabilities in Table 2, which determine the number of showers each family type take in the morning, evening, or morning and evening. Similar to showers, the probability of a low volume usage depends on the family size of a household. The tool uses multipliers to scale this probability up based on the family type. Default values of the multipliers are 1.0, 1.2, 1.6 and 2.0, respectively. The tool user can redefine these values, if required. Figure 5 gives the probability of a low volume usage occurring in a household at a given time.
Shower lengths and gaps between consecutive showers are specified by their mean, maximum and minimum values. We define minimum and maximum to discard unrealistic values (e.g., a one-minute shower) in probabilistic simulations. Normal distributions are assumed. Default values used by the tool are shown in Table 3. A low volume usage is denoted as a single 5-min draw. If required, the user can redefine these values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min (min)</th>
<th>Max (min)</th>
<th>Mean (min)</th>
<th>Standard deviation (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shower length</td>
<td>5</td>
<td>15</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Shower gap</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

The starting time of each hot water usage is specified based on probability distributions derived from actual energy metering data.

The tool employs a Monte Carlo approach to generate hot water consumption profiles for each household. First, the tool generates random values to determine specific parameters for a single household: family type, when showers are taken (morning, or evening, or morning and evening), number of showers, number of low volume usages, length of each shower and each gap between consecutive showers, starting time for each shower and each low volume usage. Next, using these parameters, the tool generates a 24-hour hot water consumption profile for a single household. The tool then repeats the profile generation process for a specified number of households using a new set of random values each time. Finally, the whole process is repeated for the required number of Monte Carlo iterations. Based on the generated hot water profiles, we can now proceed with calculating loads associated with household hot water usages. However, we need to develop a hot water cylinder model first.

2.3 Hot Water Cylinder Model

The block diagram of a DEHW system with a single heating element is shown in Figure 6.

For predicting the shower temperature and power consumption for domestic hot water systems accurately, we develop a hot water cylinder model based on the most common domestic hot water system in Tasmania, which has a 165 L cylindrical storage tank and a single 2.4 kW heating element. We validated the model with experimental data and found that predicted and measured values were closely matched. The measured and predicted values of normalized power consumption and top layer temperature over 48 hours are shown in Figure 7 and Figure 8. Figure 9 shows the measured and predicted shower temperatures during four successive showers.

We found that the mean prediction error in the total energy consumption was less than 6%, while the mean absolute error in predicted shower temperature was less than 3°C. It was considered acceptable for the model to be used in the tool.
2.4 Performance calculator

The performance calculator has two main functions: calculating peak reductions in the hot water load and estimating the customers comfort level.

First, it determines an average uncontrolled load profile for each household. The average uncontrolled load profile for a household represents an average profile of the household obtained over a specified number of Monte Carlo iterations.

Then, it determines an aggregate uncontrolled load curve $L_U$ by aggregating uncontrolled load profiles for all households. An aggregate controlled load curve $L_C$ is obtained in a similar manner after a switching program is applied to the uncontrolled loads produced by the hot water cylinder model for individual households. The peak load reduction $R_t$ of the control period $t$ is defined as

$$R_t = 1 - \frac{\text{max}[L_C(t)]}{\text{max}[L_U(t)]}$$

where $\text{max}[L_C(t)]$ and $\text{max}[L_U(t)]$ are the peaks of $L_C$ and $L_U$ the control period $t$, respectively.

The customer’s comfort level depends on the frequency (or probability) of getting a “cold shower”—the event when the shower temperature drops below the comfort temperature (e.g. 43°C) specified by the tool user. Preferred shower temperatures range from 40°C to 44°C [8]. Because of a large number of households in the controlled area, we can assume the same comfort temperature for all customers. The tool allows the user to change the comfort temperature if required.

2.5 Switching Program Optimization

Figure 10 shows a block diagram of the switching program optimizer. Here I/Os are depicted as numbered blocks. I/O 1 represents parameters of the control management system, I/O 2 optimization parameters, I/O 3 uncontrolled loads generated by the hot water cylinder model, and I/O 4 represents optimized switching programs.

The switching program generator uses user-specified control management system parameters and optimized turn-off periods from the optimizer to create switching programs, as shown in Figure 11. Here a control step is the smallest switching time interval, and a turn-off period is the time interval where the hot water system is turned off for a number of consecutive control steps. A switching cycle consists of the turn-off period followed by the turn-on period. A control period consists of multiple switching cycles (there are two control periods – for the morning peak and for the evening peak). Control groups are formed by shifting the switching cycles by one or more control steps. To ensure the time-shifted switching cycles are contained within a control period, each control group has one switching cycle less than the control period. In [9], it was demonstrated that division of households based on the family type does not significantly affect the comfort level of household residents. Therefore, the entire set of households can be divided into control groups of approximately the same size regardless of the family type of a household.

The load estimator determines the total controlled hot water load by applying a switching program to uncontrolled loads of individual households. The load estimator sets the load to zero during the turn-off periods of the applied switching program and restores the load during the turn-on periods. Water temperature is not considered in the load estimation.

The main function of the optimizer is to optimize turn-off periods of a switching program. It consists of the user-defined control period (UDCP) optimizer and the optimized control period (OCP) optimizer.
The UDCP optimizer determines turn-off periods based on the user-defined control periods and the peak load reduction targets. The control periods remain unchanged throughout the optimization process. The UDCP optimizer implements an iterative process to minimize the mean error between the user-defined peak load reduction target \( L_T \) and the estimated aggregate controlled load \( L_C \) in each switching cycle of a switching program. To calculate required changes in the turn-off period for each switching cycle, it applies proportional and integral (PI) functions to the errors. In Figure 12, \( e(j,k) \) and \( \tau_{off}(j,k) \) are the mean error and the turn-off period of switching cycle \( j \) in iteration \( k \), respectively; \( K_p \) is the proportional gain and \( T_i \) the integral time of the PI functions.

The proportional function multiplies the error by \( K_p \). The integral function sums the errors of switching cycle \( j \) from the previous \((S-1)\) iterations to the current one, and multiplies the result by \( K_p/T_i \). The sum of the current turn-off period and outputs from PI functions is converted by the limiter function into an integer between the minimum and maximum values. The final result is the turn-off period for the next iteration.

The OCP optimizer determines turn-off periods and control periods of a switching program based on the user-defined peak load reduction target \( L_T \). First, it finds the starting time \( t_i \) and finishing time \( t_f \) of the initial control period. The time \( t_i \) is found as the first intersection of the aggregate uncontrolled load \( L_U \) and the target \( L_T \), as shown in Figure 13. To avoid a high payback peak after the control period, the finishing time \( t_f \) is found by solving the following equation:

\[
\int_{t_i}^{t_f} L_U(t) \cdot dt = L_T \cdot (t_f - t_i) \tag{2}
\]

where the left hand term represents the total uncontrolled energy consumption between \( t_i \) and \( t_f \).

To further minimize the error between \( L_C \) and \( L_T \), the OCP optimizer iteratively tunes the switching program optimized by the UDCP optimizer. The OCP optimizer increases or decreases the turn-off period \( \tau_{off} \) of each switching cycle to minimize the error between \( L_T \) and \( L_C \). We define three tolerance levels: \( L_1 \) and \( L_2 \) are, respectively, 1% and 2% above \( L_T \), and \( L_3(j) \) is the difference between \( L_T \) and the estimated maximum restored load in switching cycle \( j \), if \( \tau_{off}(j) \) is decreased by one control step:

\[
L_3(j) = L_T - \max[L_U(j-2), L_U(j-1), L_U(j)] \cdot \frac{\tau_{step}}{\tau_{sc}} \tag{3}
\]

where \( \tau_{step} \) is the control step; \( \tau_{sc} \) is the switching cycle; \( \max[L_U(j-2), L_U(j-1), L_U(j)] \) is the maximum value of the aggregate uncontrolled load \( L_U \) over three switching cycles \((j-2), (j-1) \) and \( j \).

The OCP optimizer tunes the \( \tau_{off} \) of all but the last switching cycle within a control period based on the three scenarios shown below, where \( L_C(j) \) denotes values of \( L_C \) within switching cycle \( j \).

- Scenario 1. The peak of \( L_C(j) \) is above \( L_2 \).
- Scenario 2. \( L_C(j) \) stays between \( L_1 \) and \( L_2 \) for more than 15 min.
- Scenario 3. The peak of \( L_C(j) \) is below \( L_3(j) \).
Scenarios 1 and 2 represent overshooting, whereas Scenario 3 indicates over-control that can potentially create higher payback peaks. The OCP optimizer reduces $L_c(j)$ by increasing $\tau_{\text{off}}(j)$ by one $\tau_{\text{step}}$ if either Scenario 1 or Scenario 2 is met. If Scenario 3 is met, $\tau_{\text{off}}(j)$ is decreased by one $\tau_{\text{step}}$. No change is made on $\tau_{\text{off}}(j)$ if none of the above conditions are met.

Before changing $\tau_{\text{off}}(j)$, the OCP optimizer considers the current value of $\tau_{\text{off}}$ (expressed as the number of control steps) and the location of the peak of $L_c(j)$ within switching cycle $j$. For a peak located within control step $n$ of the switching cycle, increasing $\tau_{\text{off}}(j)$ of this switching cycle will reduce the peak only if the current value of $\tau_{\text{off}}$ is below or equal to $(n-1)$; decreasing $\tau_{\text{off}}(j)$ of this switching cycle will increase the peak only if the current value of $\tau_{\text{off}}$ is below or equal to $n$.

If $j$ is the last switching cycle of a control period, and either Scenario 1 or Scenario 2 is met, the control period is extended by one switching cycle; $\tau_{\text{off}}(j)$ is then set to a value equal to a multiple of $\tau_{\text{step}}$ and proportional to the error between the peak of $L_c(j)$ and $L_T$. Through iterations, the OCP optimizer tunes the switching program so that the aggregate controlled load stays below or as close as possible to the user-defined target.

### 2.6 Case Studies

We conducted several case studies to evaluate the performance of the DSM evaluation tool under various scenarios for 279 households. This set of households provided us the opportunity to use actual energy metering data in the developed tool. We used the tool to randomly generate hot water consumption profiles for 279 households and obtained an aggregate uncontrolled hot water load curve, which matched the actual data. In case studies 1 and 2, we investigated potential impacts of using constant values of ambient temperature, cold water temperature and thermostat settings on the simulation results. In subsequent studies, we evaluated the performance of switching programs produced by the optimizer in terms of the peak load reduction and customer comfort level. We used 43°C as the preferred shower temperature for all households. The default switching program configuration had 30 min switching cycles and 5 min control steps. The turn-off period in a switching cycle varied from 5 min to 25 min in the 5-minute step. The households were divided into six control groups of almost equal size.

#### 2.6.1 Case Study 1

This case study compares results of two simulations. In the first simulation, we use actual values of ambient temperature $T_a$ and cold water temperatures $T_c$, shown in Figure 14. Shaded areas indicate peak periods of hot water usage (06:00 – 09:00 and 16:30 – 18:30). The profile of $T_a$ is obtained from historical climate data for Tasmania [10]; $T_a$ usually has a positive correlation with $T_a$ [11], but has a smaller range of variation. As can be seen in Figure 14, values of $T_a$ and $T_c$ vary considerably over the 24-hour period (particularly, values of $T_a$), but their variations during peak periods are rather small. Therefore, in the second simulation, $T_a$ and $T_c$ are set to constant value of 8°C.

![Figure 14. Average ambient and cold water temperatures in winter time.](image)

Figure 14 shows two aggregate uncontrolled hot water load curves obtained using variable and constant values for ambient and cold water temperatures. We find insignificant difference between the two curves. The difference in the total energy consumption is about 1%, and the mean absolute error (MAE) is about 1.3p.u. The results can be explained by the fact that a great majority of hot water usages occur during peak periods when variations of actual cold water temperature are rather small (within ± 1°C, in shaded areas of Fig. 14). On the other hand, although $T_a$ varies significantly during the day, its variation has negligible overall effect on the rate of hot water tank heat losses. An insulated hot water tank idles for a long period (usually from 13 to 15 hours) between two consecutive recharges due to heat loss. During this period, the effect of $T_a$ variation is smoothed, and the average value of $T_a$ produces similar results as its variable values. Thus, variations of $T_a$ and $T_c$ can be represented with their respective average values in further studies.

![Figure 15. Uncontrolled load curves for constant and variable values of ambient and cold water temperatures.](image)
2.6.2 Case Study 2

This case study compares the performance of the UDCP optimizer and the OCP optimizer. Both use the default switching program configuration to produce optimized switching programs that are applied to the same set of hot water loads. The peak reduction target is 15% in both cases. Figure 16 and Figure 17 show the aggregate controlled load curves produced by the UDCP and OCP optimizers, respectively. Table 4 shows the control periods and peak reductions achieved. The UDCP optimizer keeps user-specified control periods constant in its optimization process. Probabilities of cold showers for each family type are shown in Table 5 – for the uncontrolled scenario, and scenarios controlled by the UDCP-optimized and OCP-optimized switching programs.

Comparing the aggregate controlled load curves produced by both optimizers, we find that the OCP optimizer performs much better in terms of peak load reduction.

Figure 16. Result of the UDCP optimization.

Figure 17. Result of the OCP optimization.

<table>
<thead>
<tr>
<th>Family type</th>
<th>UDCP optimizer</th>
<th>OCP optimizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family type 1</td>
<td>0.02%</td>
<td>0.03%</td>
</tr>
<tr>
<td>Family type 2</td>
<td>4.37%</td>
<td>4.63%</td>
</tr>
<tr>
<td>Family type 3</td>
<td>7.96%</td>
<td>8.44%</td>
</tr>
<tr>
<td>Family type 4</td>
<td>13.85%</td>
<td>14.36%</td>
</tr>
<tr>
<td>Overall</td>
<td>5.06%</td>
<td>5.34%</td>
</tr>
</tbody>
</table>

The starting and finishing times of control periods in a switching program are vital for peak load reduction. A delayed control period produces an initial peak above the target line, as in the evening period of Figure 16. Starting a control period too early defers loads needlessly and creates slightly higher peaks in subsequent switching cycles of the same control period, as in the morning control period of Figure 16. Control periods with sufficient length allow a gradual restoration of loads below the target line. Ending a control period prematurely creates an unwanted high payback peak at the end of the control period, as seen at around 11:30 of Figure 16. Similar results were reported in [12] and [13]. Due to shorter than required control periods used for the UDCP optimization, reducing the peaks at 10:30 and 21:30 will produce higher payback peaks at the end of the respective control periods.

While both controlled scenarios produce higher probabilities of cold shower than in the uncontrolled scenario, the OCP optimizer degrades the comfort level more than the UDCP optimizer due to its longer control periods (Table 5).

2.6.3 Case Study 3

In this case study, we evaluate the tool’s ability to optimize switching programs for two different hot water load profiles. The first one has a dominant morning peak (this load profile was used in the case study 3) and the second – a dominant evening peak. The default switching program configuration (30 min switching cycle with 5-minute control steps and six control groups) is used. The peak reduction target is 15%. Figure 18 shows the aggregate uncontrolled load curve of the second hot water load profile, and the aggregate controlled load curve after the OCP-optimized switching program is applied.

Optimized morning and evening control periods are from 07:30 to 15:00 and from 17:30 to 23:30, respectively. A 9.1% peak reduction is achieved for the morning control period, and 13.4% for the evening. The default switching program configuration (30 min switching cycle with 5-minute control steps and six control groups) is used. The peak reduction target is 15%. Figure 18 shows the aggregate uncontrolled load curve of the second hot water load profile, and the aggregate controlled load curve after the OCP-optimized switching program is applied.

Optimized morning and evening control periods are from 07:30 to 15:00 and from 17:30 to 23:30, respectively. A 9.1% peak reduction is achieved for the morning control period, and 13.4% for the evening. Table 6 shows probabilities of cold showers estimated for each family type under uncontrolled and controlled scenarios. As can be seen from Figure 18, the tool cannot further reduce the payback peak detected at 14:30 as the morning control period has reached the maximum limit of 7.5 hours.
Figure 18. The OCP optimization of a hot water load profile with the dominant evening peak.

Table 6: Probabilities of cold showers for case study 3

<table>
<thead>
<tr>
<th>Family type</th>
<th>Uncontrolled</th>
<th>Controlled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family type 1</td>
<td>0.03%</td>
<td>0.05%</td>
</tr>
<tr>
<td>Family type 2</td>
<td>4.11%</td>
<td>4.48%</td>
</tr>
<tr>
<td>Family type 3</td>
<td>7.50%</td>
<td>8.31%</td>
</tr>
<tr>
<td>Family type 4</td>
<td>14.32%</td>
<td>15.81%</td>
</tr>
<tr>
<td>Overall</td>
<td>4.82%</td>
<td>5.30%</td>
</tr>
</tbody>
</table>

Comparison of the results produced by the OCP optimizer in the case studies 2 and 3 (Tables 5 and 6) reveals that customers experience similar comfort under different load profiles.

2.6.4 Case Study 4

In this case study, we use the hot water load profile of case study 2 and compare two different switching programs represented in Table 7. Results produced by the OCP optimizer for case study 3 represent the implementation of the default configuration. Results shown in Figure 19 represent the implementation of the second switching program (configuration 2), and Table 8 shows probabilities of cold showers estimated for each family type. The optimized control period is from 07:30 to 13:30 in the morning and from 17:30 to 00:00 in the evening. Peak reductions for morning and evening control periods are 14.8% and 13.2%, respectively.

Table 7: Switching program configurations in case study 4

<table>
<thead>
<tr>
<th></th>
<th>Configuration 1 (default)</th>
<th>Configuration 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control groups</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Switching Cycle</td>
<td>30 (min)</td>
<td>30 (min)</td>
</tr>
<tr>
<td>Control Step</td>
<td>5 (min)</td>
<td>10 (min)</td>
</tr>
<tr>
<td>Turn-off periods</td>
<td>5, 10, 15, 20, 25 (min)</td>
<td>10, 20 (min)</td>
</tr>
</tbody>
</table>

Table 8: Probabilities of cold showers for case study 4

<table>
<thead>
<tr>
<th>Family type</th>
<th>Uncontrolled</th>
<th>Controlled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family type 1</td>
<td>0.02%</td>
<td>0.08%</td>
</tr>
<tr>
<td>Family type 2</td>
<td>4.37%</td>
<td>4.80%</td>
</tr>
<tr>
<td>Family type 3</td>
<td>7.96%</td>
<td>8.67%</td>
</tr>
<tr>
<td>Family type 4</td>
<td>13.85%</td>
<td>14.55%</td>
</tr>
<tr>
<td>Overall</td>
<td>5.06%</td>
<td>5.51%</td>
</tr>
</tbody>
</table>

The default switching program configuration performs slightly better in peak reduction as it has smaller control steps and higher number of control groups. Switching program configuration 2 degrades the customer comfort level further as hot water systems are switched off for longer periods of time.

Figure 19. The OCP optimization with the switching program configuration 2.

3 Fast Demand Response for Enabling Higher Penetration of Renewable Energy

Customers living in remote areas often cannot be supplied from conventional interconnected power systems. These customers are usually serviced by a local electricity generation and distribution system with electricity generated using diesel fuel. Due to remoteness and consequent high cost of diesel fuel supply, the cost of electric energy in isolated power systems is high compared to conventional interconnected systems. In some locations, the price exceeds US $1/kWh, which is an obvious incentive for introducing renewable energy (RE) generation. Unfortunately, RE from the two most abundant energy sources – wind and solar – incurs significant stability and reliability issues due to the intermittency of those sources.

This section presents fast (i.e. sub-second) demand response (DR) as an RE enabling technology in isolated power systems. The advantages of DR are presented, the concept of a fast DR system is discussed, and a case study of the first implementation of the fast DR system, along with some preliminary results, is presented.

3.1 Demand Response in Isolated Power Systems

In general, DR can provide benefits to power systems and their customers by

- Supporting frequency and/or voltage regulation [14], [15];
• Reducing operational costs and emissions by increasing utilization of RE sources [16]. Note that reducing operating costs in turn leads to a greater return on investment which would incentivize expansion of the RE industry;
• Reducing operational costs and emissions caused by traditional generators installed to provide spinning reserve for RE [17];
• Relieving stress from transmission and distribution infrastructure by coordinating loads close to RE sources [17];
• Reducing utility operating costs through advanced metering infrastructure installed to enable DR [18].

Most DR applications focus on large power systems, which have a steady and predictable load profile defined by morning and evening peaks. The time and size of these peak periods can be accurately estimated using historical load data and weather forecasts. In contrast, isolated power systems not only supply less power (MWs, rather than GWs) they are also geographically much smaller. Being smaller in capacity means that demand is less predictable. Being smaller in area means that supply from RE sources is more variable, as a larger percentage of the RE generators are likely to be affected by the same weather events (e.g. a lull in the wind or passing clouds). Due to their reduced demand predictability and increased variability of RE supply, conventional generation scheduling in isolated systems with RE is more challenging. From a generation scheduling perspective - where RE generation is usually treated as a load offset – the daily load curve becomes extremely volatile. An example of an isolated power system (IPS) with RE load curve is given in Figure 20. Notice that it does not even have predetermined daily peaks.

High load variation makes scheduling of diesel generation more difficult and less efficient, as diesel engines will rarely operate at their peak efficiency, and more generator start-ups are required. This is where fast DR can help. It can smooth the variability of required diesel generation by quickly adjusting system load. To be able to do that, fast DR cannot rely on typical load patterns – it must be executed in real-time. If DR is to complement RE in an isolated system, it has to be as fast as the speed at which RE generators change their power output.

3.2 Demand Response as a Virtual Power Plant

The idea of aggregating and controlling small loads to create a large block of variable demand has been discussed in the power engineering literature where it is often referred to as a ‘virtual power plant’ [19]. Since isolated power systems generally have only one power station, the idea of adding another (virtual) power plant to this system might be a bit misleading. However, isolated systems with RE might have several generating sources (e.g. diesel, wind, solar, etc). Therefore, aggregated DR could be treated as a ‘virtual generator’. Although such a ‘DR Generator’ (DRG) is not generating real power, the power system controller perceives it as one due to the DRG’s ability to decrease the amount of energy needed from other generating sources.

![Figure 20. Daily load diagram in an isolated power system.](image)
spinning reserve is provided by the DRG, diesel generator start-up is prevented.

![Diagram of Smart Grid generator in IPS control system.](image)

Figure 21. Smart Grid generator in IPS control system.

A DRG consists of three main components, as shown in Figure 22:

- **DRG Master controller**, communications network, and slave controllers.

![Diagram of DRG architecture.](image)

Figure 22. The DRG architecture.

### 3.2.1 DRG Master Controller

The master controller collects information on available DR and aggregates it into predefined virtual loads (e.g. geographical regions). It constantly communicates available DR capacity to the power system controller (PSC) while checking for DR dispatch requests from the PSC. DR requests identify a target virtual load and the amount of demand to curtail. The master controller selects the individual loads to curtail from the virtual load, and immediately sends a switch-off signal to each.

### 3.2.2 Communications Network

A multi-protocol bidirectional communications network delivers information between all elements of the DRG. Ethernet is used within the control system. A dedicated WiMAX network provides the backhaul capability. Within individual customer sites, a WiMAX gateway is connected by Ethernet to a ZigBee gateway for the final link to the load metering and switching devices. This communications configuration is configured to ensure a sub-second round trip for DR requests from the PSC out to the load control devices and back again.

### 3.2.3 Slave Controllers

Slave controllers are located in each DR capacity providing site. They consist of a pair of gateways to provide WiMAX-Ethernet-ZigBee signal translation between the backhaul network and the individual load control devices. The load control devices perform both metering and load switching. They provide a range of power metrics and also support set-points.

A DRG providing spinning reserve must be extremely responsive and reliable. The master controller must be able to monitor and dispatch slave controllers at all times. This critical requirement becomes obvious in the two most common scenarios:

- If the PSC requests DR for extended periods of time, some slave controllers may override the dispatch as they exceed their maximum dispatch duration. In this situation the master controller must quickly identify and dispatch another device (or devices) with an equivalent load.
- If a slave controller or communication link is unreliable, the DRG may be forced to always dispatch more DR than requested to ensure a suitable margin of error in either load switching or reporting. This is not an efficient use of capacity and may reduce the overall effectiveness of the DRG.

### 3.3 Case Study

The fast DR technology discussed above was implemented in an IPS as part of the King Island Renewable Energy Integration Project [20]. King Island lies in the Bass Strait between Tasmania and the Australian mainland. It has a population of approximately 2000 people, and an economy based on agriculture and food processing.

#### 3.3.1 The King Island Power System

Customer load on King Island ranges between 1 MW and 3 MW, with an average of around 1.5 MW. The King Island power system is shown in Figure 23. There is one power station on the island with four distribution feeders delivering electricity to customers. The power station houses four diesel generators with a
total generation capacity of 5.8MW. Three fixed speed Nordex N29 (250kW each) wind turbines are installed on a nearby hill, together with two Vestas V52 turbines (850kW each) with doubly fed induction generators. Two 800kW diesel engines with flywheels are also connected to the system. In these generators the flywheels are separated by a clutch from a diesel engine and provide system with additional inertia.

Figure 23. The King Island power system.

3.3.2 The King Island Smart Grid Project

The ongoing King Island Smart Grid project has the goal of supporting higher levels of wind energy integration in the King Island power system by providing:
1. Spinning reserve by implementing the DRG concept, and
2. Fast fine-grained under frequency load shedding (based on the slave controller level).

The master controller constantly monitors the available connected DRG load and passes this information to the PLC-based power system controller. At the same time, the controller monitors the current level of power system spinning reserve. If the spinning reserve falls below a predefined threshold, the controller instructs the DRG to curtail some load and effectively raise the spinning reserve. This function can be observed in Figure 24, where due to a sudden drop in wind generation, the spinning reserve falls. The reduction in spinning reserve causes the power system controller to initiate a request for all available DR, as shown in the lower graph.

The results shown in Figure 24 demonstrate that the implemented DRG was able to respond accurately to given set-points. It also shows that DR capacity can be dispatched reliably in 1 sec.

3.3.3 Preliminary Implementation Results

Currently the King Island DRG has 50 sites under management and has been fully integrated into the power system. When complete, the DRG will be extended to include 150 households and several commercial loads. Prior to roll-out on King Island, the DRG was tested with 10,000 simulated customer loads with minimal performance degradation [21]. This implies that the full DRG will supply more than 100 kW of sub-second DR capacity.

Figure 24. Results of the King Island DRG operation.

The effectiveness of the King Island DRG depends largely on its integration with the power system controller. Figure 25 demonstrate how the controller requests for the DRG and uses it as a tool for regulating demand accurately. During the period of over 2 hours, the King Island power system was running in zero-diesel operation and the DRG was used under small and short dips in wind power generation. In this mode of operation, the power system controller prioritizes DR dispatch, and thus postpones unnecessary start-ups of a diesel generator.

Figure 25. Operation of the King Island DRG.
4 Conclusions

With a strong drive for energy conservation, demand-side management and demand response are becoming vital for the implementation of the smart grid concept. This paper outlines some experience obtained at the University of Tasmania, Australia in developing DSM and DR systems.

The evaluation tool to recommend optimum DSM switching programs for domestic hot water systems has been developed. The tool assesses the performance of a DSM switching program by estimating potential peak load reductions and customer comfort characterized by the probability of cold showers. The starting time and the length of control periods are crucial in peak reduction. However, the length of control periods must be limited to minimize negative impact on customer comfort. The developed tool aims to assist distribution system operators in designing their DSM programs. An operator uses this tool to determine the available domestic water heating load in a controlled area, and predict the potential reduction in peak load. The tool described in this paper has been implemented in the Tasmanian power system since June 2013.

The paper also presented DR as an enabling technology for higher penetration of renewable energy in isolated power systems. These systems are often based on diesel generators. However, due to high costs of diesel fuel supply, the cost of electricity in isolated power systems is much high compared to conventional interconnected systems. This presents an incentive for introducing renewable energy generation in isolated power systems. Unfortunately, the integration of renewable generation presents significant stability and reliability challenges due to their intermittency. The solution proposed in this paper is based on centralized two-way communication and control of residential and commercial loads. DR can be dispatched and confirmed within 1 second. The technology has been installed and successfully tested in an isolated power system on King Island in Australia.

References


Future Intelligent Civil Structures: Challenges and Opportunities

J. C. Li, Y. Li, M. Askari and Q. P. Ha
Centre for Built Infrastructure Research, University of Technology Sydney, Australia
E-mail: Jianchun.li@uts.edu.au, Yancheng.li@uts.edu.au, Mohsen.askari@uts.edu.au

Abstract -
An intelligent civil structure offers ultimate protection to its structure, contents and occupants in terms of safety and functionality against undesired dynamic loadings and structural deficiency. In this paper, the concept of the future intelligent civil structure featuring self-adaptive, self-prognostic, self-sensing, self-powering and self-repairing abilities, is proposed. A decade research efforts from Centre for Built Infrastructure Research, University of Technology Sydney, towards the development and concept proof of such intelligent structure is reviewed.

Keywords -
Intelligent civil structures; Structural control, Structural health monitoring, Self-adaptive; Self-prognostic; Self-sensing; Self-powering; Self-repairing;

1 Introduction

Emerging from recent advance in smart material technology, advanced sensing, signal processing and advanced control theory, intelligent civil structures potentially offer ultimate protection to the civil structures as well as their contents and occupants in terms of safety and functionality against undesired dynamic loadings and damage or structural deficiency. Intelligent structure related research and development, therefore, possess enormous theoretical and practical values and thus are becoming a rapidly emerging area of research recently [1-2].

An intelligent civil structure can be defined as a structural system with a higher-level of autonomy. Such intelligent structures rely on the embedded functions of sensors, actuators and micro-processors with optimizations and integrated structural control algorithms that can automatically adjust structural characteristics to optimally adapt to external disturbance and environmental changes. Thus, structural safety and serviceability as well as the prolongation of structural service life are assured [2].

There have been many research and development in the field to develop smart civil structures. Benefiting from recent advances in smart materials, i.e. piezoelectric material[3-4], shape memory alloy, electro-sensitive materials and magneto-sensitive materials[5-8], various smart materials have been utilised in development of smart devices to facilitate semi-active/active structural control. Adaptive structural control of civil structures forms critical component for the development of intelligent civil structures.

In this paper, after a brief review on the recent development on intelligent civil structure research, the authors will present a new concept of intelligent civil structure with “5S” features. An introduction on the testing facilities available in Centre for Built Infrastructure Research (CBIR) will be given and the research activities in CBIR toward advancing intelligent civil structure research will then be reported. At the end of the paper, research activities to be carried out in next stage are also highlighted.

2 Future Intelligent Civil Structure Research

An intelligent civil structure of the future should possess the following characteristics: optimally respond to a changes of both external and internal environment (self-adaptive); instantly and reliably detect structural change or damage and evaluate their impacts (self-prognostic); automatically sense and collect information of structural members (self-sensing); be independent from external power sources for the required operations (self-powering); utilizing various means in real-time to minimise impact of damage and restore structural load carrying capacity (self-repairing). These 5S features of an intelligent structure represent authors’ definition of true intelligent civil structures and impose a series of challenges to the field of research. The challenges may be described as followings:

• Design and development of smart structural members that can be integrated as a part of the structure;
• Embedment of energy harvesting into smart structure design to provide independent power supply for the smart structure system;
• Utilisation and integration of sensing ability of smart materials to accomplish self-sensing ability;
• Optimization of numbers and locations of smart structural members in the structure;
• Online and real-time identification of structural characteristics after either graduate or sudden changes.
• Developing novel algorithm to integrate and optimise structural control and structural monitoring;
• When structural damage occurs, utilizing controllability of smart structural members to alter...
loading path in to minimize damage effects and further damage;

- Experimentally realize and evaluate 5S intelligent civil structure in laboratories.

3 Research Facilities at Centre for Built Infrastructure Research

3.1 UTS Shake Table

The UTS shake table has been established from two consecutive infrastructural funding from Australian Research Council after 1989 Newcastle earthquake, which is aimed at encouraging earthquake engineering research in Australia. Designed by a well-known test equipment manufacturer, MTS from USA, it was established in the Structure Lab of UTS in 1995. This advanced testing facility is the only kind within Australian universities. The self-weight of the shake table is 6 tones and it is capable of carrying a payload up to 10 tonnes. The maximum horizontal displacement is ±100mm and the maximum acceleration for the bare table is 2.5g. With advanced hydraulic system and control system, this shake table can simulate any vibration sources in the range of 0.1Hz to 100Hz. For earthquake engineering research, the four benchmark earthquakes, i.e. 1940 El-Centro, the 1968 Hachinohe, the 1994 Northridge, and the 1995 Kobe earthquake records can be produce by the shake table.

Figure 1. UTS shake table

3.2 International 5-Storey Benchmark Building Model

In 1990s under auspices of the International Association of Structural Control and Monitoring, a 5-storey international benchmark building model was designed and manufactured in UTS. The facility has been adopted by the community for encouraging international collaborative research in the area of vibration control of building structures. The model has a height of 3m and five adjustable story units with typical story height of 0.6m. the mode has a footprint of 1.5m x 1.0m. It consists of two bays in one direction and a single bay in other direction. The mass of the building model can vary between 1 tonne to 3.5 tonnes. The joint connections of the structure can be changed between fixed and pinned. The building height can also be modified by the innovative connection design.

4 Intelligent Civil Structure Research at Centre for Built Infrastructure Research

In developing future 5S intelligent civil structures the critical issues to be explored are: 1) design and characterization of smart structural members; 2) development of control algorithm for self-adaptive ability for intelligent structures; 3) development of energy harvesting unit for self-powering and self-sensing; 4) integration of structural control and structural health monitoring for self-prognostics and self-repairing. To date, Centre for Built Infrastructure Research has made progress in advancing each of the critical technology individually. Future research in integration of the system to create an ultimate intelligent civil structure with 5S features will be carried out in the future work.

4.1 Development of Semi-Active Structural Members

4.1.1 Smart Pin-Joint using Magnetorheological Fluids

In the types of beam-column structures, type of joint connections, i.e. fixed and pinned connections, affect stiffness of the structure. A fully controllable beam-column connection joint allows civil structures to alter its stiffness matrix adaptively in a wide range thus achieve structural control [9-10]. Such concept was firstly proposed by UTS researchers [9]. The research aimed to design and develop a variable smart pin-joint by utilizing a magnetic sensitive material, i.e. magnetorheological fluids (MR fluids). MR fluids possess reversible, controllable and real-time viscous property when subjected to external magnetic field. With an electromagnetic coil to provide varying magnetic field, a smart pin-joint was capable of producing controllable moment resistance in a range of less than 1Nm to 30 Nm. Steady rotational testing [11-12] and dynamic cycling testing [13-14], shown in figure 4, were conducted using experimental setups shown in Figure 3. Theoretical evaluation [15] and hysteresis modelling [13-14] were carried out to characterise the performance of the MR pin-joint. In the next stage work, the smart pin-joint will be installed in a two-story building model for applications on structural control and online system identification.

Figure 2. International 5-Storey Benchmark Building Model

Figure 3. MR pin joint: (1) cross-section; (2) prototype
4.1.2 Adaptive Magnetorheological Elastomer Base Isolator

As most important means in seismic protection of civil structures, base isolation is used to decouple superstructure from damaging ground motion, i.e. earthquakes, by installing isolation layers between them. Recently, base isolation systems encountered their challenges for its incapability in adapting versatile earthquakes due to its passive nature. This research proposed a novel adaptive base isolator utilising magnetorheological elastomer (MR elastomer). Similar as MR fluids, MR elastomer can change its material property, i.e. shear modulus, when external magnetic field is applied. The proposed adaptive base isolator, inheriting critical laminated design, possesses large compressive stiffness with controllable lateral stiffness [16-21], shown in figure 5. Two prototypes of MRE adaptive isolators, i.e. with 37 MRE layers (2mm thick each) and 25 MRE layers (1mm thick each), have been fabricated and tested (Figure 6). the experimental results revealed that they can induce stiffness increase of 37% [17] and 1630% [21], respectively, as shown in Figures 7-8. To characterise the behaviours of the devices, different hysteresis models, i.e. Bouc-Wen model [22] and Strain-stiffening model [23], have been proposed. Large scale experimental testing will be conducted in late 2014 and early 2015 to prototype smart base isolation system utilizing UTS shake table and international 5-story benchmark building model. This research already won several awards in Australia, including 1st prize in UTS Uniquest Trailblazer and 3rd prize in National Grand Final of Research and Innovation by Uniquest. Intellectual property on magnetic field generation for multiple MRE layers is protected by Australian PCT patent [24].

4.2 Energy Harvesting Technology for Self-Sensing and Self-Powering Ability

Scavenging energy from vibrations, known as energy harvesting, is an emerging technology of alternative energy solution in engineering fields. It can be used to power the sensors in the structural health monitoring systems as well as provide power for semiactive control devices. Due to the intrinsic relationship between the harvested energy and external vibration, it can be also used as means to collect dynamic information on external environment. Thus, development of energy harvesting units is the key towards realization of self-powering and self-sensing abilities.

4.2.1 Energy Harvesting Device using Piezoelectric Material

This project focuses on the development and testing of energy harvesting unit using piezoelectric material. A compression-based piezoelectric energy harvester using a multilayer stack configuration, shown in Figure 9, is designed and tested in Centre for Built Infrastructure
Research. The stack configuration makes this design suitable for environments with large compressive force and small deformation. In this research, theoretical analysis is explored to evaluate its capacity [25]. Dynamic testing (figure 10) was conducted to find the relationship between the harvested energy and external loadings [25-26]. The experimental results show that the proposed piezoelectric stack harvester can generate up to 200 mW electrical power under harmonic excitation. Linear relationship was found between the harvested power and the external vibrations [26].

![PZT stack harvester](image1)

Figure 9. PZT stack harvester

![Experimental testing](image2)

Figure 10. Experimental testing

4.2.2 Energy Harvesting Device using Electromagnetic Induction

This project addresses the development of energy harvesting unit utilizing electromagnetic induction (EMI). The principle of this design is to allow moving permanent magnets to pass through the magnetic field induced by electromagnetic coils. Detailed design is shown in Figure 11. In [27] and [28], detailed design is introduced on the two-phase EMI energy harvester. Theoretical analysis was conducted to find the relationship between harvested energy and external vibration. Further experimental evaluation is to be conducted to verify the theoretical findings.

![Cross-section view of EMI energy harvester](image3)

Figure 11. Cross-section view of EMI energy harvester

4.3 Multi-Objective Optimal Placement of Structural Control Devices

The optimal design and placement of control devices, is an important problem that affects the control of civil engineering structures. According to literature, most of the reported works in this area have focused on optimal assignments of active control devices with one or two structural indices as the objective functions to be minimized through the optimization process [29]. However, it is necessary to do a comparative study between optimal locations of actuators and MR dampers and show the differences. Moreover, since optimal placement of control devices in a structure deals with integer adjustable parameters (each actuator/damper can be assigned to an integer value which is the floor number), an advanced and accurate integer coded optimization algorithm is needed. In addition, a set of Pareto fronts, as the final result of optimization, gives more flexibility to the user to choose one’s own design based on one’s own criteria.

To this end, a modified integer coded version of non-dominated sorting genetic algorithm II (MI-NSGAII) was introduced in this research study and applied to find optimal places of actuators and MR dampers in a nonlinear 20-storey benchmark building [30], shown in Figures 13-14. The method uses the best features of a recently developed integer coded algorithm named MI-LXPM into the framework of NSGAII. Using the proposed approach, a Pareto front can be generated using three considered objective indices, i.e. peak inter-storey drift ratio (J1), peak acceleration (J2) and peak base shear force (J3). The results of optimal placement of active actuators were compared to the benchmark problem definition in which 25 actuators have been located in non-optimal locations. Results showed the effect of proposed strategy where the same level of structural performance, in terms of proposed objective functions, was obtained by use of only 7 actuators in an optimal layout [31], shown in Figure 15.

![Optimal places of different number of actuators](image4)

Figure 13. Optimal places of different number of actuators respect to J₁, J₂ and J₃
4.4 Semi-Active Control of MR Dampers for Self-Adaptive

One of the challenges in the application of MR dampers is using an appropriate control algorithm to determine their command voltage[32]. Two essential steps for the development of semi-active control algorithms are: 1) development of appropriate hysteresis model for MR damper; and 2) evaluation of semi-active control algorithm for structural control.

This part of research aimed at addressing the aforementioned issues through numerical simulations and experimental testing. In the modeling of hysteresis of MR dampers, various models were proposed, i.e. Bouc-Wen model [33], Hyperbolic model [34], inverse model[30] and curve fitting technique [35]. Bouc-Wen model was modified based on classical Bouc-Wen model in which an evolutionary variable $z$ is used to model the nonlinear hysteresis [32]. In the hyperbolic model, a hyperbolic function was used to describe the nonlinearity [33]. The inverse model is to model the inverse dynamics of MR damper using a novel multi-objective optimal subtractive Fuzzy C-Mean clustering technique[30]. In the development of control strategies, several control algorithms have been used, such as Sliding Mode Control[34], Lyapunov-Based Control[35], optimal LQG control[30] and direct control [36-38]. Figure 16 shows the applications of proposed controllers to a 20-storey nonlinear building with 25 optimally placed MR dampers, illustrated that the proposed new control algorithms can effectively track the desired control force and perform much better than the original and modified versions of clipped optimal controllers (COC and MCOC) in terms of structural response reduction with less control force and power.

4.5 Online Real-Time Structural Identification for Self-Prognostics

Structural Identification and Damage prognosis is the prediction in near real time of the remaining useful life of an engineered system, given the measurement and assessment of its current structural state and accompanying predicted performance in anticipated future loading environments. An important objective of health monitoring systems for civil infrastructures is to identify the state of the structure and to evaluate its possible damage[39]. Frequently, damage can be inferred from the changes of structural parameters, such as the stiffness and damping coefficients.

This project focuses on online real-time structural identification [40] to acquire the status of the structures. Several real-time identification methods, i.e. Kalman filter, Extended Kalman Filter (EKF), iterated Extended Kalman Filter (IEKF) have been studies and results found that these methods are not effective in the case of highly nonlinear problems. To overcome the problem, two filtering techniques, namely, unscented Kalman filter (UKF) and iterated unscented Kalman filter (IUKF), have been recently developed to handle any functional nonlinearity. Comparative study was also carried out on the aforementioned methods for their effectiveness and efficiencies through a highly nonlinear SDOF structure as well as a two-storey linear structure.

The Recursive Least Square (RLS) based methods are also another category of identification approaches which have been used widely in estimation of structural system parameters. In this research [41-42], a new adaptive tracking technique, based on RLS with adaptive multiple forgetting factors (AMFF-RLS), was presented which can estimate the unknown structural parameters as well as unknown input, e.g. earthquake signal. The method was applied to different structures, with different excitations and damage scenarios. It is found from the results that the proposed algorithm can effectively identify the time-varying parameters such as damping, stiffness, as well as unknown excitations with high computational efficiency, even when the observed data were contaminated with different types and significant levels of noise[43].
5 Conclusions

Intelligent civil structure represents an emerging research trend in the field for protection of civil structures against undesired motions and disturbances. Centre for Built Infrastructure Research at UTS has embraced this new concept of the intelligent civil structures by conducting a series research on concept proof of Self-adaptive, Self-prognostic, Self-sensing, Self-powering and Self-repairing abilities of the intelligent civil structures. This paper provides a brief report on these activities and challenges in developing such intelligent structures. Further research towards integration and experimental evaluation of the intelligent civil structure has been planned and will be carried out soon.

References


[19] Li J., Li Y., Li W. and Samali B., Development of


[40] Askari M., Li J. and Samali B, Adaptive multiple forgetting factor recursive least square (AMFF-RLS) for real-time structural identification with unknown


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

AUTOMATION AND CONTROL
Control of an Adaptive Light Shelf Using Multi-Objective Optimization

Benny Raphael

Civil Engineering Department, IIT Madras, Chennai, India
E-mail: benny@iitm.ac.in

Abstract -

Users are increasingly expecting intelligent behavior from automated building systems. However, incorporating and integrating user preferences in an automatic control algorithm is a complex task. In most commercial systems, control is done through local search techniques such as Proportional Integral Derivative (PID) control. Multiple criteria cannot be accommodated in such techniques. The objective of this paper is to evaluate the use of multi-criteria optimization in the control of a lighting and shading device, an adaptive light shelf. This study is done in two stages. In the first stage, the performance of the light shelf is evaluated using simulations. EnergyPus and Radiance are used to simulate thermal load and daylighting. These are used to determine the optimal control actions using a multi-objective optimization algorithm. The energy savings for electrical lights is compared with a traditional control strategy. In the second stage, a prototype is constructed to evaluate the actual performance. Results from simulation show that significant energy savings can be achieved through multi-objective optimization control strategy.

Keywords -

Building automation, Control, lightshelf, multi-objective optimization, energy efficiency

1 Introduction

Traditionally control is treated as a single objective optimization problem and the best solution according to a selected criterion is chosen as the control action. However, building systems are getting increasingly complex and there is a need to accommodate multiple criteria for the best performance. For example, an external window blind might be partially opened or closed in order to satisfy requirements such as eliminate glare, provide adequate daylight, reduce cooling load and accommodate user preferences.

In multi-criteria optimization, several objectives are considered simultaneously and usually these conflict with each other. For example, minimizing the lighting energy might require opening the window blinds which would increase the radiant heat entering into the building. Therefore, in multi-objective optimization the emphasis is on identifying solutions that achieve the best trade-offs among multiple objectives. An early application of this concept can be found in Radford and Gero [1] where they applied dynamic programming in the multi-criteria design optimization with four performance criteria namely, thermal load, daylight availability, construction cost and usable area.

A popular approach to multi-objective optimization is the generation of a Pareto Front. In this approach, a population of solutions known as the Pareto set is generated in which all the solutions are Pareto optimal or non-dominated. By definition, there is no solution better than a Pareto optimal solution with respect to all the criteria simultaneously. Several applications of this concept are found in the literature (eg. [2,3]). Some of these techniques have already been used in the optimization of building systems. Caldas [4] studied the trade-off between the initial cost of a building and the energy performance of the building using a multi-objective optimization approach.

While the Pareto approach is useful for design tasks in helping engineers make decisions, it does not support the selection of a single best solution. Automatic selection of solutions is needed in tasks such as control where decisions have to be taken several times a minute or second. The issue of selecting a single solution from the Pareto set has been largely ignored by previous researchers. This paper presents a new algorithm called RR-PARETO3 for selecting the best solution that makes reasonable trade-offs among conflicting objectives.

2 RR-PARETO3 Algorithm

RR-PARETO3 algorithm is the third generation of the multi-criteria decision making algorithm presented in [5]. In this algorithm, the solution with the best trade-offs among all the objectives is chosen using two pieces of information:
• Order of the objectives according to their importance
• The sensitivity of each objective

The sensitivity of an objective specifies what level of increase in its value is acceptable to the user. For example, consider the objective of minimizing the power consumption. If increasing the power consumption above 5% is not acceptable to the user, the sensitivity of this objective is defined as 5%. In this case, the algorithm attempts to select solutions that are within 5% of the best solution. In the following discussion, the term solution point or point is used to denote a set of values of all the objectives as well as the decision variables (optimization variables).

The RR-PARETO3 algorithm works by iteratively eliminating the worst points according to maximum number of criteria. This filtering is done in two stages. In the first stage, the solution point with the best value for the current objective is chosen from among all the points. All the points that lie outside the sensitivity band of the chosen point are eliminated from the set. If the sensitivity is not specified for any objective, no filtering is done for this objective and all the solutions are retained. At the end of Stage 1, one or more points might remain in the solution set. If a unique solution is not identified, Stage 2 filtering is performed.

In Stage 2 filtering, the hypercube containing all the remaining solutions is trimmed. This is done by dividing the hypercube volume into half by bisecting each objective axis one by one according to their order of importance. Let $y_{min}$ and $y_{max}$ be the minimum and maximum values of the i-th objective among all the solutions in the current set. The threshold is computed as $(y_{min} + y_{max})/2$. In the minimization problem, the region containing values greater than this threshold is considered as the bad half with respect to this objective. Worst solutions are eliminated using the algorithm described below:

• Stage 2.1: Points that lie within the bad half of most objectives are eliminated, taking combinations of k most important objectives at a time (repeated for $k = N$ to 2)
• Stage 2.2: Points that lie in the bad half of individual objectives are eliminated according to the order of importance of objectives
• Stage 2.3: Iteratively remove the worst point according to each objective based on the order of importance

The process of bisection of hypercube helps to remove visibly obvious bad solutions. When there are still many points left, each objective is given a chance to remove the worst candidate in Stage 2.2. The most important objective is given the first chance. Stages 2.1 and 2.2 use values of objective function for elimination, while Stage 2.3 uses ranking of solutions. The process stops when a single solution remains in the set or all the remaining solutions have the same values for all the objective functions. By repeating Stage 2.2 for each objective, each criterion is given an opportunity to eliminate bad solutions and the final selection is a trade-off among all the objectives. It is emphasized that the process does not favour the best solution according to any objective. For example, if the best solution according to the first objective lies within the bad half of the second objective, this solution is eliminated. Since the process is driven by the order of importance of objectives, the users’ preferences in the selection process are also respected.

The algorithm can be used to filter a set of solution points that are generated by any optimization algorithm or even pure random sampling procedures.

3 Application to lightshelf control

A lightshelf is a horizontal or inclined projection with a high reflectivity meant to increase the depth of daylight penetration into a room. It operates by reflecting sunlight off to the ceiling from where it is further reflected to the work plane. Several studies have been conducted on the effectiveness of light shelves and other daylighting features. For example, Aghemo et al. [6] present a case study for the comparison of lighting performances of different traditional shading devices. If the light shelves are static and horizontal, they are able to increase daylight penetration only when the sun is at a low angle. A proposal to making the light shelf rotatable is presented in [7]. In this case, the angle can be adjusted to the position of the sun to produce deeper daylight penetration (Figure 1).
The control of the light shelf is a complex task. When the external light shelf is rotated in the clockwise direction, initially there is an increase in daylight penetration and later at greater angles, daylight decreases depending on the angle of the sun. There is an optimal position of the light shelf at which adequate shading is provided near the window and maximum daylight is produced in the interior. In addition, the amount of heat reflected by the light shelf into the interior of the building as well as the amount of heat blocked by it have to be computed in order to evaluate its performance with respect to energy required for cooling the buildings. This makes it a multi-objective control task.

3.1 Choosing values of algorithm Parameters

The only algorithm parameters that need to be specified in RR-PARETO3 are the order of objectives and the sensitivity of objectives. The order of objectives depend on the priorities of the user. In this application lighting is given higher priority, because the primary objective of using a light shelf is to improve the lighting level. The second parameter, sensitivity, depends on how much decrease in performance is acceptable to the user. In the case of light shelf, the user has to decide on what percentage of increase in cooling load is acceptable. This is a subjective decision. If the user does not have any preference, the parameter can be omitted allowing RR-PARETO3 to apply the default filtering algorithm.

3.2 Case study

In order to evaluate the control strategy, a hypothetical office building of plan area 8.0 x 6.0 meter is considered here. The dimensions are shown in Figure 2. A rotatable light shelf of width 1.0 m. is present on the west facade at a height of 2.0 m. Dimmable lights of 56 Watts are installed along a grid of spacing 2 m in both directions. Other data related to the case study are given in Table 1.

![Figure 2. Case study of an office building](image)

The daylight available in the room is calculated using lighting simulation software Radiance [8]. The thermal load is computed using EnergyPlus Version 8.1 [9]. In order to correctly calculate the effects of reflections from the light shelf, the "Solar Distribution" parameter in EnergyPlus is set to "Full Interior and Exterior With Reflections". Lighting power for the dimmable lamps was calculated from the illuminance values obtained through Radiance by assuming a linear dimming curve such that the power varies between 10% and 100% for minimum and maximum brightness of the lamp. In order to prevent glare, it is assumed that window blinds are completely closed if the maximum illuminance exceeds 2000 lux at a distance of 2 meters from the window. Hourly simulations are done for a typical summer day for different angles of the light shelf in order to determine the optimal position with respect to lighting and cooling energy criteria. As an illustration, the lighting and cooling loads for various angles of the light shelf are shown for 3 pm (Figures 3,4). The angle 0 corresponds to a horizontal light shelf and the angle 90 degrees corresponds to vertically upward position. The power does not increase or decrease monotonically because of the effects of local reflections as well as discretization and other errors in the simulations. It is clear that small random variations in the load should not be allowed to influence the optimal position. This justifies the use of the sensitivity parameter in the RR-PARETO3 algorithm. Also it is noted that the non-monotonicity of the objective function and the small random variations create problems for conventional local search algorithms. For example, a small step might seem to increase the value of the objective function and the algorithm might terminate assuming that a minimum is reached, while actually it might be a local effect due to errors. If such small random variations are ignored, a general increasing trend in lighting power is observed above 30 degrees. At the same time, there is a general decreasing trend in the solar thermal load.
After computing the lighting power and solar thermal load for all the angles of the light shelf from 0° to 90°, the compromise solution was obtained using RRPARTEO3 filtering. Three scenarios were considered. In the first scenario, lighting criterion was given higher priority over solar thermal load. Both objectives were given sensitivity values of 5%. In the second scenario, the solar thermal load criterion was given higher priority while keeping the sensitivity values unchanged. In the third scenario, the lighting criterion was assigned higher priority, while the sensitivity parameters were increased to 10%. The results are presented in Table 2.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Angle (degree)</th>
<th>Lighting Power (W)</th>
<th>Solar Thermal load (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>453</td>
<td>1940</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>540</td>
<td>1799</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>464</td>
<td>1930</td>
</tr>
</tbody>
</table>

A single objective optimization control strategy in which the lighting power is minimized would result in an angle of 21 degrees with lighting power of 435 W and Solar Thermal load of 2164 W. This solution has 11.5% higher solar thermal load while the decrease in lighting power is only 4% compared to Solution 1. Therefore the solution identified by RRPARTEO3 algorithm for scenario 1 represents a good trade off between these two conflicting criteria. The Pareto optimal set is shown in Figure 5. The compromise solution obtained for scenario 1 is near the kink where the curve flattens in the Pareto plot. Intuitively, this is a good location that provides reasonable balance between the conflicting objectives.
Figure 6. Parallel axis plot showing the trade-off between lighting power and solar thermal load
Trade-offs between lighting and thermal load are easily visualized using the parallel axis plot shown in Figure 6. In a parallel axis plot, multiple parameters are displayed using vertical axes. In Figure 6, the first axis represents the lighting power, the second axis the solar thermal load and the third axis the optimization variable, the angle of the light shelf. The range of values of these variables are seen in the plot. Each series of lines connecting the axes represent one solution point. For example, the lines connecting the solid dots represent a point close to the compromise solution obtained in Scenario 1. They connect the points lighting power = 447 W, Solar thermal load = 1799 W and angle = 31 degrees. The many lines with positive slope starting from low values of lighting power and meeting high values of solar thermal load indicate that these two criteria conflict with each other. That is, lighting power can only be improved at the expense of solar thermal load. Similarly, the lines with negative slope indicate that the solar thermal load can only be reduced by increasing lighting power. The results for simulations at 10 am are presented in Figures 7 and 8.

At 10 am, the sun is on the east side and there is negligible effect of the light shelf on direct solar radiation. However, energy simulations show some reduction in the solar thermal load especially at high angles. This is mainly due to the effects of diffuse radiation from the sky. In general, closing the light shelf (by increasing the angle) reduces the solar thermal load while increasing the lighting power consumption due to reduced light transmission.

The compromise solutions obtained for 10 am are summarised in Table 3. The compromise solution obtained in Scenario 1 is the light shelf angle of 2 degrees. This corresponds to lighting power of 597 W and thermal load of 795 W. For scenario 2, the compromise solution has light shelf angle of 90 degrees. In this case, the light shelf blocks thermal radiations coming from outside allowing no daylight through the upper part of the window.

For scenario 3, the compromise solution is the same as in scenario 2. This is because the light shelf is not very effective as a lighting device at this time and closing the shelf completely does not reduce the lighting power by more than 10%. It is more effective as a shading device that blocks diffuse radiations.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Angle (degree)</th>
<th>Lighting Power (W)</th>
<th>Solar Thermal Load (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>597</td>
<td>795</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>629</td>
<td>693</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>629</td>
<td>693</td>
</tr>
</tbody>
</table>

From tables 2 and 3, it is seen that the solutions obtained are different according to the time of the day, the priorities of the user, and the acceptable deterioration in the performance parameters.

In order to evaluate the performance of the adaptive light shelf with respect to a static horizontal light shelf, the energy consumption for all the solutions are compared for 10 am and 3 pm in Tables 4 and 5. The energy for the hour is computed by adding the lighting power and the power required for air conditioning assuming a coefficient of performance (COP) of 6. At 3 pm when the sun is on the west, up to 7.6% savings in total energy is obtained compared to a static light shelf. Even in the morning when the sun is on the east, the light shelf is able to save 1.2% of the total energy in scenario 1. At 10 am for the scenario 2, there is an increase in the total energy because energy was not
explicitly used as an optimisation objective. In this case, solar thermal load was the primary objective and the reduction in cooling load actually caused the total energy to increase due to the increase in lighting load.

### Table 4 Energy Savings at 3 pm

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Angle (degree)</th>
<th>Energy (WH)</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31</td>
<td>845</td>
<td>7.6%</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>894</td>
<td>2.3%</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>863</td>
<td>5.6%</td>
</tr>
</tbody>
</table>

### Table 5 Energy Savings at 10 am

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Angle (degree)</th>
<th>Energy (WH)</th>
<th>Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>829</td>
<td>1.2%</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>894</td>
<td>-1.3%</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>863</td>
<td>-1.3%</td>
</tr>
</tbody>
</table>

### 4 Future work

A full-scale laboratory prototype has been constructed to evaluate the actual performance of the adaptive light shelf. The control actions identified by the RRPARETO3 algorithm will be applied to the light shelf. Measurements will be taken to evaluate the lighting and thermal performance under different sky conditions and position of the sun.

### 5 Conclusions

The concept of an adaptive light shelf looks promising. Results from simulations show that significant energy savings can be achieved in comparison to a static light shelf. The multi-objective optimization algorithm RRPARETO3 is able to identify good compromise solutions that make reasonable trade-offs between two conflicting criteria namely, lighting power consumption and solar thermal load.

### References


Automatic Positioning and Alignment for Hole Navigation in Surface Drilling


Sandvik Corp.
Pihtisulkatu 9, 33110 Tampere, Finland
E-mail: tuomo.pirinen@sandvik.com
juha.lassila@sandvik.com
mikko.loimusalo@sandvik.com
juha.pursimo@sandvik.com
sami.hanski@sandvik.com

Abstract

Satellite-based positioning is becoming the prevalent solution for positioning and alignment of drill holes in open pit mines, quarries and construction sites. This is due to the accuracy and convenience of the systems, leading to improved blasting quality and thereby reduced costs and improved productivity. In a conventional hole navigation system, the operator is assisted by precision sensors and a satellite positioning receiver which compare the boom position and attitude to those pre-defined by a drilling plan. While such an approach provides a superior accuracy of the holes in comparison to conventional systems, the boom still has to be positioned manually using the rig controls. This work describes an automatic hole positioning and alignment system developed for drill rigs. The novelty of the solution is in the surface drilling application and the versatility of a typical drill boom. The system has been tested both in a simulated and a prototype environment. The main result is that automatic positioning improves the positioning accuracy and reduces the time required for the positioning task. This is because the automatic system overcomes the main restrictions of manual operation, namely the time and attention needed to position the boom given its dynamics and the actuator-based controls. Furthermore, the operation of the automatic positioning system is effortless and fast. These developments lead to further improvements in drilling quality, which in turn yields better blasting results and higher productivity in excavation. The developed system is also an essential part of remote operated or autonomous drill rigs.

Keywords

Automation and control, drill rig, hole navigation, surface drilling, positioning, alignment, autonomous drilling, drill automation

1 Introduction

Drill and blast method is the dominant practice of excavation for moderately to very hard rocks with a low fracture density [1]. This is due to the speed of the method and its cost/yield ratio per excavated volume, i.e. excavation productivity.

Figure 1. A modern surface drill rig.

According to its name, the method involves drilling holes into rock for charging of explosives and consequently detonating the explosives to fracture and move the rock mass. In open pit mines, quarries and construction sites, the holes are drilled using surface drill rigs, see Figure 1 for an example.

To achieve the desired results in blasting, the drill holes need to be positioned and aligned accurately using the boom of the drill rig. This lies at the responsibility of the operator controlling the drill rig.

Utilization of Global Navigation Satellite System (GNSS) positioning has become the prevalent technical solution to the placement of holes. In such a setup, the rig operator is assisted by a system comprising a satellite positioning receiver and precision sensors for the boom. The system measures and compares the boom’s position and attitude to those pre-defined by a drilling plan stored into the system.

While such an approach provides a superior accuracy of the holes in comparison to conventional systems, the boom of the rig still needs to be positioned manually using the rig controls. Typically the boom movements are controlled by a number of 1- or 2-axis levers, joysticks or buttons. The controls may be hydraulic or electric.

The drill rig boom is a complex structure in terms of mechanics and kinematics. It may, for example, consist of six two-directional actuators. Therefore, positioning...
and alignment of the boom includes 12 distinct movements. Regardless of the actual controller configuration, positioning and alignment is a challenging three-dimensional task, especially for inexperienced operators [2].

This paper describes a more user-friendly solution to control the boom automatically using the control system of the drill rig. The feature simplifies the operator’s tasks and provides a faster and more accurate positioning of holes in comparison to manual operation.

Automatic positioning makes it possible to operate one or more drill rigs more conveniently from a remote control station. The automatic mode is also an enabler technology for autonomous drill rigs. Remote operation and autonomy minimize the need for human intervention and presence at the work site. This provides solutions to the present issues of safety, labor cost and limited availability of skilled workforce.

From the dimensioning standpoint the main load for the boom system consists from the drill, the feed system and other components, including the drilling tools. Typical boom masses are in the range of three to ten metric tons. The boom structure is relatively elastic compared to the mass it carries, which has to be taken into account in designing the control algorithms.

The paper is organized as follows. Section 2 explains the positioning and alignment task in more detail and its significance to excavation results. Section 3 describes the principles of satellite positioning systems currently in use, followed by a description of the conventional manual positioning mode in Section 4. The developed automatic positioning and alignment solution is described in Section 5, and its performance and benefits are described in Section 6. Some notes on system development are given in Section 7. The paper is summarized in Section 8.

2 Positioning and Alignment in Drilling

2.1 Positioning and alignment

A drill hole is determined by its starting point on the rock surface, alignment, and depth, as defined in a drilling plan. Prior to actual drilling, the drilling boom must be correctly placed to achieve desired position and alignment of the hole, see Figure 2.

Positioning of a drill hole has to be done with a centimeter level accuracy. A typical requirement is that deviation from the planned location should not exceed the hole diameter.

Positioning is the process of moving the bottom of the drill feed boom, namely the drill bit, to coincide with the above-ground extension of the hole line. That is, the drill bit is positioned such that when pushed against the ground, the bit lies at the desired starting point of the hole. Positioning is typically conducted with a sufficient margin between the lower end of the feed boom and ground surface, in order to avoid the boom from colliding into ground while being moved. Ground may be uneven and there may also be rocks and other obstacles in the area of movement.

Alignment is the process of turning the feed boom, parallel to the hole line. This is conducted in order to drill the hole in the direction defined in the drilling plan.

Note that positioning and alignment are, in principle, independent of each other. In practice, however, changing the alignment of the boom typically affects the drill bit location, and vice versa, due to the mechanics of the boom structure.

Depending on the situation and the operator’s work practices, the order of positioning and alignment tasks may vary and they may also happen simultaneously.

2.2 Typical errors

Figure 3 provides examples of typical errors in drilling related to positioning and alignment. Incorrect positioning (Figure 3.a) yields a hole that is otherwise correct but at an incorrect location.

Incorrect alignment (Figure 3.b) yields a hole that starts at the right location, but the rest of the hole cylinder...
and ultimately the bottom of the hole are at incorrect locations. Note that an extensive error in alignment will result in the hole bottom to lie at an incorrect depth, too. If both the position and alignment are erroneous (Figure 3.e), the above effects are even more severe.

Hole deviation (Figure 3.d) is not directly attributed to incorrect positioning and alignment. In rock conditions where deviation is likely to occur, an incorrect alignment may trigger or increase the deviation, especially if the desired alignment has been specifically planned to minimize the deviation. Hole deviation may also add to the errors caused by incorrect position or alignment.

2.3 Contribution to excavation results

A high accuracy of drilling is crucial for controlled, safe blasting of the rock, which directly affects the costs of rock excavation and thus the productivity.

Positioning and alignment errors in the first row may result into reduced burden (distance between the first row of holes and the free edge of bench), increasing the risk of flyrock and consequent hazards during blasting.

Positioning and alignment errors affect also the distribution of explosives in the rock, and thus the specific charging changes. This results into suboptimal fragmentation of rock producing unevenly sized rock containing oversized boulders and/or an excess fraction of fines. In addition, the floor of the rock may become uneven. These factors complicate and add cost to loading and crushing operations. Furthermore, additional breaking or blasting of boulders may be required. The fragmentation of rock also affects the end product quality and yield in aggregate quarries and is thus a major economical factor.

There are also secondary effects from drilling accuracy and blasting results. An uneven or highly fractured rock floor is more difficult to operate on afterwards and may require filling to make the ground surface even. Drilling holes through fractured or filled rock floor contains a risk for holes to collapse or get blocked quickly after drilling has been completed. Controlled blasting assists also in conserving slope stability and thus in making the operation of the quarry safer and easier.

In [3], it has been estimated, based on three case studies, that the cost savings in drilling and blasting only attributed to satellite positioning are 25%. The savings arise, e.g., from productivity, reduced idle times, improved mine planning, and improved fragmentation.

3 Satellite-based navigation in drills

GNSS-based hole navigation systems have become an industry standard for accurate positioning of holes. See [4] for a system example. In such systems, the required accuracy is attained with real-time kinematic (RTK) positioning. This approach uses two satellite antennas: one for the position information and the second for obtaining the compass direction of the drill rig frame.

In addition to the GNSS instrumentation, the drill rig is equipped with precision sensors for measuring the orientations of all boom joints. These are needed to calculate the drill bit position and feed boom alignment in the selected coordinate system.

The desired hole locations are provided by a drill plan that is downloaded into the navigation system. The plan is a file containing the locations and inclinations of the holes in a format understood by the system. Typically the plan is formatted according to the IREDES standard [12]. By combining the data from the positioning and kinematics calculation and comparing this to the information attained from the drill plan, the system guides the operator while steering the drill feed boom into the correct position and inclination. See Figure 4 for an example GUI.

In practice there is always some difference between the planned and realized hole locations. These differences are documented by storing the actual hole positions and inclinations into the navigation system. The data can be used, for instance, to review the planned charging of explosives, to monitor the quality of drilling or to optimize the overall drill and blast process.

In many worksites the documentation and review of hole locations data prior to blasting is mandatory. The data can be further supplemented by storing and combining measurement while drilling (MWD) data into the analysis.

The GNSS hole navigation system can also be utilized without a drilling plan. In these cases the hole locations are determined by the operator or have been marked on the ground. The system then documents the realized hole locations for further reference.
4 Manual positioning and alignment

The conventional method of hole positioning and alignment is manual direct hydraulic control of the boom that is assisted by the GNSS hole navigation instrumentation as described in Section 4. The feed boom positioning and alignment work phase begins after the operator has driven the drill rig into a location from which it is possible to reach the hole to be drilled. Here, “to reach” means that the feed boom can be aligned and positioned to correspond the drill plan.

The operator uses the controls provided by the drill rig to adjust the boom position and inclination. A typical drill rig boom has six degrees of freedom (six two-directional joints). A distinct control, e.g., a lever or a button, is required for each of the joints. As 12 distinct movements are needed, a (2x3), (3x4), or (6x2) control configuration is necessary.

From the user control commands, the control system of the drill rig forms control signals to the hydraulic valves regulating the oil flow to the boom actuators, which are typically hydraulic cylinders. In the direct control mode, each valve and actuator is controlled individually.

The resulting movement is measured by the boom sensors. The position and attitude of the boom are computed with forward kinematics based on the sensory data and provided to the operator via the system display.

There are downsides in the manual direct control of boom actuators. Firstly, regardless of the control configuration, positioning the boom manually is a complex three-dimensional task requiring considerable attention from the operator.

Secondly, movement of one actuator and joint has an effect on the overall boom position or attitude and therefore the target angles for the other joints are affected. For example, a change in the boom turn angle may change the feed boom alignment and drill bit position. If a joint or actuator is moved, further corrective movements in other joints may be needed. Therefore, it is not possible to reach the desired position and alignment by just controlling each of the joints respectively.

For these reasons, manual positioning and alignment is a time consuming task, especially for inexperienced operators. While the GNSS system greatly assists in the task, the resulting hole navigation speed accuracy depend ultimately on the operator’s skill and experience. These drawbacks can be overcome with an automatic positioning system.

5 Automatic positioning system

The GNSS-based hole navigation system described in Sections 3 and 4 enables to automate the positioning and alignment process. In the automatic mode, the control system calculates the desired joint angles for the boom structure and controls the valves and actuators jointly so that the desired joint angles are met. With an automated control, the operator’s task is greatly simplified as the automation can be activated, for example, by using a hold-to-run -type switch. With such a control mode the operator just supervises the actions and interferes only when necessary.

Block diagram of the automatic positioning system is given in Figure 6. Target values for the feed boom position and alignment are obtained from the drill plan. Combining the targets and forward kinematics data on the current boom position, the target joint angles can be computed using inverse kinematics (IK) [5]. Once the
current and target values for joint angles are known, automatic positioning and alignment can be done by using controllers to regulate each of the valves.

The challenges to control the boom can be divided into three categories. First, the boom kinematics has to be controlled in a reasonable way. The second challenge is to measure the boom angles accurately and fast enough in a hostile environment. The third challenge is to maintain the stability and the performance of the nonlinear feedback loop including hysteresis, state dependent inertias and a fast IK solver.

Drill rigs operate in hostile environments where heavy vibrations, reactive chemicals, dust and other sharp hard particles may be present. This requires the instrumentation to be as minimal as possible for reliability reasons. In addition to reliability, a minimal instrumentation is important for keeping the purchase price of the rig as low as possible. Thus, the solution cannot be based on improving the instrumentation.

The kinematics and the feedback control have a tight link together. Inverse kinematics is a part of the control loop and a major factor in the observed response of the system. Consequently, the inverse kinematics and the feedback control problems cannot be solved separately.

The IK problem is nonlinear and typically leads to an iterative solution. The solution of the joint angle problem is typically ambiguous. However, inside a control loop the solution should be made unambiguous to avoid random discontinuations that cause severe control problems. Our approach to solution selection is to minimize an $l^2$-norm. One commonly used method is to use Tikhonov regularization [6]. See [11] for some typical solution methods for IK problems.

In our case coordinate system is fixed to the operator’s position in cabin. The hole coordinates are given in the global coordinate system. These two coordinate systems have to be kept constantly synchronized during the positioning.

The boom control is a nonlinear control. A Bayesian probability [7] based linear-quadratic-Gaussian (LQG) control or a robust control method could have been used. However, in a nonlinear case the error distribution is state dependent. There is relationship between the nominal performance, the robust performance and the robust stability [8]. However, the characterization of the modeling errors is not easy and robust methods tend to result in control laws that do not fulfill the requirements for accuracy and speed.

To achieve the requirements, while lacking a formal approach, a heuristic ad hoc control law was utilized. The control law is basically a simple proportional control with gain scheduling [10] with a few heuristic modifications.

The use of heuristics was attempted to keep as little as possible with maximum simplicity, because excessive use of heuristics makes it difficult to generalize the system. The main factor guiding the development was to make sure that the system is robustly stable at all stages, including mode changes.
6 System benefits and performance

The automatic positioning feature has to be faster and more accurate in comparison to the manual positioning mode. The operation has to be responsive with reasonable trajectories. Given the mass of the boom and the forces involved, the movements have to be initiated and stopped smoothly to avoid unnecessary loads on structures.

The accuracy and speed of positioning was measured using operators as test subjects in a drill rig simulator. Simulated environment was selected to allow repeatability of experiments and accurate measurement of precision and positioning times.

The tests were conducted in two separate sessions. In the first session, a version of the algorithms was tested. In the test, six operators were asked to perform a pre-defined positioning task on three respective holes, resulting into N=18 test cases for both systems. In the second test session, a modified version of the algorithms was tested, now with six holes for each of the operators (N=36).

The averages of measured positioning errors are given in Figure X and positioning times in Figure Y. For both test cases it is clear that the automatic mode is faster and more accurate in comparison to the manual mode. The error of the manual mode is about 11.5 cm in both cases and is reduced to 5.5 cm in the first case and 6.7 cm in the second case. The average positioning time is reduced in the first case by 64% from 2:19 [m:ss] to 48 seconds, and by 76 % in the second case from 1:29 (m:ss) to 21 seconds. The variations in the positioning times in the two test cases are explained by the differences in the positioning setup.

The time required for manual positioning is explained by the precision of the instrumentation. The use of a GNSS hole navigation provides a good accuracy, especially in comparison to traditional systems where hole positions were determined visually without any instrumentation. However, it takes time to position the boom manually within the precision provided by the GNSS system. The automatic system does not have this limitation, and is able to achieve an even higher accuracy. In addition to these quantifiable results, the automatic mode is more pleasant to use for the operator, see [2] for more details on usability.

7 Development notes

The drill rig control system is a complex distributed system. Development and integration of a new feature into a distributed control system is challenging with regard to development time and performance verification. The algorithms developed herein were complicated, making these challenges even harder.

Had the new functionality been developed directly into the control system, the time required for each change-test iteration would have been long. This would have increased the total development time and cost unnecessarily. These issues were tackled with a rapid control prototyping (RCP) system developed in collaboration with a third party.

The RCP system uses an external computer connected to the drill rig control system. The computer communicates with the control system, reads the necessary inputs for the development-stage algorithms, computes the valve control commands and transmits them back into the control system. With the developed RCP system, the external algorithms could be run without interference or modifications to the actual drill rig control system or its environment.
software.

Early phases of development were done within a simulated environment (see Figure 9) to warrant safety and to avoid the practical constraints related to testing on an actual drill rig. Once the functionality of the algorithms was proven with the simulator, the setup was transferred onto an actual rig for testing and refinement. At both stages the observed modification needs could be almost instantaneously implemented and tested.

Once complete, the developed features were integrated into the control system using code generation. In this manner, there was no need to give further specifications or to re-program the algorithms for integration. This resulted into further savings in development time. Furthermore, there is no risk of introducing additional errors in programming as the algorithms are contained in the code generation results and no modifications are needed. The time savings were not formally measured but observed to be major, and in line with other cases where high-level programming languages and code generation is used [13].

8 Summary

This paper has introduced an automatic positioning and alignment system for surface drill rigs. The system is based on a conventional GNSS-based hole navigation system that has been supplemented with an automatic control mode comprising of target generation, inverse kinematics and control algorithms that have been integrated into the system.

The system was developed utilizing RCP methods that resulted in reduced development lead time. The main challenge in the development was to achieve a responsive yet stable control behavior, given the mechanical and dynamical properties of a drill rig boom. The system does not require additional sensors in comparison to the existing GNSS hole navigation system.

The developed automatic mode is more accurate and faster in comparison to manual positioning. It is also more pleasant for the operator to use. This provides an improved workflow and productivity, as well as an improved drilling quality with consequent benefits of production and cost to the overall excavation process. Authors believe that the developed system will become an industry standard in terms of safety, performance and automated equipment.

References


A Study on the Thermal Crack Control of Large Turbine Foundation using Automated Curing System

Ju-hyung Ha\textsuperscript{a}, Ok-pin Na\textsuperscript{a}, Chang-keun Lim\textsuperscript{a} and Yun-gu Cho\textsuperscript{a}

\textsuperscript{a} Hyundai Engineering & Construction, Research & Development Division
102-4 Mabuk-dong, Giheung-gu, Yongin-si, Gyeonggi-do, 446-716, Korea
E-mail: jhha@hdec.co.kr, okpin.na@hdec.co.kr, cklim@hdec.co.kr, yungu.cho@hdec.co.kr

ABSTRACT

The thermal crack occurrence from hydration heat is one of the most important factors that significantly affect structural quality and construction period in mass concrete. Therefore, appropriate methods to control the thermal crack are necessary for mass concrete. In this study, the probability of thermal cracking was checked by FEM analysis prior to the construction of turbine foundation in a combined thermal power plant. Subsequently, the change of concrete mix design and application of automated curing system were proposed to prevent thermal crack occurrence.

The proposed concrete mix design and automated curing method have been applied to actual turbine foundation construction site and the effect of the proposed thermal crack control methods has been evaluated through field measurement of the temperature, strain, thermal crack occurrence.

Keywords

Mass concrete, Automated curing system, Thermal crack control

Introduction

Recently, it is unlikely to predict electricity demands due to abnormal temperature, uncertainty of the increase rate of electric charges, expansion of economic volatility, etc. Under the circumstances, 28.8 % of the whole existing generators are more than 20 years old, and the number of failures rapidly increases, which expands the uncertainty of power supply.[1]

In constructing combined thermal power plants under the recent rise trend, a turbine is the most core building, and the foundation structure of the turbine takes most important role. The foundation of the turbine should be able to withstand the electromotive force caused by its own weight and engine operation, and the vibration caused by operation should not make harmful effects on other equipment or buildings. In order to reduce the transmission of vibration, we generally increase the weight of the foundation, and the concrete foundation is a very massive structure, thus in the massive concrete, the temperature stress caused by temperature difference of the structure due to cement hydration heat frequently generates cracks in the structure or makes considerable effects. The measures of the thermal cracking control include appropriate materials, mix design, mixing temperature, curing method selection, etc.[2]

As a part of thermal cracking control measures of large turbine foundation concrete, this study conducted hydration heat and temperature stress analysis with variables including changes in concrete mixing, and application of curing methods actively controlling inside and outside temperature differences.

Moreover, based on analysis results, we applied the changed concrete mixing and curing methods to actual construction, and evaluated thermal cracking control effects in the turbine foundation concrete structure through concrete temperature and strain rate monitoring and crack investigations.
FEM Analysis Method and Condition

Analysis Model

In thermal analysis, finite element method (FEM) is widely used for the analysis of temperature distribution in concrete foundation and FEM can take into account geometry shape, material properties, and arbitrary boundary conditions. In this study, a multi-purpose FEM software package in civil engineering, DIANA, was employed. The size of mass concrete model is 34.3m long, 8.7m wide, and 2m deep. In Figure 1, a half model of mass concrete was visualized with 8-node solid finite elements due to the symmetric structure instead of whole model.

Material Properties of Concrete

The mix design for concrete is shown in Table 1. For the thermal analysis, Portland cement Type 1 (OPC) and slag cement (40% replacement) were used, respectively. The characteristics of concrete adiabatic temperature rise used in numerical analysis were determined by actual test. The concrete material properties such as thermal conductivity, specific heat, tensile strength, young’s modulus referred to Manual of concrete practice.[3-4] (ACI 207.2R, 209R)

Curing Condition

For thermal analysis, automated curing method was applied to control the thermal cracking of mass concrete and was compared with the results of general curing method.

(1) General curing method

In general, wet covering with blanket is widely used for the concrete moist-curing treatment. The period of curing and mould release is about 7 days after concrete placement.

(2) Automated curing method

Automated curing method is a novel curing method to control thermal crack occurrence based on real-time temperature monitoring equipment. The principle of this system is the automated water circulation system of heated curing water. Applied curing water on the surface of mass concrete keeps the temperature difference between center and surface of the structure below the threshold of specification.[5] This system can reduce the occurrence probability of thermal cracks, and the validity of this system has been already turned out from the application of mock-up test[6] and construction field.[7-9]

Analysis Results and Discussion

Temperature Analysis Results

The measuring locations of thermal analysis were marked in Figure 2.

Table 1 Concrete mix proportion

<table>
<thead>
<tr>
<th>Mpa</th>
<th>W/C</th>
<th>S/A</th>
<th>Unit Weight(kg/m³)</th>
<th>Air (W C S G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>49.6</td>
<td>46.8</td>
<td>170</td>
<td>346</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>834</td>
<td>970</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.3</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1-Modelling for Thermal Analysis

Figure 2-Locations of Temperature Results
There are two types of analysis parameters: cement (OPC, Slag cement) and curing method (General curing, Automated curing). The temperature difference between surface and center of concrete over time is indicated in Figure 3 and Table 2.

As a result of temperature analysis, the maximum temperature of slag cement concrete can be dropped down up to 3°C rather than that of OPC concrete. The time of maximum hydration temperature of slag cement concrete was between 48 hours and 50 hours after placing and this time period was delayed by about 24 hours comparing with the period of OPC concrete.

In British standard, the temperature difference for controlling thermal cracking is prescribed at 20°C.[10] As shown in Figure 3, in general curing method, the temperature difference rises up more than 20°C within the early curing period of about 1~2 days. At the age, thermal cracks could be initiated due to the early low tensile strength of concrete. On the other hands, in automated curing method, the maximum temperature difference between centre and surface of concrete is much lower than general curing method, and the time exceeding over 20°C of temperature difference criterion is 7 days after concrete placement when the application of the automated curing method was finished and the early tensile strength could be more developed relatively. Therefore, the possibility of thermal crack can be much lower than in case of applying general curing method.

**Thermal Stress Analysis Results**

The locations of thermal stress analysis are shown in Figure 4.

<table>
<thead>
<tr>
<th>Type</th>
<th>Curing</th>
<th>Max. Temperature</th>
<th>Max. Temp. Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Temp (℃)</td>
<td>Age (day)</td>
</tr>
<tr>
<td>OPC</td>
<td>General curing</td>
<td>74.2</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Auto curing</td>
<td>74.2</td>
<td>1.1</td>
</tr>
<tr>
<td>Slag</td>
<td>General curing</td>
<td>70.8</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Auto curing</td>
<td>70.9</td>
<td>2.1</td>
</tr>
</tbody>
</table>
of applying general curing in OPC concrete. On the other hand, the automated curing system can reduce the thermal stress by 26% as shown in Figure 5(b).

![Figure 5-Results of Thermal Stress - OPC Concrete](image)

(a) General Curing

(b) Automated Curing

Figure 5-Results of Thermal Stress - OPC Concrete

In Figure 6, the external restraint stress of slag cement concrete can be decreased by 22% rather than that of OPC at 30 days age. At the beginning stage of concrete curing, the tensile stress of concrete surface resulting from the internal restraint can also be reduced by 9% in comparison with OPC. As shown in Figure 6 (b), automated curing method can be applied in order to diminish tensile stress, and the tensile stress can be reduced by 21%.

From the results of stress analysis, it was confirmed that the change of mix design and curing method can have effects on the control of thermal restraint stresses, and slag cement and advanced curing method were decided to be used for increasing thermal crack resistance in the field application.

**Field Application**

**Introduction**

In this study, as the target structures for analysis and measurement, two turbine foundations of combined cycle thermal power plant with the size of 34.3m long, 8.7m wide, and 2m deep were selected. In two concrete foundations, slag cement was used as concrete binder since slag cement was proved to be effective to control thermal crack from the analysis results. The monitoring of temperature and strain and the investigation of surface cracks were conducted in both foundations.
Curing Method

In two mass concrete foundations, the insulation curing method using styrofoam was applied in one foundation (T-1) and the automated curing method was applied in another foundation (T-2). The insulation curing method is widely used for retaining the curing temperature when the ambient temperature is not too low.\textsuperscript{[11]} For the insulation in this field test, the surface of concrete specimen was covered with vinyl and then 50mm thick of styrofoam was placed. Also, the whole surface of concrete foundation was wrapped one more time with another insulation blanket in order to minimize the effect of environmental condition such as water evaporation, wind and rain.

For applying automated curing method, the curing blanket and curing water pipe were installed on the surface of concrete after concrete placement. When the temperature difference between centre and surface exceeded the criterion for crack control (generally 20°C), heated water around 60 °C was supplied automatically on the surface of structure to reduce the temperature difference.

Measurement

Table 3 shows the types of sensors embedded in concrete for measuring temperature, strain of concrete and for the operation of automated curing system. The measurement of temperature and strain was carried out on the surface, centre, bottom, and upper edge of concrete. The monitoring of temperature and strain was respectively kept for 3 weeks and for 1 week after concrete placement.

<table>
<thead>
<tr>
<th>Type</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-Button</td>
<td>Temperature Monitoring</td>
</tr>
<tr>
<td>Thermocouple (T-Type)</td>
<td>Operating for Automated Curing Method</td>
</tr>
<tr>
<td>KM-100B</td>
<td>Strain Monitoring</td>
</tr>
</tbody>
</table>

Table 4 Temperature Monitoring Results

<table>
<thead>
<tr>
<th>Type</th>
<th>Maximum Temperature</th>
<th>Maximum Temperature Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulated Curing Method (T-1)</td>
<td>71.5 °C</td>
<td>22.7 °C</td>
</tr>
<tr>
<td>Automated Curing Method (T-2)</td>
<td>67.9 °C</td>
<td>20.0 °C</td>
</tr>
</tbody>
</table>

The temperature measurement results of T-1 foundation showed that the maximum temperature of the central concrete was 71.5 °C, and during the thermal insulation curing period, temperature difference between the top surface and the centre was maintained mostly at 10°C, which implies that the probability of cracking by inside and outside temperature difference at initial placement is very low. However, it can be thought that after the end of curing, the top surface is exposed to the air and rapidly cooled down, which makes a temperature difference rise up to 22.7°C and increases the probability of cracking. In T-2 foundation, the maximum temperature of the centre increased up to 67.9 °C, and during the entire measurement period, the temperature difference between the top surface and the

Measurement Results and discussion

Temperature Measurement Results

Figure 7 and Table 4 show measurement results of temperature differences between centre and surface of concrete.
centre was maintained below 20℃. As keeping surface temperature at a certain level using heated curing water, inside heat of T-2 foundation was smoothly released rather than T-1 foundation and temperature difference could be effectively reduced at the end of curing.

**Strain Rate Measurement Results**

The results of strain rate measurement of foundation concrete applying the insulation curing method (T-1) and the automated curing method (T-2) are shown in Figure 8 and Table 5.

![Figure 8-Strain Histories - Measurement versus Analysis](image)

<table>
<thead>
<tr>
<th>Location</th>
<th>Measure / Analysis (μ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Upper Surface</td>
</tr>
<tr>
<td>Insulated Curing Method (T-1)</td>
<td>216.8</td>
</tr>
<tr>
<td>Automated Curing Method (T-2)</td>
<td>155.5</td>
</tr>
</tbody>
</table>

According to the strain rate measurement results, the maximum strain rate change in the centre and the top surface was 100~200μ in both T-1 foundation and T-2 foundation, and it was 250~450 μ in the upper edge portion, which was a larger strain rate compared to that of other locations. Comparing T-1 foundation with T-2 foundation, in the thermal insulation curing, the strain rate of the upper surface and upper edge was greater than that of T-2 foundation by 39%, 62%, respectively. Additionally, the maximum strain rate of the centre was also about 25% greater than that of T-2 foundation. This is consistent with the trend of FEM analysis results. Moreover, in T-2 foundation, when comparing measurement values with analysis ones, approximately +4~22% differences appeared, while in T-1 foundation, +41%~48% differences appeared. In T-2 foundation, the measured value difference of the strain rate between the upper edge and the centre was maximum 155.5 μ, while it was maximum 303.9 μ in T-1 foundation, which was by 96% higher than that of T-2 foundation. Therefore, it is thought that the probability of cracking due to concrete internal restraint is higher in T-1 foundation.

**Crack Inspection**

The crack inspection of mass concrete foundations cured with insulated curing method and automated curing method was conducted at 28 days after concrete placement. The results of surface crack inspection after releasing the concrete moulds are demonstrated in Figure 9 and Table 6.
Figure 9-Diagram of Cracking at T-1(Insulated Curing)

Table 6 States of Cracking

<table>
<thead>
<tr>
<th>Crack Width</th>
<th>Insulated Curing Method</th>
<th>Automated Curing Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 0.15 mm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>0.10 ~ 0.15 mm</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>0.05 ~ 0.10 mm</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>0.05mm &gt;</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Based on the results of surface crack inspection in Figure 9, the total number of cracks of concrete with insulation curing method was 8 spots: 3 of them were placed on the surface of concrete and 5 on the side. On the other hands, in case of T-2 foundation, there was no crack on all sides of mass concrete. The field test results were well matched with the simulation results of thermal analysis. Therefore, it could be confirmed that the automated curing method is valid to control the thermal cracks of mass concrete.

Conclusions

In this study, in order to control thermal cracking of the turbine foundation concrete of large power plants, we conducted hydration heat analysis, field application and measurement with variables including changes in concrete mix design, curing methods, and the results are as follows:

1. In case of mass concrete, there is a risk of cracking due to internal and external restraint caused by hydration heat, and it is thought that the effects of risks and measures can be predicted through preliminary hydration heat analysis similarly to actual construction.

2. By changing type1 normal cement concrete mixing to slag cement concrete for low hydration heat, it was possible to lower the probability of cracking caused by external restraint of large mass concrete structures.

3. By applying the curing automation system that actively controls inside and outside temperature difference, we were able to control inside and outside temperature difference below 20°C at the initial concrete placement, which controlled the surface cracking caused by hydration heat internal restraint.

4. Upon applying thermal insulation curing, we could control inside and outside temperature difference to a very small degree, but when we removed the thermal insulation materials after the end of curing, the surface rapidly cooled down, which caused surface cracking from thermal shock. To prevent this, the thermal insulation should be appropriately controlled by preliminary hydration heat analysis.

5. The turbine foundation of large power plants is a very important structure, and due to the size, the risk of cracking caused by hydration heat is high. However, after reviewing the probability of cracking by hydration heat analysis, when required, we can control cracking appropriately through applying low heat concrete mix design and active temperature control curing.

References


[2] Korea Concrete Institute, Concrete standard specifications, Korea Concrete Institute, 2009.


of Automated Curing System for Mass Concrete. 


Integrated Approach to Machine Guidance and Operations Monitoring in Tunnel Construction

X. Shen\textsuperscript{a}, M. Lu\textsuperscript{b}, S. Mao\textsuperscript{b} and X. Wu\textsuperscript{b}

\textsuperscript{a}School of Civil and Environmental Engineering, University of New South Wales, Australia
\textsuperscript{b}Department of Civil and Environmental Engineering, University of Alberta, Canada
E-mails: x.shen@unsw.edu.au, mlu6@ualberta.ca, mao1@ualberta.ca, xwu10@ualberta.ca

Abstract -
Tunnel construction method using a tunnel boring machine (TBM) is commonly adopted for building underground infrastructure, such as railways, roads, sewers, or utility pipelines. TBM tunnelling entails precise guidance of the machine in the underground space, as well as effective construction management and project control. The research aims to address critical engineering and management problems during the course of tunnel excavation, including TBM guidance, automated as-built data acquisition, real-time data processing and 3D visualization. In this paper, we propose an integrated TBM guidance and operations monitoring solution for tunneling applications. A robotic total station is employed to automate the continuous processes of TBM tracking and guidance inside the tunnel. Wireless sensor networks are particularly implemented for on-site data communication. The results of TBM’s position state, tunnel alignment and construction progress are processed and presented in straightforward, user-friendly interfaces on the fly. Working with the City of Edmonton, Alberta, Canada, the integrated system has been implemented in the construction of a 2.4-meter-diameter and 1-km-long sewage tunnel project and undergone seven-month field testing in 2013. The solution lends substantial support to TBM operators and project managers in making critical decisions on a near real-time basis.

Keywords -
Automation and Control; Tunnel Construction; Machine Control and Guidance; Wireless Sensor Networks; 3D Visualization; Project Management and Control

1 Introduction

Tremendous demands for sustainable development of urban areas have driven widespread applications of tunneling techniques for underground construction. The resulting subsurface infrastructure, such as metro railway networks, traffic tunnels, drainage and storm water pipelines, electricity cables and gas lines, etc., are vital to the economy and security of the community [1]. Compared against the conventional hand-digging and cut-and-cover methods, tunnel construction using a tunnel boring machine (TBM) is more preferable especially in well-developed areas, thanks to its enhanced productivity, safer environments provided to the working crews, and less disruption to surface activities [2].

For operators in the tunnel construction field, steering a TBM is like driving a vehicle in complete darkness. It is crucial to precisely position the TBM in the underground space, so as to guide its advancement closely along the as-designed alignment [3]. For tunnel quantity assurance, tolerances of alignment deviations in both horizontal and vertical planes are generally specified in tight margins, such as ±75 mm for the entire tunnel project [4]. However, unforeseen underground obstacles and variable geologic conditions present significant challenges to tunnel alignment control. In practice, it is not unusual that TBM operators and site managers are caught by surprise with excessive out-of-tolerance tunnel alignment errors. It will take weeks or longer time to determine the exact alignment deviations by surveying specialists, and then figure out a practical strategy to steer the TBM back on track. Sometimes, the TBM can be trapped in the ground, requiring considerable time, cost and effort for recovery; in worst-case scenarios, the TBM has to be abandoned due to prohibitively high cost and overwhelming effort in rescuing it.

TBM guidance control currently relies on a laser station which projects a laser beam onto a laser target board mounted on the TBM [5]. However, laser target boards lack accuracy and reliability, thus potentially leading to increased risk and uncertainty in tunnel construction [6]. Meanwhile, current manual methods for tunnel progress data collection undermine both efficiency and effectiveness of project management and
control. The operations data, like tunnel alignment deviations and TBM’s advance rate, are manually recorded by TBM operators and project manager on a daily basis. The resulting data are often kept as separate paper-based records, which need days or even weeks to be compiled, post processed and analyzed. The analysis results often come back too late to be helpful.

The research aims to address critical engineering and management problems during the course of tunnel excavation, including TBM guidance, automated as-built data acquisition, real-time data processing and 3D visualization. We propose an integrated TBM guidance and operations monitoring solution for tunneling applications. A robotic total station is employed to automate the continuous processes of TBM tracking and guidance inside the tunnel. Wireless sensor networks are particularly implemented for on-site data communication. The results of TBM’s position state, tunnel alignment and construction progress are processed and presented in straightforward, user-friendly interfaces on the fly.

The remainder of this paper first reviews the current practice for TBM guidance, followed by a description of the integrated system design. Finally, field evaluation of the proposed system in a 2.4-meter-diameter and 1-km-long sewage tunnel project in Canada is presented.

2 Current Practice for TBM Guidance

Laser systems currently predominate in TBM guidance applications. Typically, a laser guidance system consists of a tunnel laser station and a laser target board. As shown in Figure 1, a laser station is mounted on precast concrete tunnel lining segment. The coordinate of the laser station is first determined by traditional tunnel surveying. The orientation and gradient of the laser beam are then calibrated to be parallel with the as-designed tunnel alignment. Figure 2 shows a laser target board installed on a TBM. When the laser beam is projected onto the target board, TBM operator can infer the current line and grade tunnel alignment deviations by reading offset of the laser spot from the centre of the target board.

Despite extensive applications and low cost of the traditional laser guidance systems, several major technical limitations contribute to relatively low accuracy and reliability of the technology, including (1) potential manual errors in initializing or calibrating the laser station, (2) dispersion and refraction of the laser beam over a long distance, and (3) difficulty to receive laser’s projection over a long distance or due to excessive deviations of the TBM [6]. As such, maximum application distance for the laser guidance system is around 200 m. For long tunnel projects, the laser station needs to be relocated every 200 m of tunnel excavation. The surveying work for moving the laser station in the tunnel is tedious and time consuming, requiring a crew of specialist surveyors spending one hour to calibrate the laser beam, and around five hours to set up a new laser station. Meanwhile, all of the tunnelling operations have to be halted during this work process. As a result, the tunnelling productivity can be considerably undermined by operation and maintenance of the laser guiding system.

Figure 1. Laser station mounted on precast concrete tunnel liner

Figure 2. Laser target board installed on TBM

Corporations specialized in tunnel guidance have developed advanced laser guidance systems by integrating electronic or video laser target units to digitize the laser spot position on the target [7, 8]. On the down side, the high complexity in system design considerably increases the system’s price tag and consumption cost, including system maintenance and technical service [6].
Integrated System for TBM Guidance and Tunnel Operations Monitoring

An integrated solution is proposed in this research to automate TBM guidance and monitor tunnel operations in real time. The main functions include TBM positioning, wireless field data collection, online analytic data processing and visualization in a three dimensional (3D) environment. Given in Figure 3, a robotic total station is employed to automate the continuous processes of TBM tracking and spatial data collection inside the tunnel. Two reference targets with known coordinates are utilized to initialize and calibrate the total station inside the tunnel. The real-time position of the TBM in the underground space is then fixed accurately by the total station through continuously tracking a target installed on the TBM, as shown in Figure 3. Meanwhile, wireless sensor networks are particularly implemented for on-site data communication among the total station, TBM guidance tablet computer, and a monitoring laptop (see Figure 3). The analytical results of TBM’s position state, tunnel alignment and construction progress are presented in straightforward, user-friendly 3D interfaces, which aid TBM operator and project managers in making critical decisions on a near real-time basis.

3.1 Surveying Automation and Real-Time As-Built Data Acquisition

The advanced surveying instrument of robotic total station is utilized in the research, serving as a critical geomatics tool for automated TBM positioning and as-built data collection. Compared against traditional surveying instrument, the robotic total station incorporated on-board computer, digital signal processing and self-driven motor. Therefore, sophisticated functions can be provided by the robotic total station, such as: (1) automatic target recognition, (2) automatic target tracking, and (3) programmability for computer controlled automatic surveying and data transmission.

The robotic total station is remotely controlled by a rugged tablet computer in the system. During tunnel operations, control commands are sent in wireless to trigger the surveying procedures, including automatic tracking of the target on the TBM, as shown in Figure 3. TBM’s coordinates with time stamped are returned in serial data package, which are further processed to derive the TBM guidance information in the tablet. Meanwhile, intelligent self-calibration functions are incorporated in the system design, by which the robotic total station can automatically detect any displacement of its own standpoint by checking two reference targets.

The low-cost and power-efficient wireless sensor networks are implemented for data communication inside the tunnel. As such, the total station can be wirelessly linked up with the tablet computer in the
TBM and the monitoring laptop on the surface. Figure 4 shows a prototype of the surveying automation system we developed in the Construction Automation Lab at the University of Alberta, Canada.

The guidance information is updated in real time, assisting TBM operator making crucial decisions in an intuitive and straightforward manner.

3.2 Virtual Laser Target Board (VLTB) Program

A VLTB program is developed in house to automate TBM surveying and generate guidance information on the tablet computer. Three subsystems are integrated in the program, namely, the surveying subsystem, the communication subsystem and the control subsystem [9]. In the surveying subsystem, the robotic total station locks the coordinates of the target by a pre-scheduled plan or on request from the control subsystem. The survey data are sent via the communication subsystem to the control subsystem. The communication subsystem is responsible for data transmission through the wireless sensor networks infrastructure. The communication subsystem acts like a black box, and handles input/out data using standard serial data communication protocol. The control subsystem handles user interaction, survey control, data integrity check and failure recovery.

Figure 5 shows the VLTB user interface providing guidance information to the TBM operator. The as-designed tunnel alignment passes through the center of cross. A red circle and a green triangle represent the rear end and the cutter head of TBM, respectively (overlapped in Figure 5). The steering guidance is neatly simplified as a process to keep the red circle and green triangle within the square boundary as much as possible. Moreover, the red triangle arrows suggest the direction of the next manoeuvres for the TBM operator (turning right and downward as in Figure 5). The numbers on the right side of the interface indicate line-level deviations, yaw-roll-pitch angles, advancing speed of TBM and chainage distance of the as-built tunnel.

3.3 Data-driven 3D Tunnel Visualization

A straightforward, user-friendly 3D visualization platform is further developed to present the analytical results of TBM's position state, tunnel alignment and construction progress [10]. The visualization program running on the monitoring laptop aids project managers in making critical decisions on a near real-time basis.

The architecture of the 3D tunnel visualization program is given in Figure 6. Two categories of input data are combined to model the as-built tunnel and its surrounding environment, including static data and time dependent data. Static data are design parameters derived from construction drawings and geotechnical reports, for instance as-designed tunnel alignment, soil layers, existing utilities, and other related information. These data are inputted to the database manually. Time dependent data refer to TBM tracking data, which are sourced from the VLTB system and are autonomously inputted by a data feed program. The resulting tunnel visualization interface is shown in Figure 7.

Figure 4. Robotic total station remotely controlled by a tablet computer through wireless sensor networks

Figure 5. User interface of the VLTB program

Figure 6. Architecture of tunnel visualization program
Field Evaluations

From August 2012 to March 2013, the integrated TBM guidance and tunnel operations monitoring system was implemented and underwent seven months testing on a 2.4-meter-diameter, 1-km-long sewage tunnel project (West Edmonton Sanitary Sewer - WESS, Stage W13) in Edmonton, Alberta, Canada. Figure 8 shows the LOVAT open-face TBM utilized in the project.

The robotic total station (model: Leica TS15) was mounted on the as-built tunnel liner using a specially designed surveying bracket, as shown in Figure 9. Constrained by the narrow surveying window available in the tunnel, one tracking target (model: Leica GMP101 mini prism) was installed at the center of the existing laser target board (see Figure 10a). Figure 10b shows one of the two reference targets (model: Leica GPR121 round prism) installed in the tunnel.

The maximum range to enable the automation function of the robotic total station for surveying the tracking target was found to be around 250 m inside the tunnel. Therefore, similar to the operations of the traditional laser station, the total station had to be moved forward from time to time as tunnel operations unfold. During the testing, the two reference targets were utilized to initialize the coordinates for a newly installed total station, which is enabled by a traditional surveying method called resection. As such, relocation of the total station can be completed within 30 min, as opposed to five hours required to move the laser station.

Wireless sensor nodes (model: SENA ZigBee ZS10) were deployed onsite every 150 m from the tunnel face along the tunnel lining segments, and access shaft to the surface trailer office, as illustrated in Figure 11a-d. A total of 12 sensor nodes established the wireless communication network covering the 1-km-long tunnel and the site on the surface. It is found from the extensive site testing that the wireless sensor networks can provide reliable data communication among the robotic total station, the tablet computer and the monitoring laptop. Figure 12 shows the tablet computer installed inside the TBM, which provided real-time guidance information to TBM operator during the testing.
Figure 11. ZigBee wireless sensor nodes deployed on site: (a) at rear end of TBM’s trailing grantry; (b) along precast concrete tunnel liner; (c) at bottom of access shaft; and (d) in the site office

Table 1 shows consecutive testing results for 70 min on March 13, 2013. It is found the tunnel had been excavated over 934 m. The TBM advanced 0.809 m during this testing period, while the line and level deviations were maintained about 13 mm and 63 mm, respectively. It is noteworthy that some intervals of data were not available, e.g. 32 min between 10:29 and 11:01. This is mainly due to short period blockage of the line-of-sight by workers and tunnelling facilities. The VLTB program was designed to handle the situation properly, resulting in continuous operation afterward. When there are no readings from the system, a message will pop up and remind the TBM operator for checking the visibility of the tracking target.

Figure 12. TBM guidance tablet computer mounted in the steering panel of the TBM

5 Conclusions

Today, utility and traffic tunnels are commonly constructed using the mechanized tunnelling method. Steering control of a TBM remains a challenging task due to complicated ground conditions, such as unforeseen obstacles and variable geologic conditions. Traditional laser guidance systems fall short in terms of accuracy and reliability. The lack of effective TBM guidance and operations monitoring solutions adds to high risks in executing tunnelling projects, potentially leading to out-of-tolerance alignment deviations, project delay and budget overrun, or even failure of the tunnel project.

An integrated system for TBM guidance and tunnel operations monitoring has been developed in this research. Through embedding automation control mechanisms and innovative computing algorithms, the

<table>
<thead>
<tr>
<th>Date (yyyy/mm/dd)</th>
<th>Time (hh:mm:ss)</th>
<th>TBM coordinates (m)</th>
<th>Chainage (m)</th>
<th>Deviations (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013/03/13</td>
<td>10:24:01</td>
<td>27743.706 5934141.476 644.721</td>
<td>934.514</td>
<td>15 -62</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>10:25:07</td>
<td>27743.703 5934141.435 644.721</td>
<td>934.555</td>
<td>18 -62</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>10:29:37</td>
<td>27743.705 5934141.314 644.721</td>
<td>934.676</td>
<td>17 -62</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:01:26</td>
<td>27743.714 5934140.717 644.719</td>
<td>935.273</td>
<td>11 -65</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:02:08</td>
<td>27743.714 5934140.694 644.72</td>
<td>935.296</td>
<td>11 -64</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:04:16</td>
<td>27743.713 5934140.633 644.718</td>
<td>935.357</td>
<td>12 -66</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:17:24</td>
<td>27743.715 5934140.53 644.721</td>
<td>935.460</td>
<td>11 -63</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:19:40</td>
<td>27743.717 5934140.531 644.72</td>
<td>935.459</td>
<td>9 -64</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:33:35</td>
<td>27743.713 5934140.667 644.72</td>
<td>935.323</td>
<td>12 -64</td>
</tr>
<tr>
<td>2013/03/13</td>
<td>11:34:42</td>
<td>27743.711 5934140.667 644.722</td>
<td>935.323</td>
<td>14 -62</td>
</tr>
</tbody>
</table>
system transforms a mature survey tool, the robotic total station, into a tunnelling control robot which precisely tracks and positions the TBM. Accurate TBM positions are automatically determined by the robotic total station, so as to derive line and grade deviations of tunnel alignment in real time. Meanwhile, intelligent self-calibration functions are incorporated in the system design, by which the robotic total station can automatically detect any displacement of its own standpoint by checking two fixed “landmark” points with known coordinates inside the tunnel. The low-cost and power-efficient wireless sensor network solution is implemented for data communication inside the tunnel. As such, the total station can be wirelessly linked up with a control computer on the surface, where data are analyzed to generate (1) real-time 3D visualization of the working TBM, (2) as-designed vs. as-built tunnel sections, and (3) any anticipated obstructions and geological information in the underground workspace.

Field testing of the prototype system on a 1-km-long sewer tunnel project in Canada validated the technical feasibility and effectiveness for real-time data acquisition. It is also found that the automation system can save on project shut-down periods for relocation surveying, from the current 5 hours to less than 30 minutes. Therefore, application of the new system would lead to significant enhancement of tunnelling productivity. The solution also lends substantial decision support to not only tracking the construction progress but also visualizing any tunnel alignment deviations on the fly. With sufficient project data accumulated, the cases and knowledge gained from the new solution can be used for industry training.

The future research will enhance system reliability through more field applications. Meanwhile, this research mainly investigated TBM guidance in straight tunnel alignment. Applications in curved tunnel sections deserve rigorous further investigations.

Acknowledgment

This research was substantially supported by the TECTERRA’s Alberta University Applied Research Funding Program (1108-UNI-008). The writers thank Mr. Siri Fernando of the City of Edmonton for his vision and unreserved support. The writers also thank Mr. Tim Schneider, Mr. Chris Pratt, Mr. Junhao Zou, Mr. Darsh Nawaratna, and Mr. Ray Davis of the City of Edmonton for facilitating field testing.

References

Dynamic Risk Assessment in Construction Projects Using Bayesian Networks

Limao Zhang\textsuperscript{a,b}, Xianguo Wu\textsuperscript{a}, Miroslaw J. Skibniewski\textsuperscript{b,c}, Jingbing Zhong\textsuperscript{a}

a. School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China
b. Department of Civil & Environmental Engineering, University of Maryland, College Park, MD, 20742-3021, USA
c. Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Poland

E-mail: limao_zhang@hotmail.com, wxg0220@126.com, mirek@umd.edu, jb_zhong@126.com

Abstract -

This paper presents a systemic Bayesian network (BN) based approach for dynamic risk assessment for adjacent buildings in tunnel construction. This approach consists of four steps in detail, namely, hazard analysis, BN learning and BN-based risk analysis. In the dynamic risk analysis framework, the predictive, sensitivity and diagnostic analysis techniques in the Bayesian inference are used to conduct the feed-forward control in the pre-construction stage, intermediate control in the construction stage and back-forward control in the post-accident stage, respectively. A case relating to dynamic safety risk analysis of some existing buildings adjacent to construction of the Wuhan Yangtze Metro Tunnel in China is presented. Results demonstrate the feasibility of the proposed approach, as well as its application potential. The proposed approach can be used by practitioners in the industry as a decision support tool to provide guidelines on the conservation of adjacent buildings against tunnel-induced damages, and thus increase the likelihood of a successful project in a dynamic project environment.

Keywords -
Dynamic risk assessment; adjacent buildings; tunnel construction; Bayesian network; case study

1 Introduction

In the last ten years, urban tunneling projects have increased substantially as a result of rising populations, space restrictions, and growing environmental concerns. Tunneling excavation in soft ground inevitably leads to ground movement, which may cause adjacent surface buildings to deform, rotate, distort, and possibly sustain unrecoverable damages, especially those founded on shallow foundations [1, 2]. Many existing buildings are aging and do not have complete load-bearing capability as designed, resulting in very low deformation resistance. Thus, the tunnel-induced ground movement may destroy these buildings, unless accurate risk analyses are conducted and appropriate protection measures are implemented [3]. In order to assure the safety and serviceability of adjacent buildings in tunnel construction, it is therefore necessary to explore the safety risk mechanism for the tunnel-induced damage to adjacent buildings.

To prevent heavy casualties and property losses caused by safety violations due to tunnel-induced damages, application of a probabilistic risk assessment (PRA), which is a systematic and comprehensive methodology to evaluate risks associated with a complex engineered technological entity, has been widely reported in literature [4]. However, conventional PRA has the disadvantage of being static and fails to capture the variation of risks as deviations or changes during the life of a continuous process [5]. When associated parameters, such as geological, design and construction parameters are changed with the development of construction projects, the above traditional PRA methods cannot accurately illustrate the updated features of dynamic environments as the construction progress evolves. Accordingly, accurate risk analyses and decision support cannot be conducted in real time as these parameters are updated.

To overcome disadvantages of conventional PRA methods, dynamic risk analysis provides a possible way to cope with the dynamic nature of the risk profile. Dynamic risk analysis is a continuous process of identifying risk, assessing, and determining a way to reduce or eliminate risks in a dynamic manner [6]. Recently, there have been many efforts to simulate the dynamic nature of process behaviors. With the capacity of integrating prior knowledge and sample data, Bayesian network (BN) provides a powerful tool for knowledge representation and reasoning under a dynamic environment [7]. BN allows explicit modeling of changes over time, and can therefore model the evolution of the probabilistic dependencies within a random system. Basically, BN allows designers to easily update the prediction when additional information becomes available, and is especially suitable for engineering applications, where statistical data is often sparse [8]. This paper therefore investigates the possibility of using BN techniques to address the
potential dynamic nature underlying the risk analysis and management in tunnel construction. A systemic BN-based approach with detailed step-by-step procedures is proposed for dynamic risk analysis, including predictive, sensitive and diagnostic analysis throughout the overall construction process. A case in relating to the dynamic safety risk analysis of some existing buildings adjacent to construction of the Wuhan Yangtze Metro Tunnel in China is presented. Results demonstrate the feasibility of the proposed approach, as well as its application potential.

2 Methodology

Dynamic risk assessment is a complex activity and requires several steps. Taking advantage of BN inference techniques, a BN-based approach is developed for dynamic risk analysis, making efforts to improve the effectiveness and accuracy of safety management in a dynamic project environment. In the proposed approach, the overall workflow includes the following three steps:

2.1 Hazard analysis

The risk assessment process starts with questions, such as “What can go wrong and how can it go wrong?” The identification of what and how it can go wrong entails defining hazards, risk events, and risk scenarios [9]. Hazard analysis involves determining which risks/risk factors might affect the project safety and documenting these characteristics. The identification is considered a difficult task for a complex system, particularly in tunnel construction. Nevertheless, past experiences provide extensive prior knowledge for risk identification. With the development of tunnel construction practice worldwide, large amounts of scattered knowledge were accumulated [10]. During the hazard analysis process, the failure modes, internal variables, exogenous factors, explicit cause and effect relationships are determined. The main outputs will be used as the input to the BN model in the next step.

2.2 BN learning

A BN model is defined by two components: structure and parameters. The structure is a graphical and qualitative illustration of the relationships among the nodes using directed arcs, while the parameters represent the quantitative probabilistic relationships among the nodes using probabilities. In this stage, a BN model is developed to simultaneously integrate the structure of the system, the variables and causal mechanisms (or interdependencies) analyzed by the last step. The design of a BN model involves determining the network structure and its parameters. There are two procedures in this step, as follows:

1. Structure learning. Structure learning aims to determine the proper DAG, confirming the relationship between nodes. Every variable in the real-world situation is represented by a Bayesian variable. The BN variables (nodes) should be first created according to the results of hazard identification in the last step. Then the network structure can be developed by creating the directed edges from the node corresponding to fault causes to the node representing its consequence, which is indicative of a conditional dependency between the variables it links. In the meantime, considered as one of the most commonly used techniques for risk and reliability studies, FTA is a logic diagram that displays the interrelationships between a potential critical event and the causes in a system. Thus, the approved fault trees in construction fields can be applied to provide effective prior knowledge for BN structure learning.

2. Parameter learning. Parameter learning aims to determine the conditional probability distribution of each node under the established BN structure [11, 12]. The conditional probability tables can be determined by learning the parameters on the database using a learning algorithm. For instance, the K2 algorithm is a well-known algorithm for BN parameter learning [13, 14], and can be adapted under the established BN structure. Also, expert judgment is an alternative when certain database information is unavailable.

2.3 BN-based risk analysis

When the established BN model is validated within the acceptable range, the BN model can be used to conduct various types of analysis. Ren et al. [15] indicated that the most important use of BN is in revising probabilities in light of actual observations of events. It is therefore possible to calculate the probability distribution of potential safety risks and identify the most likely potential causes in occurrence of accidents. In this paper, we mainly discuss the predictive, sensitivity and diagnostic analysis using the Bayesian inference.

1. Predictive analysis

Predictive analysis aims to capture the probability distribution of the risk event (T) under a combination of root nodes (X_1, X_2, ..., X_n). The states of each root node and intermediate node can be treated as evidence input into the BN model. Compared with traditional FTs/ETs, the Bayesian inference in BN models does not need to get minimal cut sets, which increases greatly the computational efficiency. The probability distribution of T, represented by P(T=E), can be calculated by Eq. (1). At the same time, in light of actual observations of events, for instance, X_i is observed to stay in the state of q_i (X_i = x_i^q), and the probability distribution of T, represented by P(T=E|X_i = x_i^q), can be calculated by
Eq. (2) under given evidence. Both \( P(T=t) \) and \( P(T=t\mid X_i = x_i^k) \) can serve as indicators to evaluate the risk of \( T \), assisting construction decision makers to take proper preventive measures in advance.

\[
P(T=t) = P(T=t\mid X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) \times P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)
\]

\[
t = \{t_1, t_2, \ldots, t_p\}, \quad x_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,0}\}, \quad i = 1, 2, \ldots, n
\]

\[
P(T=t\mid X_i = x_i^k) = \frac{P(T=t, X_i = x_i^k)}{P(T=t)}
\]

where, \( t \) stands for the state of a risk-prone event \( T \) with \( P \) states; \( \{t_1, t_2, \ldots, t_p\} \) is a range of \( P \) states for the risk event \( T \); \( x_i \) stands for the state of risk factor \( X_i \) with \( Q \) states; \( \{x_{i,1}, x_{i,2}, \ldots, x_{i,0}\} \) is a range of \( Q \) states for a root node \( X_i \); \( P(T=t\mid X_i = x_i^k) \) represents the conditional probability distribution of \( T \); and \( P(X_1 = x_1, \ldots, X_n = x_n) \) represents the joint probability distribution of the root nodes.

(2) Sensitivity analysis

Sensitivity analysis is particularly useful in investigating the performance of each risk factor’s contribution to the occurrence of an accident. The most natural way of performing sensitivity analysis is to change the values of input parameters, and then monitor the effects of changes on the output probabilities. In this research, a performance-based indicator, Sensitivity Performance Measure (SPM) is proposed to measure the contribution of each risk factor \( X_i \) to risk event \( T \). Key risk factors can then be identified to help the decision maker determine the main checkpoints in the construction phase. Under the prior probabilities, the SPM of each root node \( X_i \), represented by \( SPM(X_i) \), can be calculated by Eq. (3). In light of actual observations of events, for instance, \( X_i \) is observed to stay in the state of \( q_i( X_i = x_i^q ) \), and \( SPM(X_i) \) can be calculated by Eq. (4) under given evidence. \( SPM(X_i) \) can be used as an indicator to measure the degree of sensitivity of the root node \( X_i \) in the accident occurrence. Factors that are very sensitive to the accident occurrence should be given more attention during the construction process to reduce the risk limit.

\[
SPM(X_i) = \frac{1}{Q_i} \sum_{q_i} \frac{P(T=t\mid X_i = x_i^q) - P(T=t)}{P(T=t)}
\]
bearing capability as designed initially, and some kinds of structural damages are likely to occur in the process of long-term operations [19, 20]. The structural health condition of an existing building provides a basis as to how much additional deformation or load it is able to bear, which is very important for the safety of adjacent buildings in tunnel construction. Such parameters as Building Value (X9), Building Intact Conditions (X10) and Structure Configuration (X11) are all related to the quality of the building health condition. Furthermore, the Horizontal Distance (X8) between the tunnel structure and the adjacent building is another factor that should be included, since the magnitude of the tunnel excavation effect appears to be slowed down as the building foundation is further away from the tunnel structure [21].

(4) Mechanical variables (B4). In the process of shield-driven tunneling excavation, engineers pay close attention to the measurement of some mechanical variables, to maintain the face stability of the excavation and minimize settlements [22]. Some pressure and speed sensors are installed on the top and middle of the cutter head. These monitored parameters, including Driving Speed (X12), Thrust Force (X13), Cutter Torque (X14), Cutter Speed (X15), Cut Slurry Pressure (X16), Soil Pressure (X17), Grouting Pressure (X18) and Grouting Amount (X19), are very sensitive to geologic conditions, and should be adjusted to adapt to the changing surrounding environments.

3.2 Risk modelling

In tunnel construction practice, the daily measurement of ground settlement is reviewed as a basic means for the safety assurance of surface and subsurface buildings. According to some technical specifications in China, such as “Technical code for monitoring measurement of subway engineering (DB11/490-2007)”, the ground settlement should be controlled within 30 mm, and the nearby buildings are regarded unsafe if exceeding this control standard. However, the actual observed value turns out to be random due to the uncertainties and complexities underlying complex project conditions, and a single predicted value is met with significant limitations. In this situation, we use the fuzzy set theory to divide the predicted value into several ranges. The safety status of shield tunnel construction can then be assessed by analyzing the chance of the predicted value among different ranges. As to ground settlement with a general distribution range of 0~70 mm, we divide its predicted value into the following five ranges, namely, I (Very Safe, 0~20 mm), II (Safe, 20~30 mm), III (Dangerous, 30~40 mm), IV (Very Dangerous, 40~50 mm) and V (Extremely Dangerous, 50~70 mm). Each range corresponds to one risk level in regard to the tunnel-induced building damage (T). Between 2006 and 2013, researchers at Huazhong University of Science and Technology developed some safety control systems for metro tunnel construction and operation tasks for Shenyang, Zhengzhou, Shenzhen and Wuhan Metro systems. The researchers have also developed early warning web-based systems for safety control of each project. Large amounts of monitoring records have been accumulated during the work progress on these projects [23, 24]. According to the BN learning process as mentioned in Section 2, the accumulated 1000 training samples are used to conduct the TIBDN, as seen in Fig. 1.

4 Results

A case of three surface buildings adjacent to the construction of the Wuhan Yangtze Metro Tunnel (WYMT) in China is presented in this research. WYMT, known as “the first metro tunnel across the Yangtze River in China”, is an important route connecting two large cities, comprising the metropolitan area of Wuhan, namely Wuchang and Hankou. According to the site investigation, there are 35 buildings within 30 m offset the tunnel centerline in total, among which 11 buildings are located directly above the tunnel structure. Due to the complex failure mechanism of tunnel-induced building damage and poor geological conditions, the safety risk analysis and management of existing building adjacent to WYMT is considered a challenging task. In this case study, three buildings, denoted by $B^{3}$, $B^{5}$ and $B^{19}$, are randomly chosen and taken as examples to present the detailed computation process.

Each risk event has a life cycle, namely before, during and after an accident or failure. Therefore, safety analysis of risk-prone events can be divided into three stages in the overall work process: namely, feed-forward control in the pre-construction stage, intermediate control in the construction stage and back-
forward control in the post-accident stage. Taking advantage of powerful reasoning features in the Bayesian inference, predictive, sensitivity and diagnostic analysis techniques are used to conduct the safety control of the above three stages, respectively. In this way, TIBDN is offered as a decision support tool, and thus, real-time and effective support can be available for decision makers in a dynamic manner in the entire life cycle of risk-prone events.

4.1 Feed-forward control

Feed-forward control aims to determine the probability distribution of the tunnel-induced building damage ($T$) using the predictive analysis technique of the Bayesian inference in the pre-construction stage of tunnel construction. Since most of the important decisions are made in the pre-construction stage, this stage plays a significant role in guaranteeing the safety of the tunnel construction and adjacent buildings. To be specific, in the conceptual design, a deep understanding about the factual situation related to the potential safety status of adjacent buildings is lacking, since no definite information about the tunnel construction is provided. However, in this situation (defined by Scenario A), prior probabilities of root nodes ($X$) can be entered into TIBDN as input evidence. The probability distribution of $T$ within each risk level can then be obtained using Eq. (1), as seen in Table 1. The results indicate that the potential safety status of $T$ corresponds to a level of IV [Very Dangerous] under Scenario A, since $P(T=IV) > P(T=III) > P(T=II) > P(T=I)$. Scenario A can be viewed as a general situation where all existing nearby buildings are involved. In this way, the impact of the tunnel excavation on general existing buildings can be assessed without much given information.

During the construction survey and design phases, the values of other influential variables for a specific building can be obtained. In the preliminary design phase, the state of root nodes ($X_1$-$X_19$) can be determined (see Table 1), and subsequently used as given evidence in the Bayesian inference. For simplification, the situation regarding $B^A$, $B^B$ and $B^C$ can be represented by Scenarios $B$, $C$ and $D$, respectively. Thus, we list the variable values of each scenario, enter their current variable states (I, II,...,V) into TIBDN as given evidence, and then calculate the probability distribution of the risk event ($T$) using Eq. (2). The results as seen in Table 1 indicate that the safety risk of all these three buildings is rated at a level of III [Dangerous]. In other words, in locations of the mentioned three surface buildings, the ground settlement induced by tunneling excavation is likely to fall into a range of 30-40 mm, which is still beyond the allowed safe range of the safety control standard. Thus, to reduce the risk limit, the construction decision makers will make some further adjustments and optimizations based on the previous scheme according to the calculated results. Using the same Bayesian inference process, the updated calculation results of the probability distribution of $T$ are shown in Table 1. As one might expect, the safety risk of these three buildings then tends to decrease to a level of II [Safe]. In this way, the construction scheme can be optimized continuously until the high potential safety risk is under control.

<table>
<thead>
<tr>
<th>Stages</th>
<th>Scenarios</th>
<th>Evidence {X1,X2,…..X19}</th>
<th>$P(T=I)$</th>
<th>$P(T=II)$</th>
<th>$P(T=III)$</th>
<th>$P(T=IV)$</th>
<th>$P(T=V)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual design</td>
<td>A</td>
<td>Prior probabilities</td>
<td>0.017</td>
<td>0.082</td>
<td>0.316</td>
<td>0.432</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>{III, II, V, IV, III, V, II, I, III, III, IV, III, II, III, IV}</td>
<td>0.088</td>
<td>0.144</td>
<td>0.257</td>
<td>0.224</td>
<td>0.287</td>
</tr>
<tr>
<td>Preliminary design</td>
<td>C</td>
<td>{IV, IV, IV, II, V, III, II, IV, IV, IV, IV}</td>
<td>0.072</td>
<td>0.176</td>
<td>0.354</td>
<td>0.257</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>{II, III, V, III, I, III, V, III, IV, IV, IV, IV}</td>
<td>0.108</td>
<td>0.169</td>
<td>0.308</td>
<td>0.259</td>
<td>0.155</td>
</tr>
<tr>
<td>Optimized preliminary design</td>
<td>B</td>
<td>{III, II, V, IV, III, V, III, II, I, II, III, III, II, III, III}</td>
<td>0.150</td>
<td>0.269</td>
<td>0.177</td>
<td>0.246</td>
<td>0.158</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>{IV, IV, IV, IV, IV, IV, IV, IV, I, II, III, II, III, II, III}</td>
<td>0.017</td>
<td>0.432</td>
<td>0.316</td>
<td>0.082</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>{II, III, V, III, I, III, V, III, IV, IV, II, II, III, II, III, III}</td>
<td>0.088</td>
<td>0.224</td>
<td>0.257</td>
<td>0.144</td>
<td>0.287</td>
</tr>
</tbody>
</table>
4.2 Intermediate control

Intermediate control aims to identify critical and sensitive factors in occurrence of construction failures using the sensitivity analysis technique of the Bayesian inference. In the construction stage, some variables, including tunnel related variables (B1), geological variables (B2) and building related variables (B3), have been determined and cannot be changed as a matter of fact, and thus, engineers pay much more attention to mechanical variables (B4) which have flexibility to be adjusted in order to maintain the face stability of the excavation and minimize settlement. In the intermediate control, the most sensitive mechanical variables can be identified as key check points in the process of shield-driven tunneling excavation.

With regard to general surface buildings in Scenario A (prior probability) of this case, the values of all influential variables are unknown, and Eq. (3) can be used to calculate the performance sensitivity of all root nodes $PSM \left(X_i\right)$ ($i=12,13,...,19$). The results as seen in Fig. 2 (a) indicate that X15 (Cutter Speed), X14 (Cutter Torque) and X12 (Driving Speed) become the top three sensitive factors when the tunnel-induce building damage ($T$) falls into a risk level of V (Extremely Dangerous), since $X_{15} > X_{14} > X_{16} > X_{13} > X_{17} > X_{19} > X_{18}$ in the sensitivity analysis results in case of $P(T=V)=1$. As seen in Fig. 6 (a), X15, X14 and X12 are more likely to become the most sensitive factors when $T$ falls into a risk level of IV (Very Dangerous). Meanwhile, when $T$ falls into a risk level of III (Dangerous), X13, X15 and X12 turn out to be the top three sensitive factors. In general, under the prior probability situation, X12, X13, X14 and X15 should be considered as the critical check factors for real-time measurement and adjustment until the high potential safety risk ($P(T=III, IV \ or \ V)$) is under control.

Fig. 2. Results of sensitivity analysis for mechanical parameters in: (a) Scenario A (Prior probabilities); (b) Scenario B; (c) Scenario C; and (d) Scenario D.

With regard to specific buildings in Scenarios B, C and D of this case (see Table 1), the values of all influential variables are determined and entered into TIBDN as given evidence, and Eq. (4) can then be used to calculate the performance sensitivity of all root nodes $PSM \left(X_i\right)$ ($i=12,13,...,19$). The results as seen in Fig. 2 (b)-(d) indicate that there are some changes in the sensitivity of root nodes when the states of influential...
variables are observed, and contribution of each root node \(X_i\) to the leaf node \(T\) varies in different scenarios. In order to simplify the sensitivity analysis, the sensitivity of each variable can be evaluated in terms of average sensitivity measure (as shown in a polyline in Fig. 2 (b)-(d)), given the tunnel-induced building damage \(T\) lies in a high risk level \(P(T=\text{III}, \text{IV} \text{ or V}).\) In Scenario B, X16, X15 and X14 can be considered as the most sensitive variables to the occurrence of a high safety risk level of \(T\), as seen in Fig. 2 (b). In Scenario C, X14, X15 and X16 can be regarded as the top three sensitive variables in case of a high safety risk level of \(T\), as seen in Fig. 2 (c). In Scenario D, X15, X14 and X13 should be the top three sensitive variables in case of a high safety risk level of \(T\), as seen in Fig. 2 (d). As a consequence, key check points should be updated given the observed states of influential variables are different among existing buildings. Accordingly, the major focus of concern for safety management strategies can be updated among different scenarios during the construction process.

4.3 Back-forward control

In current construction practice, construction managers are likely to invite domain experts to join an expert group meeting in case of an accident, and then the experts discuss proposing some control measures. This is likely to miss the critical opportunity of handling problems, causing more serious losses. Back-forward control aims to identify the suspected causes using the diagnostic analysis technique of the Bayesian inference, in order to facilitate the real-time fault diagnosis once an accident occurs.

With regard to general surface buildings in Scenario A of this case, Eq. (5) is used to calculate the posterior probability distribution of the risk factors (X12-X19), given that the tunnel-induced building damage \(T\) lies in a high risk level, that is \(P(T=\text{IV})=1\) as an example.

The results as seen in Fig. 3 (a) indicate that X14=IV (with a 45.2% chance) and X19=IV (with a 46.9% chance) are most likely to occur in case of \(P(T=\text{IV})=1\). For this reason, the fault diagnosis should concentrate on these two factors, and the practical check confirms our deduction. As a consequence, both X14=IV and X19=IV can be entered into TIBDN as additional evidence for the subsequent diagnostic analysis. The results as seen in Fig. 3 (b) show that both X12=I and X13=IV are more likely to occur in the second diagnosis cross, which should be the focus of practical diagnosis in the next cross.

With regard to specific buildings in Scenarios B, C and D of this case, the values of all influential variables are determined and entered into TIBDN as given evidence, and Eq. (5) can then be used to calculate the posterior probability distribution of the risk factors (X12-X19), given \(T\) lies in a high risk level of IV (Very Dangerous), as an example. The results in relating to \(B^{3}\), \(B^{2}\) and \(B^{0}\) are shown in Fig. 4 (a), Fig. 5 (a) and Fig. 6 (a), respectively. The suspected causes leading to an occurrence of \(P(T=\text{IV})=1\) can be easily detected, and the practical check against these suspected causes is followed subsequently. According to the similar Bayesian inference, the observed values of the suspected causes are then entered into TIBDN as additional evidence for the next diagnostic analysis and the posterior probability distribution of other variables are shown in Fig. 4 (b), Fig. 5 (b) and Fig. 6 (b), respectively. In general, the diagnostic analysis results can provide new evidential information for the diagnostic analysis of the next cross, and the posterior probabilities of relevant factors can be updated in a dynamic manner. In this way, the evolution route of accidental occurrence can be extracted in real time, and at the same time, the high dependency on domain experts can be reduced.

Fig. 3. Fault diagnosis under Scenario A in: (a) the first diagnosis cross \(P(T=\text{IV});\) and (b) the second diagnosis cross \(P(T=\text{IV}|X15=\text{X19}=\text{III}).\)
AUTOMATION AND CONTROL

5 Conclusions

In the past ten years, tunnel construction has presented a powerful momentum for rapid economic development worldwide, especially in China. Tunnel excavation produces a significant disturbance to adjacent buildings, and safety violations occur frequently due to complex tunnel-soil-building interactions. This paper presents a systemic BN-based approach with detailed step-by-step procedures regarding dynamic risk analysis for adjacent buildings in tunnel construction. A case in relating to the safety risk analysis of some existing buildings adjacent to construction of the Wuhan Yangtze Metro Tunnel is used to verify the applicability of the proposed approach. There are also some limitations to the developed systematic approach. Large quantities of monitoring records which serve as training and testing samples have been obtained from web-based systems developed for this research. Numerous engineering technicians have participated in the monitoring work, making an essential contribution to securing regularly
scheduled input for the daily monitoring data into the system from project sites. This process is laborious and susceptible to human error. Future work will focus on developing a real-time intelligent monitoring system using automatic data acquisition technologies.

Acknowledgements

The National Science and Technology Support Plan (51378235), Wuhan City Construction Committee Support Projects (201208, 201334) and China Scholarship Council (CSC) are acknowledged for their financial support of this research.

References


[23] Zhang L, Wu X, Ding L, Skibniewski MJ, Yan Y. Decision support analysis for safety control in...

Nonlinear Flatness-Based Controller for Permanent Magnet-Excited Synchronous Motor

Pham Tam Thanh\textsuperscript{a} and Nguyen D. That\textsuperscript{a, b}

\textsuperscript{a}Faculty of Electrical and Electronics Engineering, Vietnam Maritime University, Vietnam
\textsuperscript{b}Faculty of Engineering and Information Technology, Sydney, NSW 2007, Australia
E-mail: phamtamthanh@vimaru.vn, That.NguyenDinh@uts.edu.au

Abstract:
This paper addresses the problem of nonlinear discrete-time flatness-based controller design for a Permanent Magnet-Excited Synchronous Motor (PMSM). To eliminate the static errors of the system state variables and consider the nonlinear characteristics of a PMSM, a cascaded flatness-based control scheme is proposed. Simulation results are provided to illustrate the effectiveness of the proposed control structures, in terms of better performance.

Keywords:
Flatness-based control; Permanent Magnet-Excited Synchronous Motor (PMSM); nonlinear control; real-time control.

1 Introduction
With the advantages of superior power density, high performance motion control with fast speed and enhanced accuracy, Permanent Magnet-Excited Synchronous Motors (PMSM) have been increasingly used in robotics, precision machining and many automation processes. Therefore, the problem of control design for PMSM has received considerable attention. However, there remain interesting questions as to how to design a controller so that the static errors of the system state variables are minimized in which the nonlinear characteristics of the PMSM are taken into account.

The concept of differentially flat nonlinear systems was first introduced by Fliess et al., 1992. The system is considered to be flat if a set of outputs can be found such that all states and inputs can be determined from these outputs without integration. The main purpose of the flatness-based control method is first to design an open-loop nominal control corresponding to the predicted trajectory of the flat output. Then, a feedback control law is applied to stabilize the real trajectory around the predicted trajectory of the flat output. The flatness-based control has been recognized as a promising method to deal with nonlinear systems (see, e.g., Levine, 2009 and references therein). Based on the combination of the natural energy dissipation properties of the permanent magnet stepper motor system with its differential flatness property, a nonlinear feedback controller was proposed in Ramirez, 2000. To minimize the copper loss at all operating points of the PMSM, a hierarchical flatness-based control scheme was developed in Delaleau et al. 2004. In Dannenh and Fuchs 2006, a nonlinear differential flatness-based control was proposed to the induction machine fed by a voltage source converter in which the flatness-based control was used for the inner current and outer flux and speed loops. Another approach for drive systems with elastically coupled loads was reported by Thomsen and Fuchs 2010. By using fuzzy logic technique to eliminate the effects of the time-varying nonlinearities of an induction

\begin{tabular}{|l|l|l|}
\hline
Symbol & Unit & Description \\
\hline
A, B, N, S & & Matrices of model \\
\hline
f & & Nonlinear function \\
\hline
L_s, L_r & H & Stator, rotor inductance \\
\hline
L_m & H & Mutual inductance \\
\hline
R_s, R_r & \Omega & Stator, rotor resistance \\
\hline
u_{sd}, u_{sq} & V & dq components stator voltage \\
\hline
i_{sd}, i_{sq} & A & dq components of stator current \\
\hline
T_s, T_r & s & Stator, rotor time constants \\
\hline
L_{sd}, L_{sq} & H & d axis, q axis inductance \\
\hline
\omega_s, \omega_s & rad/s & Stator circuit velocity, Mechanical rotor velocity \\
\hline
\Theta_s & rad & Angle of flux orientated coordinate system \\
\hline
\sigma & Wb & Total leakage factor \\
\hline
\psi_p & Wb & Pole flux \\
\hline
J & kg m^2 & Torque of inertia \\
\hline
m_W & Nm & Load torque \\
\hline
z_p & & Number of pole pairs \\
\hline
\end{tabular}
motor, a fuzzy differential flatness-based controller was developed by Fan and Zhang, 2011. In Houari et al. 2012, a new differential flatness-based control method was presented for a three-phase inverter with an LC filter. Recently, quasi-continuous implementation of structural nonlinear controller based on direct-decoupling for PMSM was reported in Thanh and Quang, 2013.

It should be noted that in the aforementioned papers, the problem of nonlinear flatness-based control has not been fully investigated and the minimization problem of the system state variable static errors has not received considerable attention. These facts have been motivated us to the present study.

In this paper, the problem of nonlinear discrete-time flatness-based controller design for PMSM is investigated. By considering the nonlinear characteristics of PMSM, a control strategy based on flatness theory is proposed to suppress the system state static errors. Simulation results are given to illustrate the effectiveness of the proposed approach.

The paper is organized as follows. Section 2 presents a nonlinear flatness-based controller design method for PMSM. Simulation results for a standard PMSM are given in Section 3. Section 4 concludes the paper.

2 Nonlinear flatness-based controller design for PMSM

2.1 Flatness-based control

As mentioned in Fliess et al. 1994 and Fliess et al. 1995, the main property of differential flatness is that the state and input variables can be directly expressed, without integrating any differential equation, in terms of the flat output and a finite number of its derivatives. Therefore, the trajectory of input can be determined from desired trajectory of flat output. The general flatness-based control structure consists of a nominal feedforward controller combined with a feedback stabilizing controller as shown in Figure 1.

In this structure, the feedback controller is crucial of importance to compensate the effects of external disturbances and model uncertainties.

2.2 Flatness-based controller for PMSM

To eliminate the system state static errors and consider the nonlinear characteristics of PMSM, a cascade nonlinear flatness-based control structure is proposed in Figure 2.

As shown in Figure 2, the cascade control structure includes two loops which are coupled to each other. The outer loop (speed loop) consists of a proportinal-integral controller (PI-controller) and a current Feedforward block while the inner loop (current loop) containing another PI block combined with a voltage Feedforward block. Here, the current controller for the current loop is first designed to guarantee that $i_{sd} \rightarrow i_{sd}^*, i_{sq} \rightarrow i_{sq}^*$ sufficiently fast with respect to the variations of the desired trajectories $\omega \rightarrow \omega^*$ which will be achieved by the mechanical subsystem (speed loop). Then, a speed controller is synthesized. It should be noted that these PI blocks are used to compensate the current and speed static errors.

2.2.1. Stator current controller

Motivated by Quang and Dittrich, 2008, we consider the continuous-time model current of PMSM as follows

$$\dot{x} = f(x) + H(x)u + h_1(x)u_1 + h_2(x)u_2 + h_3(x)u_3, \quad y = g(x),$$

where

- State vector:
  $$x = [x_1 \quad x_2 \quad x_3]^T = [i_{sd} \quad i_{sq} \quad \theta_s]^T,$$

- Input vector:
  $$u = [u_1 \quad u_2 \quad u_3]^T = [u_{sd} \quad u_{sq} \quad \omega_s]^T,$$

- Output vector:
  $$y = [y_1 \quad y_2 \quad y_3]^T = [x_1 \quad x_2 \quad x_3]^T,$$

$$f(x) = [-c_1 x_1 - d_1 x_2 \quad 0]^T,$$

$$H(x) = \begin{bmatrix} h_1(x) & h_2(x) & h_3(x) \end{bmatrix},$$

$$h_1(x) = [a_{i} \quad 0 \quad 0]^T; h_2(x) = [0 \quad b_{i} \quad 0]^T,$$

$$h_3(x) = \begin{bmatrix} \frac{d_{i}}{b_{i}} \frac{a_{i}}{b_{i}} x_2 - \frac{b_{i}}{a_{i}} x_1 - h_{s}\psi_{p} \end{bmatrix},$$

$$g(x) = [g_1(x) \quad g_2(x) \quad g_3(x)]^T = [x_1 \quad x_2 \quad x_3]^T,$$

and temporary parameters:

$$a = \frac{1}{L_{sd}}; b = \frac{1}{L_{sq}}; c = \frac{1}{T_{sd}}; d = \frac{1}{T_{sq}}.$$

Note that the functions $f(.)$ and $H(.)$ in equation (1) are nonlinear in nature, the ordinary differential equation (1) cannot be solved exactly, and hence the exact form of the discrete-time differential equation is difficult to obtain. Therefore, to obtain the discrete-time current model of PMSM, Taylor’s series expansion is used

$$x(k+1) = x(k) + x(t) \bigg|_{t=t_T} T + \Xi(T),$$

where
where $T$ is sampling period and $\Xi(T)$ is the higher-order terms of the Taylor’s series expansion which can be expressed as follows

$$
\Xi(T) = \sum_{n=2}^{\infty} \frac{1}{n!} x^{(n)}(kT)T^n + \frac{1}{(n+1)!} x^{(n+1)}(\zeta)T^{n+1}, \zeta \in (kT, kT+T).
$$

As the sampling period in the advanced electric drive systems is very small, the higher-order terms in equation (6) can therefore be neglected. By substituting (1) into (5), the discrete-time current model of PMSM is obtained as

$$
x(k+1) = x(k) + T f(x(k)) + TH(x(k))u(k) + \Xi(T),
$$

where $a = \frac{1}{L_{sd}}$; $b = \frac{1}{L_{sq}}$; $c = \frac{1}{T_{sd}}$; $d = \frac{1}{T_{sq}}$.

Note that the nonlinear characteristics in the current model of PMSM (8) is considered in terms of the products between the state variables (current components $i_{sd}(k)$, $i_{sq}(k)$) and input variable ($\omega(k)$).

From (8), by using the property of differential flatness whereas the trajectory inputs $u_{ref,ff}(k)$ and $u_{eq,ff}(k)$ can be directly determined from desired trajectory of flat outputs $i_{sd}^*(k)$, $i_{sq}^*(k)$ and $\omega(k)$. Controllers for voltage feedforward block is proposed as

$${\text{Figure 1. The general flatness-based control structure}}$$

$${\text{Figure 2. Cascaded control structure of flatness-based control of PMSM}}$$
From (9), it can be seen that the coupling effects caused by decoupling current component $i_{sd}$ and $i_{sq}$ can be eliminated by \( \frac{a}{b} \tau_{i_s}^s(k) \omega_s^s(k) \) and \( \frac{b}{a} \tau_{i_q}^q(k) \omega_s^q(k) \) terms.

Denote the winding time constants $T_{sd} = \frac{L_{sd}}{R_s}$ and $T_{sq} = \frac{L_{sq}}{R_s}$, the control parameters of feedback controller (PI controller) in the current loop is therefore determined as

\[
K_{p,s} = \frac{2T_{sd}}{\varepsilon_s T_{sd}^2} R_s, K_{i,s} = \frac{L_{sd}}{\varepsilon_s T_{sd}^2},
\]

\[
K_{p,q} = \frac{2T_{sq}}{\varepsilon_q T_{sq}^2} R_s, K_{i,q} = \frac{L_{sq}}{\varepsilon_q T_{sq}^2},
\]

where $\varepsilon_s, \varepsilon_q$ are positive scalars such that $\varepsilon_s T_{sd} > 0, \varepsilon_q T_{sq} > 0$.

Thus, the current controller is obtained as

\[
\begin{align*}
&u_{sd,s}(k) = u_{sd,s}(k-1) + r_{sd,s} \left[ i_{sd,s}^s(k) - i_s(k) \right] + r_{sd,s} \left[ i_{sd,s}^q(k) - i_q(k-1) \right], \\
&u_{sq,s}(k) = u_{sq,s}(k-1) + r_{sq,s} \left[ i_{sq,s}^s(k) - i_s(k) \right] + r_{sq,s} \left[ i_{sq,s}^q(k) - i_q(k-1) \right],
\end{align*}
\]

where

\[
\begin{align*}
r_{sd,s} &= K_{p,s} + T_{sd} \frac{K_{i,s}}{2}; r_{sd,s} &= K_{p,s} + T_{sd} \frac{K_{i,s}}{2}, \\
r_{sq,s} &= T_{sq} \frac{K_{p,q}}{2} - K_{p,q} + r_{sq,s} &= T_{sq} \frac{K_{i,q}}{2} - K_{i,q}.
\end{align*}
\]

and $T_{si}$ is the sampling time for current loop.

**2.2.2. Speed controller**

The motion equation of PMSM is considered as

\[
J \frac{d\omega}{dt} = \frac{3}{2} z_p \left[ \psi_p + i_s \left( L_{sd} - L_{sq} \right) \right] i_q - m_w. \tag{12}
\]

Similarly, by adopting the discretization approximation (5), the discrete-time speed model of the PMSM is obtained as follows

\[
\begin{align*}
J &\frac{1}{2T_s} \left[ 3\omega(k) - 4\omega(k-1) + \omega(k-2) \right] = \\
&= \frac{3}{2} z_p \left[ \psi_p + i_s \left( L_{sd} - L_{sq} \right) \right] i_q(k) - m_w. \tag{13}
\end{align*}
\]

Based on the property of differential flatness, the controller of the current feedforward block is obtained as

\[
\dot{i}_{sd,s}(k) = \frac{J}{2T_s} \left[ 3\omega(k) - 4\omega(k-1) + \omega(k-2) \right] + m_w,
\]

\[
\dot{\tilde{i}}_s = \frac{3}{2} z_p \left[ \psi_p + i_s \left( L_{sd} - L_{sq} \right) \right] i_q(k) - m_w, \tag{14}
\]

and a feedback controller (PI controller)

\[
\dot{\tilde{i}}_s = \dot{i}_{sd,s}(k) + \dot{i}_{sq,s}(k) \tag{15}
\]

Finally, the speed controller can be obtained as

\[
\dot{\tilde{i}}_s(k) = \dot{i}_{sd,s}(k) + \dot{i}_{sq,s}(k) \tag{16}
\]

\[
\dot{i}_s(k) = \dot{i}_{sd,s}(k) + \dot{i}_{sq,s}(k) \tag{17}
\]

where $r_w = K_{p,s} - T_{sd} \frac{K_{i,s}}{2}$ and $K_{p,s}, K_{i,s}$ are respectively the proportional gain and integral gain of PI controller in the speed loop.

### 3 Simulation results

In this work, to verify the effectiveness of the proposed control strategy, we consider a standard PMSM with the following parameters as shown in Table 1.

**Table 1. PMSM parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rated output power $P_{\text{rated}}$</td>
<td>0.4kW</td>
</tr>
<tr>
<td>Rated voltage $U_{\text{rated}}$</td>
<td>220V</td>
</tr>
<tr>
<td>Rated current</td>
<td>2.7A</td>
</tr>
<tr>
<td>Instantaneous peak current</td>
<td>8.1A</td>
</tr>
<tr>
<td>Rated torque</td>
<td>1.27Nm</td>
</tr>
<tr>
<td>Number of poles $p_p$</td>
<td>8</td>
</tr>
</tbody>
</table>
Rated rotated speed 3000rpm
Stator resistance $R_s$ 2.35Ω
Equivalent inertia $J$ 0.000031kgm$^2$
Stator inductance $L_s$ 0.0065H

Several scenarios were considered for simulation to assess the capability of the proposed controller. Here, the speed of PMSM will be controlled with a constant load and this load will be involved after 0.5sec.

**Case 1:**

The reference speed is set at 157.1 (rad/s)-forward speed and -157.1 (rad/s)-reverse speed and the motor will reverse rotation at 0.3s. The constant load involves at 0.5sec.

Simulation results are shown in Fig 3-8.

**Case 2:**

The reference speed is set from 157.1 (rad/s) to 314.2 (rad/s) and the motor will reverse rotation at 0.7s. The constant load also involves at 0.5sec.

Simulation results are given in Fig 9-14.

It can be seen that the motor speed tracked the desired speed after 0.12 sec. The tracking error curves for both cases are depicted in Fig 2, 4, 6, 8, 10, 12, 14. It is apparent that the tracking performance of the proposed method in this study is good when the static errors including the speed and current errors converge into zero after 0.01 sec. It is also observed that the tracking errors are still within an acceptable level even when reversing the rotation and starting-up with a constant load.
4 Conclusion

In this paper, the problem of nonlinear flatness-based controller design of PMSM has been addressed. Based on Taylor series expansion and by using the property of differential flatness, two controllers of the current and speed loops are proposed to eliminate the static errors. Simulation results are provided to illustrate the feasibility of the proposed approach.

References

Abstract -

Few robots are found on the worksite today and the construction process has remained unchanged since pre-industrial times, the only major advancement being the replacement of human and animal power with motors, despite research efforts in construction automation. Contrarily, the planning stages of projects have benefited from advances in simulation and modeling technologies. We propose a mechanism for leveraging information rich models to automate construction operations by using Discrete Event Simulation (DES) models of operations to drive the process in the real world using autonomous robots. This endeavor requires a new method to process DES models as they are currently used for the fundamentally different task of analyzing operations. Technically, this method is necessitated by the fact that activity durations are not known a priori when the DES model is controlling a real world operation. We put forth a novel algorithm for processing DES that enables construction automation. Our lack of prior knowledge of activity durations are accounted for by assuming that initiated activities continue until pre-empted to a stop by communication from the autonomous robot performing it. The primary contribution of our presented research is enabling the unprecedented use of DES models for construction automation. This approach allows for the automation of any construction operation, given the application breadth of DES. This modeling approach also enables new construction visualization techniques with the use of robot simulators that free up the operation modeler from consideration of low level equipment and geometric details and from the collection of activity duration data.

Keywords -

Automation and Control, Simulation, Robots, Construction Operations

1 Introduction

The level of automation in the construction industry stands in stark contrast to other engineering industries wherein most tasks are performed by machines with minimal oversight from human operatives. Several papers [1, 2, 3] have recognized the need to automate construction worksites in order to counter the threats of an aging workforce, unsafe conditions, declining/stagnating productivity, and huge cost and time overruns on projects. Tiecholz [4], in his comparison of productivities across industries, concluded that the construction industry labor productivity index is in decline, as opposed to an annual increase of the same in other non-farm industries. All of the above problems elicit the proposition of an overhaul of the construction process as it is being performed today by introducing the notion of automating construction operations as opposed to construction tasks.

A fully automated construction site remains a dream today in spite of several research attempts to introduce automation into construction. Most of these previous attempts have been inspired by the success of automation in the manufacturing industry and thereby adapting concepts from manufacturing into the construction context, such as Computer Integrated Manufacturing (CIM) to Computer Integrated Construction (CIC), etc. However, a judicious juxtaposition of construction and manufacturing reveals the following key differences that have proven critical in the failure of automation to catch on in construction: Every construction project is unique, unlike the engineered products in manufacturing. Construction generally happens in an uncontrolled outdoor environment as opposed to a controlled factory setting,
which renders the planning for construction projects to be fraught with uncertainty and makes the worksite itself difficult to control. The fact that construction is usually spread over a great area involving the movement of large quantities of materials, equipment, and labor resources requires the oftentimes remote collaboration of multiple disparate resource entities to get work done, which is different from an assembly line setting wherein the raw materials are engineered in turn by a series of local resources (machines) until completion.

It is with these issues in mind that we developed a framework for construction automation that uses the power and flexibility of Discrete Event Simulation (DES) models to orchestrate autonomous construction equipment on worksites to perform construction operations, as shown in Figure 1. Our method is inspired by the fact that DES has been established to be the most effective tool for modeling and analyzing construction operations [5]. It can be seen in the figure that the solid arrows indicate commands issued from the model to the robots and the dashed arrows represent the feedback from the robots to the DES-based command center. This framework facilitates the communication between robots and the DES model, which allows for the timely advancement of the operation based on the model’s operational logic dependent on the communication received from the robots.

Figure 1. Conceptual design of DES orchestrated worksite automation

However, DES models have hitherto been used only for the analysis and design of operations and never as a control system for construction operations. Thus, a new mode of processing DES models is required in order for them to be able to control an operation in the real world. The technical reason required for this modification is the fact that DES models for analysis require the duration of activities to be input into the system a-priori, i.e. before the activity occurs in the simulation. However, it is impossible to know the duration of any construction activity in the real world until after it has happened due to the prevalent uncertainty in worksites. In this paper, we present a novel methodology that allows for the use of DES models as control mechanisms to enable the automation of complete construction operations. The key challenge addressed by the presented framework is obviating the need for a-priori duration data. We anticipate our methodology to be a prime enabler for our overarching goal of realizing automated construction sites (or major operations therein) using the framework illustrated in Figure 1 that marries the power and flexibility of DES modeling with the burgeoning sophistication of autonomous equipment. We also believe that the envisioned DES processing algorithm will enable novel methods for the simulation and visualization of construction operations by alleviating the modeler from concern over low level equipment and geometric details of the worksite, thereby allowing for faster model development and focus on operation level design.

An overview of construction automation attempts is provided in this paper to highlight its challenging nature. The use of DES in the construction context is also discussed to underscore the novelty of our approach, which is then described along with its applications. The final section of this paper contains a delineation of expected research contributions and plan for future work.

2 Background

This section details the research efforts that have aimed to introduce automation into the construction worksite. A review of the use and state of the art of DES models in analyzing construction operations is also provided to set the context within which our methodology is presented.

2.1 Automation in Construction

It has been established that the application of automation technology and information technology is essential to the continued growth and survival of the construction industry. A review of existing literature on
the subject reveals how construction automation and robotics is seen to be the answer to numerous problems that the industry is facing. Two dominant themes that have characterized prior research efforts are discussed here to provide an overview of the state of automation efforts in construction: 1) The idea that concepts that were successful in implementing automation in the manufacturing sector could be translated into the construction industry; 2) The notable shift of research focus from robotics and automation to “soft robotics” that encompasses software integration, simulation and VR, sensory based monitoring and tracking, etc.

2.1.1 Manufacturing to Construction

In a study on the potential for automation and robotics in construction, Bernhold [1] concluded that even though there are a lot of institutional barriers related to the fragmented and unstructured nature of the construction industry, there exists the prospect for restructuring the approach to construction during both the design and construction phases to allow for the easier diffusion of automation technologies into construction. Bernhold referred his readers to Halpin and Woodhead’s [6] organization of construction management into a 6-tier hierarchy consisting of the levels of organization, project, activity, operation, process and work task; the last three levels of which are amenable to the concepts of repetition and standardization that led to automation and mass production in manufacturing.

Sanvido and Medeiros [7] performed a comparison of the two industries to identify potential areas for the cross-fertilization of concepts and concluded that the similarities between the two industries outweigh the differences and thus defined a strategy for implementation of successful integrative manufacturing tools and concepts, such as Computer Integrated Manufacturing (CIM), in construction as Computer Integrated Construction (CIC). Miyatake and Kangari [8] defined CIC as “a strategy, incorporating computers and robotics, for linking existing technology and people in order to optimize business activity.” They noted that there is no standard formula for CIM and that each company must formulate its own system for implementing CIC, but they did mention that CIM usually results from an integrated information flow, the widespread application of computers, and high levels of automation. They described a prototype construction system called the SMART system (Shimizu Manufacturing system by Advanced Robotics Technology) in Japan that implemented CIC by focusing on three major areas: 1) integrated design/construction planning, 2) site-automation system, and 3) factory automation; which when applied to an office building project in Japan in 1994 was purported to have reduced labor by a total of 30%, although it has not been repeated since. Possible explanations for this attempt at CIC not being repeated are its high implementation cost, the fact that the technologies developed were only applicable to a narrow class of high rise building projects, and that the system was not fully automated.

A more recent example of applying a manufacturing concept into construction is the introduction of 3D printing on construction sites, of which one of the foremost examples is the Contour Crafting method [9] that uses additive manufacturing methods to print houses. While it greatly reduces the cost and time to build dwellings using in-situ resources without compromising on the flexibility of design like traditional manufactured houses, it is not generic enough to be applied to a class of construction activities other than that of raising the shells of buildings. All these manufacturing-inspired research efforts, which have not yet solved the problem of automation in construction, echo the conclusions of Everett and Slocum [2] that construction needs to develop its own strategies for automation rather than copy them from manufacturing as they are fundamentally different in the relationships among product design, process design and fabrication.

2.1.2 Current State of the Art and Trends in Construction Automation

In spite of the development of several prototype construction robots [3] [10] and the introduction of CIC etc., the world has yet to witness a huge change in the way the construction industry operates. Warsawzki and Navon [3] lamented the failure of the diffusion of robotics in the 1990s in spite of early promising research efforts and cited the following reasons to be its cause: current building and construction practices not being amenable to performance by robots; insufficient development of construction robots to deal with site conditions; the lack of economic justification of robots; and the managerial environment in the construction industry. This sentiment was shared by Balagué [11]
who concluded that the industry still operates by the same philosophy as from the pre-industrial era, the only major advancements being the use of motors to apply force and the replacement of human and animal power with machines to do the same work while manual control, visual feedback, human operators, etc. still play a vital role in construction. Balaguer also noted that most of the successes of CIC were limited to the implementation of IT practices in construction in the design stages and not in the production and construction stages.

This general trend towards the softer side of robotics and automation is reflected in the findings of Son et al. [12] who attempted to describe the global trends in research and development of automation and robotics in the construction industry by analyzing International Symposium on Automation and Robotics in Construction (ISARC) papers of the last 2 decades. It was found that the percentage of papers categorized under “Construction Robotics” declined from 71% in 1990 to 33% in 2008, which stands in contrast to all of the other categories, which seemed to attract a larger research focus during the same period.

It is with due regard to the trends described above that we present our methodology. Specifically, we ensure that our proposed framework is developed for construction with due consideration for the uncertainty, diversity, and complexity of its operations. By using DES, our method exploits the sophistication achieved in modeling and simulation tools as a result of the focus of recent research. The next section describes the application of DES in construction and the current DES processing algorithm, which would serve to provide the background for our framework.

2.2 Discrete Event Simulation in Construction

Personnel involved in the design of construction operations are expected to make decisions about the complex processes which may involve several decision alternatives, each of which is associated with its own expected outcome, which can be estimated using either real world experimentation, mathematical modeling, or simulation, which involves imitating the system under study. Martinez [13] noted that of the three techniques mentioned above, simulation is the most convenient, as it is both realistic (as opposed to mathematical modeling) and inexpensive, fast and flexible (versus real world experimentation). Discrete Event Simulation (DES), a type of simulation wherein the state of the system changes at discrete points of time marked by events per the DES model, is very effective at modeling construction processes [13].

2.2.1 DES for Analysis of Construction Operations

Applications for the simulation of systems follow two main simulation strategies: process interaction (PI) and activity scanning (AS). Event scanning (ES), a third approach is often used in conjunction with either PI or AS. Martinez and Ioannou [5], in their comparison of the two dominant strategies, establish that AS models, built from the perspective of activities in the operation, are more suited to modeling construction operations than PI models which are created from the perspective of the resources that flow through the system. The authors note that PI modeling is more suited to systems wherein resources that flow through the system have many differentiating attributes as opposed to stationary entities that serve them; a situation characteristic of manufacturing and industrial systems. On the other hand, construction operations involve heavy interaction between machines that can have many attributes and be in several states, which justifies the focus on activities rather than on differentiating resources as either flowing or stationary.

Three-Phase AS is a hybrid strategy that optimizes the performance of AS by combining it with ES. Three-Phase AS is represented at the conceptual level by a network of activities and queues that make up the operation called Activity Cycle Diagrams (ACDs). Figure 2 shows the method for processing Three-Phase AS models.

Figure 2: Traditional Discrete Event Simulation Loop for Construction Simulation [13]
In Figure 2, the Future Events List (FEL) refers to a list of activity instances that have already been started, that are arranged in ascending order of their end times. In the Clock Advance Phase, the clock is advanced to the time of the activity that is ending the soonest, before that activity can be harvested (or ended) in the FEL. Thus it can be seen that the durations of activities must be known before the activity ends, which is an impossibility in real world operations.

2.2.2 DES Enabled Operation Visualization

Construction visualization at the operations level displays DES-generated animations of the resources on a work site as they build a product or perform a support service [14] and was borne of the need to verify and validate DES models. While initial implementations including VITASCOPE [15] completely separated the generating process (eg. DES model) from the visualizer, a later advancement in the art of visualizing construction so as to enable user interactivity with the underlying DES model, known as DES-Based-Virtual Reality (VR) [16], is of particular interest to our research goals. This interest is due to the fact that the interactivity could only be achieved if the simulation ran concurrently with the visualization, and allowed for unplanned changes in the underlying DES model due to user interaction.

While a complete description of DES-based-VR is beyond the scope of this paper, Figure 3 depicts the conceptual time advance algorithm that allowed for the synergistic combination of DES and animation to enable user interactivity.

![Figure 3: Conceptual Time Advance Algorithm for DES-Based-VR [16]](image)

In Figure 3, it can be seen that while the duration of activities is still required to be known a-priori in order for the regular processing of the simulation model, there is an element of uncertainty in the user interactions that are unforeseeable. Rekapalli [16], in his implementation of the incorporation of user interaction affected the model by preempting the relevant activity to a stop. As will be seen in Section 3, we borrow of the concepts that are used in enabling concurrent simulation visualization in enabling the DES model to be used as a control mechanism of real world construction operations and in enabling a novel paradigm of construction visualization using robot simulators.

3 Proposed Methodology

This section provides a conceptual overview and plan of implementation for the modified DES processing algorithm to be used as a control mechanism for the automation of construction operations. The proposed methodology is explained by revisiting the traditional method for processing Three-Phase AS models. In addition to the goal of construction automation, we also describe the plan for implementation to enable a new mode of construction simulation and visualization.

The primary assumption made in the development of our methodology is the availability of entities that are capable of communicating with the DES model. These “entities”, that could be robots in the real world or virtual objects, are assumed to be able to receive commands from the DES models to start doing a specific pre-programmed activity and to be able to communicate their status back to the DES model when the activity assigned is completed.

3.1 Conceptual Overview

In order to use DES models as control mechanisms to orchestrate autonomous robots to automate construction operations, it is necessary to account for the lack of a-priori information regarding the activity durations, which is a primary input for DES models. For a conventional DES, each activity’s duration is required to be known to ascertain end times of each activity instance, which determines its place in the Future Events List (FEL), subsequently driving the advancement of the simulation model as shown in Figure 2. In order to compensate for the lack of a-priori knowledge of activity durations in our control scheme, we propose the modified DES processing algorithm which is presented in Figure 4.
Both of the implementations described use STROBOSCOPE [14], an extensible general purpose simulation system specifically designed to model construction operations. The attributes of STROBOSCOPE that make it particularly suited to our goal are: its consideration of uncertainty in any aspect (not just time); its combination and allocation of resources and control of the activation tasks by subjecting them to complex logical conditions; and its usefulness to model any construction operation by virtue of being a general purpose simulation system [14]. The extensibility of the capabilities of STROBOSCOPE through the use of add-ons makes it particularly amenable to our efforts at implementing the presented methodologies.

### 3.2 Methodology Implementation

We present the following two strategies for implementing the modified DES algorithm to enable construction automation and visualization respectively.

#### 3.2.1 Construction Automation

As can be observed from Figure 4, a Current Activities List (CAL) replaces the FEL from the traditional implementation shown in Figure 2. The CAL contains an unordered list of all activities that have been started and are being performed by robots on the construction site. The activity items in the CAL contain only their starting time and not their end time. The clock advance will thus happen only when any of the robots on the work site complete their assigned activity, upon which the clock is advanced to the current time, and the corresponding activity from the CAL is harvested and forced to a stop in the DES model. Subsequent activities may be triggered by the end of one activity as per the model logic, which in turn may trigger robots on the worksite to perform more activities, until the CAL is empty.

We believe that this approach exploits the power and flexibility of existing DES modeling tools for the control of multiple robots on the work site to perform complex operations. The algorithm modification described above to account for the fact that no a-priori information is available about activity durations allows for the incorporation of duration uncertainty in our approach.

#### 3.2.2 Construction Visualization

As stated above, construction visualization was borne of the need to verify and validate DES models. Animating simulated construction operations suffers from the following limitations that places additional burden on the operation modeler: Animation instructions from the DES model need to be communicated at a very high level of detail that may not be relevant to the purpose of the model from a decision making perspective; operation modelers need to concern themselves with geometric detail, mostly irrelevant in DES analysis, to create spatially correct and convincing visualizations [17]. The collection of duration data to input to DES models is also a very time consuming and
sometimes impossible process, for activities that have not yet been performed in the real world that need to be simulated. The paradigm for simulating and visualizing operations described in this section is necessitated by the fact that conventional visualization techniques for verifying and validating DES models are not applicable to the modified DES processing algorithm, and this paradigm has the potential to overcome the limitations of conventional visualization systems described above.

We propose to replace the use of traditional virtual CAD models in 3d operations visualization with models in robot simulators [18, 19], a measure that allows for the encapsulation of all of the low level information regarding the equipment’s performance within its virtual model. Our methodology for visualization and simulation uses a hybrid approach wherein the modified DES control model of the operation is used to direct virtual equipment in the robot simulator (as per Figure 1). The virtual equipment would perform the tasks required in the operation according to their programed and encapsulated functionality, thus freeing up the modeler from animation modeling to focus on operation level modeling. Sensing capabilities on the virtual robot would allow it to react to its virtual environment appropriately, thereby obviating the need for the entry of any spatial and geometric information into the DES model. The encapsulated performance data of the robot would allow it to perform tasks without the need for accompanying duration data, which would exempt the modeler from having to collect that data from real world operations. While this methodology is generic enough to be applied to any simulation engine and robot simulator packages that are designed to be extensible, our implementation of the framework is built specifically for the use of STROBOSCOPE [13] and the Virtual Robot Experimental Platform (V-REP) [18]. The methodology provides a more accessible means to create construction visualizations, and at the same time, constitutes a means to develop and test the modified algorithm for DES-based control of construction operations as we propose.

4 Contributions and Future Work

The algorithm described herein for the processing of DES models to enable their use as control mechanisms enables a fresh attempt at the automation of construction operations. We contend that our methodology allows for the unprecedented use of the power and flexibility of information rich DES models to automate any construction operation. The primary challenge that was tackled here in enabling the transition of a DES model from analysis and planning to control mechanism was obviating the need for a-priori information regarding the performance of activities, namely durations. Instead, the resulting concurrency of the DES model and the real world operation are enabled by the use of communications from autonomous robots to advance the state of the DES model instead.

We believe that this approach to automation pays due consideration to the dominant trends that have characterized research efforts in automating construction: 1) by developing a methodology that considers the challenges of unpredictability, specificity of projects, and complex processes, that are characteristic and inherent to the construction industry; and 2) by leveraging the sophistication of modeling techniques that benefitted from research focus in the construction context. The presented approach is generic enough to be applied to a wide class of processes that can be suitably modeled using DES. The automation of entire operations also eludes the problem of islands of automation that has been characteristic of past efforts of task automation.

We also contend that the use of the modified DES algorithm in conjunction with robot simulators for the validation and verification purposes of the control mechanism presents a new approach to the simulation and visualization of construction processes that has numerous benefits over traditional methods which require the durations of activities as input. Instead, the use of encapsulated performance data of equipment (easier to acquire from manufacturers) in their corresponding models in the robot simulator relieves the operation modeler from 1) the intensive time and effort and sometimes impossible task of collecting duration data for activities and 2) concern over geometric data that is otherwise irrelevant to the purposes of the DES model from the decision making perspective, but essential to creating convincing and spatially correct visualizations of the operation. The encapsulation of data and functionality within the robot simulator model also permits the operation modeler to communicate animation instructions to the visualizer at the activity level instead of the elemental motion level.
While we have developed the implementation strategy for both the automation and simulation-visualization of construction operations using the modified DES models, we are currently in the process of developing a protocol to describe the capabilities of the robot/robot model to the operation modeler. This would enable the separation of operational level and low level equipment details and thus alleviate the responsibility on the concern of the operation modeler over aspects that are irrelevant to the simulation. We have completed the development of an add-on to STROBOSCOPE that implements the modification to the DES processing algorithm as described in this paper. We are working on developing robot models, both in the real world and in V-REP, to test the overall methodology for the automation and visualization respectively of typical construction operations. A future step in the development of the presented methodology is to utilize the feature of STROBOSCOPE, which allows for access into the state of the model and its resources at any point, to access highly contextual real time information about the operation being performed. This advancement would enable automated and real time monitoring and control of the already automated operation.

5 References


[16] Rekapalli, P. V. Discrete-event simulation based virtual reality environments for construction operations, Purdue University, 2008.


Performance Test for Rapid Surface Modeling of Dynamic Construction Equipment from Laser Scanner Data

C. Wang\textsuperscript{a} and Y.K. Cho\textsuperscript{b}

\textsuperscript{a}Ph.D. Candidate, School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason Building 3136, 790 Atlantic Dr., Atlanta, GA 30332-0355
\textsuperscript{b}Associate Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Mason Building 4140A, 790 Atlantic Dr., Atlanta, GA 30332-0355

E-mail: cwang2@gatech.edu, yong.cho@ce.gatech.edu

Abstract -
In modern dynamic construction fields, more attention has been paid on safely operating heavy construction equipment such as cranes, excavators, and concrete pump. In order to improve the safety in complicated jobsites, it is highly required to provide heavy equipment operators with accurate measurement of on-site objects in near real time. In this paper, a rapid surface modeling method and its performance evaluation through on-site tests are introduced. The performance of the proposed method was tested with a heavy equipment at construction site. The interrelationships among data size, processing time, and the size of the resultant hull segments were examined from the data analyses. The field experimental results demonstrate that the proposed dynamic surface modeling method would significantly improve the equipment operation productivity and safety by distinguishing a dynamic surface model being controlled by the operator from the point cloud of existing static environment in 3D views.

Keywords -
Construction equipment; Point cloud; Object recognition; 3D modeling; Safety

1 Introduction
Visibility-related accidents can be easily caused by the interactions between workers, equipment, and materials. This problem can lead to serious collisions without pro-active warnings. There have been a number of advances in vision-aid techniques because lacking full visibility is a major contributing factor in accidents at construction sites. 3D spatial modeling can help to optimize equipment control [1, 24], significantly improve safety [2-3], monitor construction progress [4], and enhance a remote operator’s spatial perception of the workspace [5-8]. However, the rapid processing of tens of thousand bits of range data in real time is still an unsolved problem requiring further investigation [9]. Unstructured work areas like construction sites are difficult to graphically visualize because they highly involve unpredictable activities and change rapidly. Construction site operations require real-time or near real-time information about the surrounding work environment, which further complicates graphical modeling and updating.

One commonly used method to obtain the 3D position of an object is based on 3D laser scanning technology [7, 10-11]; this method, however, has some limitations, such as low data collection speed and low object recognition rates [12]. It has always been a challenge to recognize specific objects from a 3D point cloud in unstructured construction environments because it is difficult to rapidly extract the target area from background scattered noises in a large and complex 3D point cloud.

While rapid workspace modeling is essential to effectively control construction equipment [13], few approaches have been accepted by the construction industry due to the difficulty of addressing all the challenges of current construction material handling tasks with the current sensor technologies. Thus, an innovation in rapid 3D spatial information is necessary to meet the challenges. The main objective of this paper was to validate a 3D visualization framework to collect and process dynamic spatial information rapidly at a construction job site for safe and effective construction equipment operations. Multi-video camera integrated vision-based object recognition and tracking method has been developed, based on which, a smart laser scanning method was proposed to reduce data size and scanning time.
Fig. 1. Model-based object recognition and registration [17]

2 Related Work

For the operator to monitor blind spots of the workspace from the cab, a vision-based system using a single or multiple cameras is an inexpensive option [14]. Brilakis et al. [15] introduced 2D vision-based methods that recognize new overlapping feature points and track them in the subsequent video stream. To acquire a precise 3D position of objects with additional depth information, generally two or more cameras generate a stereo view after calibration with known intrinsic parameters. Park et al. [16] achieved more accurate 3D locations of tracking objects by projecting the centroids of the tracked entities from two cameras to 3D coordinates. Notwithstanding the recent advances, there are some known drawbacks of vision-based techniques in tracking moving equipment at the sites: 1) additional infrastructure is needed to install and maintain cameras; 2) fixed camera locations have limited view angles and resolutions, and 3) the results are sensitive to lighting conditions [17].

Laser scanners have been extensively utilized to automatically obtain the “as-is” condition of the existing buildings [18]; they also can be used to classify and capture a complex heavy equipment operation as it happens or to provide automated feedback to those who are conducting the operations [7, 17, 19]. Teizer et al. presented a methodology for real-time 3D modeling using Flash LADAR which has a limited measurement range and low accuracy for outdoor use [3]. Lee et al. proposed an automated lifting-path tracking system on a tower crane to receive and record data from a laser device [13]. Bosch and Hass registered 3D static CAD objects to laser-scanned point cloud data [20], which can be utilized to efficiently assess construction processes. However, most of the algorithms were developed mainly to recognize and register static objects’ models to point clouds. Few applications have demonstrated the technical feasibility of registering dynamic models to point clouds in real or near real time. In the authors’ previous studies [17], a model-based automatic object recognition and registration method, Project-Recognize-Project (PRP), was introduced to register the CAD models with the corresponding point cloud of the recognized objects through comparing the recognized point cloud of the objects with existing CAD models in a database (shown in Fig. 1.). While the PRP approach provides very detailed, accurate solid models in a point cloud, the limitation of this method is that it only works for the objects which have corresponding models in the database. In this study, a non-model based, surface modeling method is introduced to automatically recognize and visualize dynamic objects on construction sites.

3 Overview of the Proposed Method

In Fig. 2, the framework of the proposed rapid surface modeling method is illustrated. The developed data acquisition system is composed of two 2D line laser scanners (80 meter working ranges at 100Hz scan speed, up to 2.5 sec / 360° scan, 190° for vertical line), a digital camera and three video cameras with a resolution of 0.25 degree in a vertical direction and 0.0072 degree in a horizontal direction. In this system, multiple
degree-of-freedom (DOF) kinematic problems were solved based on the mechanical installation, and 3D point cloud data and digital image streams can be collected simultaneously. Utilizing this form of flexible design together with Time-of-flight (TOF) laser scanner working type, higher scanning resolution and faster scanning rate were obtained, which is more suitable for the complicated dynamic construction environments.

Working together with the laser scanner system, the digital video cameras were employed to capture real time image streams in the jobsite for the operators. Then, the operators can select the target objects in the image streams by drawing a bounding box through the developed graphical user interface. Using a robust local feature detector, Speeded Up Robust Features (SURF) [21], image-based target object recognition and tracking algorithms were implemented.

Taking the bounding box area as an input, smart scanning and visualization processes were immediately applied to separately collect and update the data of the target object and the static site environment. The major two steps of the smart scanning method are: (1) a static jobsite environment is scanned by the laser system with a very high resolution. Those collected site data are stored in the memory; and (2) from the second scanning round, only the point clouds in the dynamic target area specified by the equipment operator are updated separately. In this process, the size of point cloud data obtained from complex large construction sites and the updating time can be significantly reduced.

Surface modeling of point clouds can be more advantageous for the equipment operators over the point-cloud-only visualization because a surface model of target object can be better distinguished from a complex point-cloud environment. The concave hull approach is one of the most widely used surface modeling methods creating a polygon that represents the area occupied by a set of points. A concave hull better details the shape of the point cloud than the convex hull does. A concave hull of a set of surfaces is the enclosing concave surfaces with smallest volume. There are several existing concave hull calculation algorithms [22], however the efficiency of these algorithms decreases significantly because of high computing queries for the large size of point cloud data like the ones obtained from construction jobsites.

The raw data size of point cloud data collected from a construction jobsite is quite a large and as a result, it is challenging to process and visualize these data in real or near real time. In this study, as shown in Fig. 4, the smart scanning technology was employed at the very beginning of the surface modeling phase in order to significantly reduce the data size of the point clouds and the surface modeling time. In addition, outliers of the point data were statistically identified and removed to improve concave hull model accuracy. Then, a data filtering and downsizing process was conducted to further decrease the number of points. The goal of data downsizing is to increase the data processing speed by reducing the amount of overly dense data being processed.

4 Field Test and Discussion

Validation of the proposed methodology was implemented on a backhoe loader at a building construction site. The data acquisition system was mounted on a mobile cart, and set up in the working area of the equipment, especially in its blind spots. It firstly scanned the whole jobsite and kept the point cloud data in the database; then smart scanning and visualization were fulfilled based on the selected tracking objects. All tests were benchmarked on an Intel Core i5 CPU with 4GB RAM on a 64 bit Windows
mobile computer. It should be noted that the required resolution, registration accuracy, and scan rate for successful surface modeling vary based on the scan range, ambient lighting conditions, properties of the target (e.g., shape, color, reflectivity, and moving speed), and the number of mounted 2D laser scanner. In this study, all data were collected from the system with two 2D laser scanners, and the maximum scan speed of which is 227,500 points/sec. The scan speed could be doubled if four 2D laser scanners were equipped [17].

In this study, a backhoe loader was chosen as a test subject. The entire equipment was selected as the tracking target, as shown in Fig. 3. Through smart scanning, the point cloud of the target area and the site environment were separately collected and visualized in Fig. 4., and 18,304 points were collected from the target. Then the collected raw data was processed by the data filtering and downsizing algorithms to decrease the surface modeling time. The resolution (the average point to point distance) of the raw data is 0.01 m, and the raw data size became 17,585 after filtering, and varied from 61 to 17,311 after downsizing based on different leaf sizes (Fig. 5. and Table 1). The leaf size here is defined as the edge length of the 3D voxel, and the greater the leaf size is, the more the data size is decreased. Table 2 shows the processing time under different data downsizing scales and the size of the concave hull segments. In Table 2, α limits the size of the resultant concave hull segments. The smaller α is, the more detailed the hull segments are. The concave hull surface modeling on raw data without data filtering and downsizing was also evaluated. As shown in Table 3, the range of surface modeling time for raw data was between 1 to 4 seconds. Thus, data filtering and downsizing was needed before surface modeling to reduce the data size for a real-time application. In Fig. 6, the total processing time includes time for data filtering, downsizing, and surface modeling. When the leaf size was 0.01 m, the total processing time was greater than 1 second, but in the other sizes the time was all within 1 second. To deliver a best available result to the operator, therefore, the smallest α value was selected. Fig. 7 shows the surface modeling result of the backhoe holder integrated with the data of the static site environment.

Table 1. Data size of each step of the proposed methodology

<table>
<thead>
<tr>
<th>Step</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw data</td>
<td>18,304</td>
</tr>
<tr>
<td>Filtered data</td>
<td>17,585</td>
</tr>
<tr>
<td>Downsized data</td>
<td></td>
</tr>
<tr>
<td>Leaf size = 50 * Resolution</td>
<td>61</td>
</tr>
<tr>
<td>Leaf size = 10 * Resolution</td>
<td>1,014</td>
</tr>
<tr>
<td>Leaf size = 5 * Resolution</td>
<td>3,036</td>
</tr>
<tr>
<td>Leaf size = Resolution</td>
<td>17,311</td>
</tr>
</tbody>
</table>

*Resolution = 0.01 m in the case 1.
Fig. 5. The point cloud of (a) the filtered data, (b) the downsized data, and (c) the created surface model.

Fig. 6. Total processing time with different $\alpha$ value.
Table 2. Processing time of each step of the proposed methodology

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>Data Downsizing Leaf Size (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>Data Filtering</td>
<td>0.133</td>
</tr>
<tr>
<td>Data Downsizing</td>
<td>0.003</td>
</tr>
<tr>
<td>Surface Modeling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α = 1.0 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 1.5 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 2.0 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 2.5 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 3.0 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 3.5 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 4.0 * Leaf Size</td>
</tr>
<tr>
<td></td>
<td>α = 4.5 * Leaf Size</td>
</tr>
</tbody>
</table>

Table 3. Test on raw data without filtering and downsizing

<table>
<thead>
<tr>
<th>α</th>
<th>Surface modeling time (s)</th>
<th>α</th>
<th>Surface modeling time (s)</th>
<th>α</th>
<th>Surface modeling time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>1.003</td>
<td>0.08</td>
<td>0.948</td>
<td>0.6</td>
<td>1.853</td>
</tr>
<tr>
<td>0.02</td>
<td>0.961</td>
<td>0.09</td>
<td>0.947</td>
<td>0.7</td>
<td>0.924</td>
</tr>
<tr>
<td>0.03</td>
<td>0.965</td>
<td>0.1</td>
<td>0.948</td>
<td>0.8</td>
<td>1.854</td>
</tr>
<tr>
<td>0.04</td>
<td>0.963</td>
<td>0.2</td>
<td>1.888</td>
<td>0.9</td>
<td>2.782</td>
</tr>
<tr>
<td>0.05</td>
<td>0.967</td>
<td>0.3</td>
<td>2.822</td>
<td>1.0</td>
<td>0.925</td>
</tr>
<tr>
<td>0.06</td>
<td>0.962</td>
<td>0.4</td>
<td>3.752</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>0.954</td>
<td>0.5</td>
<td>0.924</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7. The created surface model was integrated with the static site environment
5 Conclusion

In this study, a specially designed data capturing system was utilized to collect point cloud data from multiple laser scanners; while multiple video camera arrays were used to rapidly recognize and track the selected dynamic construction equipment objects including a backhoe loader and a crane. The validation of the proposed method was implemented at a real world construction jobsite. The concave hull of the crawler crane was generated in less than 0.5 seconds, and then the data were smoothly transferred to the operator in a cabin. The test results indicate that the proposed rapid workspace modeling approach can improve the heavy equipment operations by distinguishing surface-modeled dynamic target objects from the point cloud of existing static environment in 3D views in near real time.

While the surface modeling time is sufficiently fast enough for real-time operation, the data collection time should be carefully configured based on the types of equipment (e.g., size and moving speed), distance, ambient lighting, and relectivity.

For future work, the research will continue to improve the resolution of laser scanner data while reducing data collection time. With an increase in scanning speed, the scanned resolution is lowered accordingly. To resolve this issue, a smart scanning approach with differentiated scan speeds will be further developed, to allow faster rotations for the areas to be skipped, and slow the scan speed for the target areas. Improving surface modeling speed using a higher performance computer and surface model quality will be another future research focus.

Acknowledgement

This material is based on work supported by the National Science Foundation (Award #: CMMI-1055788). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References


Abstract -

The paper deals with the use of biological life support systems, initially created for long-duration space flights [1, 2], for the solution of problems with the sustainability of metropolitan cities, which must meet the requirements of decrease of environmental load. Motivated by applications in issues of design of water purification stations for new generation of high-rise buildings "Sky City" this research considers the questions of the simulation and control of artificial ecosystems. As mathematical formalism to describe and control the processes in a bio purification station, we use the general form of conservation laws of matter in the form of the Reaction-Convection-Diffusion Equation System [3, 4, 5] and optimisation problems for it. The main types of schemes of life support systems for space stations have been discussed. Much attention has been paid to the question of creating bioreactors, which can be used for water purification within residential building. The simulation and control of artificial ecosystems has been discussed.

Keywords -
Space Technologies, Life Support Systems, Metropolitan Cities, Mathematical Modelling of Ecosystems

1 Introduction

Growing constantly, modern megacities and urban agglomerations exert extremely high environmental load. Numerous skyscrapers that form city skylines and high density of buildings lead to disruption of the ecological balance of the city and the territory in which it is located, as well as and the self-repairing functions of city ecosystems. Also, one of the biggest problems of megalopolises is the human-waste disposal. In modern conditions, traditional sewer and wastewater treatment systems are very expensive, occupy vast territories, and do not always operate reliably. Occasionally, wastewater discharges to water bodies lead to exceeding the maximum permissible concentrations of pollutants. According to United Nations Environment Programme, the state of the environment is the direct cause of about 25% of ailments all around the world, and the most common among them are gastrointestinal and acute respiratory infections. Cause of 7% of deaths and disease is poor water quality and poor sanitation [6, 7].

Nowadays new generation building projects intended to 100-200 thousand inhabitants are being developed actively. Let us call them Sky-City-buildings. Under current technology, such amount of people makes sewer system load of the order of 0.6-1 m$^3$/c and requires about 8-10 hectares of treatment plant area. We propose a new concept of small-volume local treatment facilities, which provide almost complete recycling of substances, and can be used for water purification within Sky-City-buildings. The main features of the concept are the small volume of purification system, its local position on a floor in close proximity to inhabitants, recycle of substances and energy (closed-loop treatment facilities).

Small volume of treatment facilities is a significant advantage from point of view preservation of the wildlife areas, the possibility of refusal to use a sewer network, increasing the motivation to economy and responsibility of each water user. Further such kind of facilities can be integrated into self-sustaining, interchangeable and standardized platform, which is situated beneath a modular home, a so-called mainboard- inspired one, by the principle of computing, which houses and electronically controls all water installations and energy components needed for a household [8].

Idea of the article is the use of technology developed for space under ground conditions. No other field of science has influenced imagination and vision more than the research and outcomes related to space exploration, either it is in terms of science fact, or science fiction. Analysis of the opportunities and prospects of using various space technologies for designing the systems on the Earth that may be
described as being sustainable has been given in [9]. Background of our project is experience in creating closed ecological systems as parts of biological life support systems (LSSs) for space ship.

We have applied the idea of creating some isolated volume suitable for long-term survival of man to the Earth conditions, for example, a volume located in a skyscraper.

2 Life Support Systems for Space Stations

To date, life support systems that allow human to exist inside the hermetically closed volume for a long time have been created and exploited. Such systems fully provide human with water, oxygen and remove wastes. There are four kind of LSSs for a spacecraft (Figure 1): the system with reserves where all waste is removed, the system with reserves and partial physical-chemical regeneration of substances, the system with partial substance regeneration by living organisms (bioregeneration), and the closed-loop system.

Currently models with partial physical-chemical regeneration of substances is being used at the International Space Station. These LSSs can function only in the presence of constant communication with the biosphere of the Earth, as they are based on the stocks of materials, so the duration of human existence in them depends on the amount of reserves or the possibility of their constant replenishment. Since the beginning of the space age, it was evident that for long interplanetary spaceflight and planetary bases (settlements), when the connection with the biosphere of the Earth is completely absent, LSS must be built on a fundamentally different basis.

The basis of LSS for a long-duration flight should be the principle of regeneration of the human's environment. Long-term work on the creation of human's life-support system has showed that it is easier to implement regenerative system based on biological principle rather than on the basis of physical-chemical processes. Attempts to develop such kind of systems started in USSR in the late 50's. As a result, by the current moment, different models of LSS based on the biological cycle of substances have been created and tested in Earth conditions [10].

In 1975, Eugeniy Shepelev introduced the concept of closed ecological life support system as the hypothetical biological system that exists on the basis of a closed cycle of substances without material exchange across its borders. To date, the feasibility of constructing of closed ecological LSSs that meet all needs of the person has been proved.

Figure 1. Schemes of the main types of Life Support Systems
During the 60s - 80s, at Institute of Biomedical Problems and Institute of Biophysics of Russian Academy of Sciences, there were several experiments with systems based on the following links of the ecological chain: algae - mineralization - human and algae - mineralization - higher plants - human.

As the project being implemented currently, should be mentioned the MELiSSA-project (Micro-ecological life-support system) of the European Space Agency. In this project, scientists plan to include one more heterotrophic link – animals, namely fishes [1]. Whereas attempts to create a similar system on the basis of physical-chemical processes are still not successful. To date, there are only different elements (subsystems, components, assemblies, etc.) but there is no regenerative systems based on physical-chemical processes as a whole.

It should be noted that in spite of the successful experiments, the closed ecological life support system based on a biological cycle of substances are not used in practical astronautics. That is due to the facts that their operation requires high power consumption, as well as large areas, volumes, initial weight (especially for systems with higher plants), and there is no technology of work with living organisms in space flight conditions (in weightlessness). However, these factors are not essential in our case. Key processes of LSS are the recycling of inedible biomass, human waste (feces, and urine), carbon dioxide and minerals and the production of food, fresh water, and air revitalization, while functions of the system developed for use in skyscrapers are much narrower. We have only the following objective: recycling of human waste for fresh water production. There is no weightlessness, the need to produce food, oxygen deficiency. Thus, we have to solve the problem of high energy consumption only. However, on the Earth, it is possible to close the system through the process of oxidation (combustion) of biomass for generation energy. Another problem that is still constraining the design of similar systems is the fact that in contrast to the conditions at the space station, for an ordinary city building, it is pretty difficult to predict effluent precisely enough for creation a stable and efficient system.

3 The Project Concept

Making a start from the effective experience of closed ecological LLS creation in USSR, we design the idea of local small volume biological treatment facilities with maximum-closed cycle of substances. The purpose of our work is to develop of biological treatment facilities based on a closed-loop cycle for ultra-small waste effluent of about 30 people for using in civil engineering.

Since the main difference of the system is a high degree of closure of flows then the hardest part is to balance all the technological processes. The basic idea that allows us to overcome this difficulty is creation of an automatic control system of the treatment facilities on a base of sufficiently accurate mathematical model describing concentration substances in an artificial aquatic ecosystem of the purification system. Thus, the adequate simplicity and the possibility of describing the purification process via laws of matter conservation is one of the main requirements for the system being developed.

In all types of closed ecological LLSs, algae is a basic element (link) of LLS, which allows regenerating the air, and water completely and provides [2, 10]:

- Partial nitrogen cycle by way of the full use of human urine nitrogen by unicellular algae.
- Optimization of air ion and aerosol composition of the atmosphere.
- Stabilizing of content of water-insoluble impurities (methane, carbon monoxide, etc.) by way their adsorption on the algae cell and microorganism surface and the subsequent removal of the grown biomass from the system.
- Ousting microflora, including pathogenic to humans, from micro biocenosis by a competitive relationship.
- If applicable, the names of other authors, having different affiliations and addresses, in the same format used for the first author.
- Cleaning the atmosphere of a hermetic volume from various water-soluble gaseous contaminants through their full absorption and utilization in a photo-reactor by algae and attendant microorganisms (photo-reactor is an universal resettable hydrobiological filter).

Multifunctionality of biological regeneration processes of the human environment is an additional argument to use algae as a main element of water purification systems located within the building.

For our purification system, the structure Microalgae-Human-Mineralization as the simplest option has been chosen. In this case, it is possible to consider our treatment plant as a set of artificial aquatic ecosystems. Components of aquatic ecosystems is not different functionally from the components of terrestrial ecosystems. However there are some features. Organisms in the water biochemically are closely related to their environment and depend on the concentration of soluble substances. Due to the water
density, which is considerably greater than that of air, many aquatic organisms are free-floating. The water consists of spatially-distributed suspended substances, microalgae, and microbes. The water also creates the possibility of biochemical links between communities of hydrobionts through the allocation of many organisms, oxygen, carbon dioxide, and various metabolic products into the water. These substances, toxic or otherwise, stimulate other organisms, form the network by which organisms are connected implicitly, without direct contact. This feature of aquatic ecosystems allows describing their processes quite efficiently via the laws of matter conservation in the form of Reaction-Advection-Diffusion equations. As a result, we have an effective tool to create a reliable system for monitoring and controlling.

The functional diagram of the model Microalgae-Human-Mineralization is shown in Figure 2. The diagram has a simplified form but reflects the main streams of matter.

![Figure 1. Functional diagram of the model](image)

Stability and reliability of the structure Microalgae-Human-Mineralization is implemented by controlled flow circulating between a bioreactor-mineralizer and photoreactor cultivating algae. This combination of systems copies the natural self-purification processes, but which are intensified many times over.

The structure can be integrated into a residential building in the form of the following combinations of elements presented in Figure 3.

In Figure 3, there are the Arabic numerals depict the technological elements and the Roman numerals are the major flows substances. Namely: 1 is a photobioreactor cultivating algae, 2 is a solar concentrator, 3 is a circular bioaeration tank (a mineralizer), 4 is a desilter, 5 is an air-saturator, 6 is an air-blower, 7 is a circulation pump, 8 is an air valve, 9 is a distribution system, I is the waste water, II is air, III is flue gases, IV is the purified water, V is regenerated air with a high oxygen content, VI is the surplus sludge, VII is algae biomass.

Figure 3. Technological scheme of purification system

The key unit, the point of control in the system is the distribution system 9. Organic substances are oxidized in the mineralizer 3, mainly due to oxygen, which is formed in the photobioreactor. Separation of sludge biomass occurs in a desilter one part goes back into the reactor, and the excess is removed from the system. By regulating the flow rate of the oxygenated water returning into the system, we can control efficiency of the purification process. In the dark, the additional air supply can be carried out through the saturator 5. Saturator is a necessary device for dissolution of oxygen. As well as the flow rate of the oxygenated water coming into the unit 3 from the unit 4, the flow rate of air blown into the element 5 is a control parameter.

The idea of using photoreactors for tertiary sewage treatment within the traditional treatment plants appeared several years ago. There are a number of research and pilot projects, however, still no successfully implemented, cost-effective industrial objects [11]. In developing the technological schemes of functioning of photoreactors being part of the treatment facilities usually there is the need to reduce the growth rate of algae biomass. We propose to use a different concept. All biomass formed as a result of the purification process can be processed into biofuel and provide a source of clean energy. According to many researchers, buildings cultivating microalgae for energy production are the most promising direction in the construction of self-sufficient cities of the future. In addition, the algae biomass is well proven as an environmentally friendly building material. The world’s first prototype of a residential building that provides itself by electricity due to algae grown directly on the facade is the project BIQ House, which was presented at the International Exhibition in Hamburg in 2013. This project is a joint development of the international design...
Using algae for electricity and heat generation has been studied by many researchers for several years. Among them the French architectural bureau X-TU, which is supported by laboratory CNRS "Gepea" from Nantes University started the project "Symbio2". The first stage of the experiment is concentrated on designing of photobioreactors on the facade in Saint-Nazaire, near the laboratory Gepea. Then, based on the results obtained by research, it is planned to move to the industrial stage. It includes developing of reactors incinerator Alcéade Nantes. According to the developers, the use of microalgae will help reduce energy consumption by more than 50 % [13].

Applying the ideas of these technologies to our facility, we get, firstly, the ability to close flows of matter and energy in the system through the process of oxidation (combustion) of biomass for generation energy, and secondly, the absence of any limitations on the amount of produced biomass. Dynamics of processes into the phobiotoreactor depends on the illumination and the concentration of the carbon dioxide and nitrogen and phosphorus compounds. In photobioreactor 1, carbon dioxide, nitrogen and phosphorus come both from the reactor mineralizer and with the flue gases products of combustion of biofuel obtained from excess biomass. At the same time, microalgae purify the flue gases completely. Flue gases' supply mode (mode of biomass burning) can be one more control parameter of the system.

The next step is a statement of simulation and control problem of the dynamics of the structure Microalgae-Mineralization.

4 Simulation and Control of Artificial Ecosystems

4.1 Specificity of Research in the Field of Mathematical Ecology

The end of the 20th and the beginning of the 21st centuries have been marked by the rapid growth of works in the biological sciences and related disciplines. Along with the rapid accumulation of experimental material occurs the development of life science theory. Partly, the new results in the biological sciences are currently obtained with the help of mathematical modeling. New subject areas such as mathematical biology, biophysics, and ecology have been developed. However, specificity of studies in mathematical biology and ecology is such that, until recently, the main array of models includes either systems of ordinary differential equations or balance models. Presence of a number of complex models based on partial differential equations was seen as exotic supplement to tried and tested mathematical apparatus. This is due to the facts that, firstly, one of the main difficulties in the mathematical modeling of the biology and ecology is that the correct use of models requires meaningful problem statements. Obtaining such settings is possible only through close contact between experts in different subject areas. Often, without such interaction, biologists are unable to translate correctly seemingly obvious concepts of their subject area into the language of mathematics. At the same time models developed by mathematicians without the participation of biologists and environmental engineers can be pure abstractions, which are far from the practical use. Secondly, models of reaction-diffusion type may be rather "unpleasant". For example, they are able to produce various kinds of instabilities and cause spatially inhomogeneous structures and self-oscillatory regimes. All this is enhanced by the presence of flow in the system.

Unlike traditional engineering approach, we propose to use complex mathematical model of photo- and bioreactor in the form in the form of the laws of conservation of mass and momentum and a control problems for them. This mathematical model forming the basis for automated control system is a major instrument to compensate for the human factor and provide robust work of the treatment facilities of such a small volume.

4.2 Statement of a Primal and Control Problems

Dynamics of elementary volume biochemical reactor can be written in vector form using the laws of conservation of mass and momentum [3,4,5]

\[
\frac{\partial S}{\partial t} = U \cdot \text{grad}(S) + \text{div}(D \text{grad}(S)) + W, \tag{1}
\]

\[
W = N \mu, \tag{2}
\]

\[
\frac{\partial U}{\partial t} = -U \cdot \text{div}(U) + \text{div}(\nu \text{grad}(U)) - \frac{1}{\rho} \text{grad}(P) + G, \tag{3}
\]

\[
\text{div}(U) = 0, \tag{4}
\]

where \( S \) is the vector of concentrations of substances, \( N \) is the matrix of stoichiometric coefficients, \( \mu \) is the
vector of the rates of biochemical reactions, $W$ is the reaction term, $D$ is the vector of diffusion coefficients, $U$ is the flow velocity vector, $P$ is pressure, $G$ is the vector field of mass forces, $\rho$ is density.

In the final version of the model, the flow equations are reduced to Saint-Venant equations. The problem is formulated as a boundary value problem defined on a stratified set containing 1D and 2D area. Desilter model includes equations of flow and transport of suspended particles and the deformation of the layer of sludge. For photoreactor we use a model that is similar to [14]. The growth rate of biomass in photoreactor hyperbolically dependent on internal concentrations of nutrients (Droop model) [15]. Making a start from the analysis of models given in [16, 17], for reactor-mineralizer, we have chosen a model which takes into account the processes of biological phosphorus removal and nitrification inhibition by high doses of ammonium.

Further, we have to solve the problem of finding the optimal mode of the system. Boundary conjugation conditions subdomains of stratified set are controlled equations. Control variables are water rates at the connection points of subdomains 1, 3, 4 and the oxygen concentration at the 3. Also the flue gases flow rate can be added to the control parameters.

All equations are linearized. The time axis is divided into sections with assumption about the constancy of controls. Features of this method of adjoint equations are described by us in detail in [18, 19]. As a result, our problem is reduced to a series of linear programming problems, which allows us to significantly reduce the problem and perform almost real time control despite the complexity of underlying model.

5 Conclusion

The paper proposes a concept of treatment facilities, which can be used for water purification within residential building. The basic idea started from the conception of closed ecological LLS initially created for long-duration space flights. A distinctive feature of the project is a high degree of closure of matter flows in the system. The proposed technological scheme integrated in a building, is able to provide an autonomous existence for residents for a long time, for example as in a case of the harsh climatic conditions, with partial closure of the needs for water, pure atmosphere, decontamination and disposal of waste water. It also has the high efficiency of wastewater treatment, as it allows to partially replace the aeration by the natural process of water oxygenation during photosynthesis. A biomass formed as a result of the cleaning process can be processed into biofuel and provide a source of clean energy.

To create the robust work of closed exchange cycle and its components it is necessary to forming simulation model of processes. In this article, as the simulation model, we propose to use an initial boundary value problem for the linearized equations of Reaction-Advection-Diffusion and flow defined on stratified set. As a method for solving the optimal control problem, the best way is to use the adjoint equations method with an assumption about the constancy of controls during of some periods. In this case, the problem is reduced to the linear optimization problem and can be underlying for real-time control.

Further, a controlled closed-loop purification systems are being integrated in a building through the self-sustaining, interchangeable and standardized platforms, so-called mainboard-inspired one, situated beneath a modular home and controlling all water installments and energy components needed for a household, that being developed in [8].

Acknowledgments

The authors express their gratitude to Pavel Ermachenko for productive discussion of material of the article.

This work is in part funded by the Erasmus Mundus Action 2 programme of the European Union

References


Offshore Container Crane Systems with Robust Optimal Sliding Mode Control

R. M. T. Raja Ismail a, Nguyen D. That a,b and Q. P. Ha a

aFaculty of Engineering and Information Technology, University of Technology, Sydney, NSW 2007, Australia
bFaculty of Marine Electrical and Electronics Engineering, Vietnam Maritime University, Vietnam
E-mail: Raja.RajaIsmail@uts.edu.au; thatkd@vimaru.vn; Quang.Ha@uts.edu.au

Abstract -
Open-sea ship-to-ship transfer operation is an alternative way to avoid port congestion. This process involves a relatively small vessel which transports containers between the harbour and a large cargo ship equipped with a container crane. However, the presence of disturbances and uncertainties caused by harsh open-sea conditions can produce an excessive sway to the hoisting ropes of the crane system. This paper addresses the problem of robust sliding mode control for offshore container crane systems subject to bounded disturbances and uncertain parameters. The mathematical model of an offshore container crane is first derived whereas the effects of ocean waves and gusty winds are taken into account. Then, based on the linear quadratic regulator (LQR) design, a sliding surface is obtained to meet the required performance and stable dynamics for the closed-loop system. Finally, a robust sliding mode controller is proposed to drive the state trajectories of the offshore container crane system towards the sliding surface in finite time and maintain them subsequently on that surface. Simulation results are given to show that the proposed controller can significantly suppress the effects of uncertainties and disturbances from the vessel’s wave- and wind-induced motion and wind drag force on the payload.

Keywords - Sliding mode control, Offshore crane, Ocean wave, Gusty wind, Lyapunov, Linear Quadratic Regulator (LQR).

1 Introduction

Recently, along with rapid developments of the logistics industry, port congestion becomes a major issue all around the world due to the explosive increase in the trading volume [1, 2]. To solve this problem, some ambitious plans of port expansion have been proposed. However, it can be seen that expanding outwards is not a feasible option due to land constraints [3]. Consequently, a new method of transportation, the so-called ship-to-ship cargo transfer operation, is introduced [4]. This operation involves the transfer of containers between a huge container ship and a relatively smaller ship, known as a mobile harbour by using an offshore crane mounted on the container ship [5]. This method, emerging to become a promising solution to improve the port’s efficiency and productivity and reduce the operational cost, could enable the ports to stay competitive [6]. For this, the offshore container crane system has to meet stringent safety and efficiency requirements since open-sea waves, winds, and ocean currents easily cause the vessel to move away from a desired position both horizontally and vertically. Therefore, the design of the crane controller is a challenging problem due to the presence of parameter variations and disturbances, e.g. changes of load during the stevedoring operation, from wave- and wind-induced vibrations. These seriously affect the control system performance.

During the past decade, several control schemes have been developed to improve the crane operation performance such as input shaping control [7, 8], optimal control [9, 10] and fuzzy control [11]. However, conventional control methods, applied for crane systems in general, are inapplicable to offshore cranes which are subject to variations in the system matrices due to the changes of the payload mass and rope length and wave- and wind-induced disturbances of a large amplitude. To ensure the safety and reliability of offshore container crane operation, it is necessary to mitigate the effects of these uncertainties and disturbances.

Sliding mode control (SMC) has been recognised as a strong control methodology for the Lagrangian systems, for example the robotic excavator [12], construction robot [13], and tunneling shield machine [14]. The SMC design begins with the selection of a stable sliding surface with desired performance characteristics. Then, the discontinuous control is designed to drive the state trajectory towards the sliding surface and maintain it on this surface over time. The dynamic characteristics of the resulting closed-loop control system will be equivalently determined by the sliding surface design. In [5], the model of an offshore container crane was first introduced where the length of the hoisting rope was considered as a constant. Then, an extended mathematical model was developed in [15] by taking into account the variation of the rope length. For this, a second-order sliding mode controller was also proposed to deal with the problem of trajectory tracking of the crane trolley position and its hoisting rope length. Recently, a further improved result was reported in [16] by considering the offshore crane control problem in three dimensions. However, there remain interesting questions as to how to deal with the payload mass, rope length variation and input disturbances due to nonlinear frictions, strong waves and gusty winds which can significantly affect the payload sway angle of offshore cranes and their safe operations.
In this paper, the problem of robust sliding mode control is investigated for offshore container crane systems with bounded disturbances and uncertainties. Here, based on the dynamics analysis, the mathematical model of the ship-mounted container crane subject to parameter variations and disturbances, is derived for the first time. By utilizing the LQR design approach, a desired sliding surface is obtained. A robust sliding mode controller is then synthesized to drive the state trajectories of the closed-loop system towards the sliding surface in finite time and maintain it there after subsequent time. Simulation results are given to illustrate the feasibility of the proposed approach in terms of reducing the effects of uncertainties and disturbances caused by the changes in payload mass, length of rope and also the vibration effects of ocean waves and gusty winds.

2 Modelling and Problem Statement

The offshore container crane system considered in this study consists of a gantry crane mounted on a ship vessel as visualize in Figure 1, where \( \{O_{BxOyOz0}\} \) and \( \{O_{BxOyOzB}\} \) are the coordinate frames respectively of the ground and the vessel. The offshore crane system motion is represented by three generalized coordinates; in which, \( y \) is the position of the cart, \( l \) is the length of the rope, \( \theta \) is the swage angle induced by the motion of the cart, \( h \) is the vertical position of the cart, and \( m_c \) and \( m_p \) are respectively the masses of the cart and the payload. Let \( \zeta(t) \) be the heaving and \( \alpha(t) \) be the rolling angular displacement of the vessel. Thus, the position vectors of the cart and the payload with respect to the vessel coordinate frame \( \{O_{BxOyOzB}\} \) are

\[
p_c^B = [y, h]^T, \quad p_m^B = [y + l \sin \theta, h - l \cos \theta]^T.
\]

The position vectors of the cart and the payload with respect to the ground coordinate frame \( \{O_{0x0y0z0}\} \) can be obtained by multiplying the augmented position vectors with a homogeneous transformation matrix. Let \( \tilde{p}_c^B \) and \( \tilde{p}_m^B \) are respectively the augmented position vectors of \( p_c^B \) and \( p_m^B \) such that

\[
\tilde{p}_c^B = [(p_c^B)^T, 1]^T, \quad \tilde{p}_m^B = [(p_m^B)^T, 1]^T.
\]

Hence, the augmented position vectors of the cart and the payload with respect to \( \{O_{0x0y0z0}\} \) are given by

\[
\tilde{p}_c^0 = T_0^B \tilde{p}_c^B, \quad \tilde{p}_m^0 = T_0^B \tilde{p}_m^B,
\]

where

\[
T_0^B = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

Thus, the kinetic energy, potential energy and Lagrangian of the crane system are obtained respectively as

\[
\mathcal{K} = \frac{1}{2} m_c \|	ilde{p}_c^0\|^2 + \frac{1}{2} m_p \|	ilde{p}_m^0\|^2,
\]

\[
\mathcal{P} = m_c g(\tilde{p}_c^0)_z + m_p g(\tilde{p}_m^0)_z,
\]

\[
\mathcal{L} = \mathcal{K} - \mathcal{P},
\]

where \( (\cdot)_z \) denotes the \( z \)-component of the vector. By applying the Euler-Lagrange formulation

\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = \tau_i, \quad i = 1, 2, 3
\]

the dynamic model of the offshore crane system can be obtained in the following form:

\[
M(q)\ddot{q} + f(q, \dot{q}) + \Delta f(t, q, \dot{q}) = \tau(t) - \tau_c(t), \quad (1)
\]

where \( q = [y, l, \theta]^T \in \mathbb{R}^3 \) is the vector of generalized coordinates and \( \tau(t) = [F_p(t), F_l(t), 0]^T \in \mathbb{R}^3 \) is the vector of input forces. In (1), \( M(q) \in \mathbb{R}^{3 \times 3} \) is the inertia matrix, \( f(q, \dot{q}) \in \mathbb{R}^3 \) is the vector of centrifugal-Coriolis and gravity forces, \( \Delta f(t, q, \dot{q}) \in \mathbb{R}^3 \) is the vector of disturbances in the system due to open-sea wave-induced vibrations, and \( \tau_c(t) \in \mathbb{R}^3 \) is the vector of disturbances due to Coulomb friction and wind drag force on the payload.

The system matrices are derived as follows:

\[
M(q) = \begin{bmatrix} m_c + m_p & m_p \sin \theta & m_p l \cos \theta \\ m_p \sin \theta & m_p & 0 \\ m_p l \cos \theta & 0 & m_p l^2 \end{bmatrix},
\]

\[
f(q, \dot{q}) = \begin{bmatrix} 2m_p l \dot{\theta} \cos \theta - m_p l^2 \dot{\theta}^2 \sin \theta + K_{cy} \ddot{y} \\ -m_p l \dot{\theta}^2 - m_p g \cos \theta + K_{cl} \dot{l} \\ 2m_p l \dot{\theta} + m_p g l \sin \theta + K_{cr} \dot{\theta} \end{bmatrix},
\]

\[
\Delta f(t, q, \dot{q}) = \begin{bmatrix} \Delta f_1 \Delta f_2 \Delta f_3 \end{bmatrix}^T,
\]

\[
\tau_c(t) = \begin{bmatrix} F_{cy} \text{sign} \ddot{y} \\ F_{cl} \text{sign} \dot{l} \\ \tau_\theta \text{sign} \theta + \tau_{wd}(t) \sin 2\pi f_w t \end{bmatrix},
\]

where

\[
\Delta f_1 = (m_c + m_p) \left( -y \ddot{a}^2 + \dot{h} \dot{a} - (g + \ddot{\zeta}) \sin \alpha \right) + 2m_p (\dot{\theta} \dot{a} \sin \theta - \dot{\alpha} \cos \theta) - m_p l (\sin \theta \dot{a}^2 + \cos \theta),
\]

\[
\Delta f_2 = m_p \left( \left( y \cos \theta + \dot{h} \sin \theta \right) \ddot{a} + 2g \dot{a} \cos \theta + 2 \dot{\theta} \dot{a} - \left( y \sin \theta - h \cos \theta + l \right) \dot{a}^2 + g \cos \theta - (g + \ddot{\zeta}) \cos (\theta - \alpha) \right),
\]

\[
\Delta f_3 = m_p \left( -2l \ddot{\alpha} + 2y \dot{a} \sin \theta - (y \dot{g} \sin \theta - l \dot{h} \cos \theta - l^2) \dot{a} - g l \sin \dot{\theta} + l (y \cos \theta + h \sin \theta) \dot{a}^2 + (g + \ddot{\zeta}) l \sin (\theta - \alpha) \right),
\]

in which \( K_{cy}, K_{cl} \) and \( K_{cr} \) are the viscous friction coefficients, and \( F_{cy}, F_{cl} \) and \( \tau_\theta \) are the Coulomb friction coefficients. The term \( \tau_{wd}(t) \) represents the magnitude of the torque produced by the wind drag force on the payload and it can be calculated by using the following formula [17]:

\[
\tau_{wd}(t) = \frac{1}{2} \rho_w v_w^2 c_d A_p [l(t) + L_c],
\]

where \( \rho_w \) is the density of air, \( v_w \) is the velocity of the wind, \( c_d \) is the drag coefficient, \( A_p \) is the effective surface area of the payload, and \( L_c \) is the average distance from the hook to the payload center of gravity.
System (1) is a typical underactuated system since the number of control inputs is less than the number of generalized coordinates. Thus, we can partition the vector of generalized coordinates into actuated and unactuated parts, such that \( q = [q_a^T, q_u^T]^T \), where \( q_u = [y, l]^T \) and \( q_a = \theta \). It follows that, the matrices and vectors in (1) can be partitioned as

\[
M(q) = \begin{bmatrix} M_{aa} & M_{au} \\ M_{ta} & M_{uu} \end{bmatrix}, \quad f(q, \dot{q}) = \begin{bmatrix} f_a \\ f_u \end{bmatrix},
\]

\[
\Delta f(t, q, \dot{q}) = \begin{bmatrix} \Delta f_a \\ \Delta f_u \end{bmatrix}, \quad \tau(t) = \begin{bmatrix} u_a(t) \\ 0 \end{bmatrix}, \quad \tau_c(t) = \begin{bmatrix} \tau_{ca}(t) \\ \tau_{cu}(t) \end{bmatrix}.
\]

The sub-blocks of the above matrices and vectors are obtained based on the dimensions of vectors \( q_a \) and \( q_u \).

For the purpose of control design, we introduce a state variable vector \( x \in \mathbb{R}^n \) as

\[
x = [x_1, x_2, x_3, x_4, x_5, x_6]^T = [y, l, \theta, \dot{y}, \dot{l}, \dot{\theta}]^T.
\]

Hence, by substituting \( [q^T, \dot{q}^T] \) in system (1) with the state vector \( x \), the offshore container crane dynamics can be expressed in the state-space as follows:

\[
\dot{x} = F(x) + \Delta F(t, x) + G(x)u_a(t) + H(x)\omega(t), \quad (2)
\]

where

\[
F(x) = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ \tau_a \\ -M_{au}^{-1}(M_{uu}^T\tau_a + f_u) \end{bmatrix},
\]

\[
\Delta F(x) = \begin{bmatrix} 0_{3 \times 1} \\ \Delta \tau_a \\ -M_{au}^{-1}(M_{uu}^T\Delta \tau_a + \Delta f_u) \end{bmatrix},
\]

\[
G(x) = \begin{bmatrix} 0_{3 \times 2} \\ M_{uu}^{-1} \\ -M_{au}^{-1}M_{uu}^{-1} \end{bmatrix},
\]

\[
H(x) = \begin{bmatrix} H_{21} \\ H_{22} \\ H_{31} \\ H_{32} \end{bmatrix},
\]

\[
u_a(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} F_v(t) \\ F_l(t) \end{bmatrix},
\]

\[
\omega(t) = \tau_c(t),
\]

\[
\tau_a = M_{aa} - M_{au}M_{uu}^{-1}M_{au}^T,
\]

\[
\Delta \tau_a = -M_{uu}^{-1}(f_a - M_{au}M_{uu}^{-1}f_u),
\]

\[
H_{21} = -M_{uu}^{-1},
\]

\[
H_{22} = M_{uu}M_{uu}^{-1},
\]

\[
H_{31} = M_{uu}^{-1}M_{uu}^T, 
\]

\[
H_{32} = -M_{uu}^{-1}(M_{uu}^TM_{uu}^{-1} + 1).
\]

The linearised dynamic model of the offshore crane system (2) about an operating point \((x_0, u_0)\) can be obtained in the form of

\[
\dot{\bar{x}} = (A + \Delta A(t))\bar{x} + B(u(t) + Du_0(t)), \quad (3)
\]

where \( \bar{x} = x(t) - x_0 \) and \( u(t) = u_a(t) - u_0 \). By choosing \( x_0 = [y_0, l_0, 0, 0, 0, 0]^T \), the state vector of the local model can be represented as \( \bar{x} = [\Delta y, \Delta l, \theta, \dot{y}, \dot{l}, \dot{\theta}]^T \), and \( u_0 \) can be found as

\[
u_0 = -G^\dagger(x_0)F(x_0),
\]

in which \( G^\dagger \) denotes the Moore-Penrose pseudo-inverse of matrix \( G \). Thus, the system matrix, the input matrix,
and the unmatched disturbance matrices are obtained as follows:

\[
A = \frac{\partial F(x)}{\partial x} + \sum_{j=1}^{2} \frac{\partial G_j(x)}{\partial x} u_j \bigg|_{x=x_0} ,
\]

\[
B = G(x_0),
\]

\[
\Delta A(t) = \frac{\partial \Delta F(t, x)}{\partial x} \bigg|_{x=x_0} ,
\]

\[
D = H(x_0),
\]

where \( G_j (j = 1, 2) \) denotes the \( j \)th column of matrix \( G \).

3 Control Design

In this section, a robust sliding mode control scheme is proposed to deal with the problem of trajectory tracking of the crane system whereas the effects of the ocean waves and gusty winds is mitigated.

3.1 Crane trajectory

For the desired trajectory of the offshore container crane, we assume that there are three stages:

Stage 1 (Picking up the container with time \( t_1 \)): The hoisting rope length decreases from \( l_D \), where the container is on the deck of the large ship to \( l_U \), where the container is in the upper position with a maximum lateral velocity \( v_R \). Thus, we have

\[
t_1 = \frac{l_D - l_U}{v_R}.
\]

It should be noted that the rope length \( l_D \) and \( l_U \) are assumed to be constant between the different containers.

Stage 2 (Placing the container on the smaller ship with time \( t_2 \)): The rope length increases from \( l_U \) to \( l_D \). This is the reverse process for lifting the container.

Stage 3 (Moving the container with a maximum longitudinal velocity \( v_R \)): Assuming the whole distance of travel of lifting, moving and placing the container is \( y_F \) with the corresponding time \( t_F \), the placing phase then will take a time of \( t_F - 2t_1 \). Denote \( y, \dot{y}, l, \dot{l} \) are longitudinal position, velocity of the container and the rope length, its rate of change respectively. We have the following diagrams, from which desired trajectories \( y_d \) and \( \dot{y}_d \) can be derived from given parameters \( l_D, l_U, v_R, v_y, y_F \). Desired trajectories of position and velocity of the container are obtained respectively as

\[
y_d(t) = \begin{cases} 
\frac{1}{2} a t^2, & 0 \leq t \leq t_1, \\
v_y t - \frac{1}{2} a t^2, & t_1 \leq t \leq t_2, \\
\frac{1}{2} a t^2 + v_y \frac{t^2}{t_1} t + y_F - \frac{1}{2} a t^2, & t_2 \leq t \leq t_F, \\
at, & 0 \leq t \leq t_1, \\
v_y, & t_1 \leq t \leq t_2, \\
- a t + v_y \frac{t^2}{t_1} t, & t_2 \leq t \leq t_F,
\end{cases}
\]

\[
\dot{y}_d(t) = \begin{cases} 
\frac{1}{2} a, & 0 \leq t \leq t_1, \\
v_y, & t_1 \leq t \leq t_2, \\
- a t + v_y \frac{t^2}{t_1} t, & t_2 \leq t \leq t_F, \\
0, & t_2 \leq t \leq t_F.
\end{cases}
\]

where \( a = \frac{v_y}{t_1} \), \( t_2 = t_F - t_1 \).

Similarly, desired trajectories of the hoisting rope length and its rate of change are

\[
l_d(t) = \begin{cases} 
- \frac{1}{2} b t^2 + l_D, & 0 \leq t \leq \frac{t_1}{2}, \\
\frac{1}{2} b t^2 - 2 v_R t + v_R t_1 + l_U, & \frac{t_1}{2} \leq t \leq t_1, \\
l_U, & t_1 \leq t \leq t_2, \\
\frac{1}{2} b t^2 - b t_2 + \frac{v_R^2}{4} t_2^2 + l_U, & t_2 \leq t \leq t_F - \frac{t_2}{2}, \\
- \frac{1}{2} b t^2 + b t_2 t - \frac{1}{2} b t_2^2 + l_D, & t_F - \frac{t_2}{2} \leq t \leq t_F, \\
- b t, & 0 \leq t \leq \frac{t_2}{2}, \\
bt - bt_2, & \frac{t_2}{2} \leq t \leq t_1, \\
bt - bt_2, & t_2 \leq t \leq t_F - \frac{b t_2}{2}, \\
-b t + b t_F, & t_F - \frac{b t_2}{2} \leq t \leq t_F,
\end{cases}
\]

where \( b = \frac{2 v_R}{t_1} \).

3.2 Problem statement

Let us consider a reference model as follows:

\[
\dot{x}_d(t) = A_d x_d(t) + B_d r_d(t),
\]

where \( x_d(t) = [y_d, l_d, \theta_d, y_d, \dot{\theta}_d]^T \) and \( r_d(t) \in \mathbb{R}^2 \) is the bounded reference input. The matrices \( A_d \) and \( B_d \) are designed in a way that there exist compatibly dimensioned matrices \( K \) and \( L \) to satisfy the matching condition:

\[
BK = A_d - A, \\
BL = B_d.
\]

Define a tracking error state \( x_e \) as the difference between the plant and the reference model state responses:

\[
x_e = x(t) - x_d(t).
\]

Consequently, the following state-space equation with the state vector error is obtained as

\[
\dot{x}_e(t) = A_d x_e(t) + (A - A_d) \bar{\tau}(t) + \Delta A \tau(t) + Bu(t) - B_d r_d(t) + D \omega(t). 
\]

For the purpose of sliding surface design, a transformation matrix \( T \in \mathbb{R}^{6 \times 6} \) is introduced such that

\[
TB = \begin{bmatrix} 0 \\ B_2 \end{bmatrix}
\]

where \( B_2 \in \mathbb{R}^{2 \times 2} \) is nonsingular. By using the coordinate transformation \( z = T \tau \), (4) can be rewritten in the form of the new coordinate:

\[
\dot{e}(t) = \bar{A}_d e(t) + (\bar{A} - \bar{A}_d) z(t) + \Delta \bar{A} z(t) + \bar{B} u(t) - \bar{B}_d r_d(t) + \bar{D} \omega(t),
\]

where \( e = T x_e, \bar{A} = T A T^{-1}, \Delta \bar{A} = T \Delta A T^{-1}, \bar{A}_d = T A_d T^{-1}, \bar{B} = TB, \bar{B}_d = TB_d \) and \( \bar{D} = TD \).

Without loss of generality, we assume that the disturbance \( \omega(t) \) and the system uncertainty \( \Delta \bar{A}(t) \) are norm-bounded as

\[
\| \omega(t) \| \leq \omega_p, \quad \| \Delta \bar{A}(t) \| \leq \beta,
\]
where $\varpi, \beta$ are known positive scalars.

Our purpose is first to design a stable sliding surface using the linear quadratic regulator design approach to mitigate the effects of the ocean waves and gusty winds on the system state. Then, a robust sliding mode controller is synthesized to guarantee that the system sliding motion converges exponentially to a ball whose radius and the rate of exponential convergence can be chosen arbitrarily.

### 3.3 Sliding surface

The sliding function is defined in terms of trajectory tracking errors as follows

$$s(t) = [s_1(t), s_2(t)]^T = C e(t) = -Ce_1(t) + e_2(t), \quad (6)$$

where $C$ is a constant matrix to be designed. Consider the following quadratic performance index

$$J = \frac{1}{2} \int_{t_s}^{\infty} e^T(t)Qe(t)dt, \quad (7)$$

where $Q = [Q_{11}, Q_{12}, Q_{22}]$ is a given symmetric positive definite matrix and $t_s$ is the starting time which indicates the induction of the sliding motion. By noting that

$$2e_1^T(t)Q_{12}e_2(t) + e_2^T(t)Q_{22}e_2(t)$$

$$= (e_2(t) + Q_{22}^{-1}Q_{21}e_1(t))^T Q_{22}(e_2(t) + Q_{22}^{-1}Q_{21}e_1(t))$$

$$- e_1^T(t)Q_{12}Q_{22}e_1(t),$$

(7) can be rewritten in the form of

$$J = \frac{1}{2} \int_{t_s}^{\infty} (e_1^T(t)Q_{11}e_1(t) + v^T(t)Q_{22}v(t)) dt,$$

where

$$\bar{Q} = Q_{11} - Q_{12}Q_{22}^{-1}Q_{21},$$

$$v(t) = e_2(t) + Q_{22}^{-1}Q_{21}e_1(t).$$

Based on the LQR minimisation of $J$ in association with the nominal system in (5), we obtain

$$v(t) = -Q_{22}^{-1}A_{12}^TPe_1(t),$$

where $P$ satisfies the following equation

$$\bar{A}^TP + PA - PA_{12}Q_{22}^{-1}A_{12}^TP + \bar{Q} = 0,$$

in which $\bar{A} = A_{11} - A_{12}Q_{22}^{-1}Q_{21}$ and $A_{12}$ is the subblocks obtained from partitioning matrix $A$ in (5). Consequently, we obtain

$$e_2(t) = -Q_{22}^{-1}(A_{12}^TP + Q_{21})e_1(t). \quad (8)$$

During the sliding motion, we have $s(t) = 0$ so that

$$e_2(t) = Ce_1(t). \quad (9)$$

Thus, by comparing (8) and (9), the design matrix of the sliding function is obtained explicitly as

$$C = Q_{22}^{-1}(A_{12}^TP + Q_{21}). \quad (10)$$

### 3.4 Robust optimal sliding mode control

Before presenting our proposed control scheme, the following definition and lemma are introduced.

**Definition 1:** The solution of system (5) is uniformly exponentially convergent to a ball $B(0, r) = \{e \in \mathbb{R}^n : \|e\| \leq r\}$ with rate $\gamma > 0$ if for any $\xi > 0$, there exists $k(\xi) > 0$ such that

$$\|e(t)\| \leq r + k(\xi) \exp(-\gamma t), \quad \forall t \geq 0.$$

**Lemma 1:** [18] Let $V(t)$ be a continuous positive definite function for all $t \geq 0$, $k^* \geq 0$ and

$$\dot{V}(t) \leq -\eta V(t) + \nu, \quad \forall t \geq 0,$$

where $\eta$ and $\nu$ are positive constants, then

$$\dot{V}(t) \leq r + k^* \exp(-\gamma t), \quad \forall t \geq 0,$$

in which $r = \nu/\eta$ and $\gamma = \eta$ is the exponential convergence rate.

The control scheme proposed here has the form of

$$u(t) = u_E(t) + u_R(t), \quad (11)$$

where $u_E(t)$ and $u_R(t)$ are the equivalent and switching control. The equivalent control which maintains the sliding motion on the sliding surface is defined as

$$u_E(t) = u_{E_1}(t) + u_{E_2}(t)$$

where

$$u_{E_1}(t) = -(\bar{C}\bar{B})^{-1}(\bar{C}\bar{A}e(t) + \Pi s(t)), \quad (12)$$

and

$$u_{E_2}(t) = \bar{K}z(t) + Lr_d(t), \quad (13)$$

in which $\Pi$ is a design diagonal matrix with real distinct positive eigenvalues and $\bar{K} = KT^{-1}$. For a given convergence ball radius $r_0$, the following switching control $u_R(t)$ is designed to force the system trajectories towards the prescribed sliding surface

$$u_R(t) = -(\bar{C}\bar{B})^{-1} \frac{\mu s(t)}{\|s(t)\| + \varepsilon}, \quad (14)$$

where

$$\mu = \frac{2r_0\lambda_{\text{min}}(\Pi)}{\varepsilon}$$

and $\varepsilon > 0$ is a small positive scalar for chattering reduction to be selected according to the theorem stated below.

**Theorem 1:** For with given bounds of the system uncertainty $\beta$ and disturbance $\varpi$, and radius $r_0$, the state trajectories of offshore container crane system (4) are driven towards the sliding function designed as in (6) under the following control law

$$u(t) = -(\bar{C}\bar{B})^{-1} \left(\bar{C}\bar{A}e(t) + \Pi s(t) + \frac{\mu s(t)}{\|s(t)\| + \varepsilon}\right)$$

$$+ \bar{K}z(t) + Lr_d(t),$$

(15)
where \( \varepsilon \) is chosen to be sufficiently small such that
\[
\varepsilon \leq \frac{2\tau_0 \lambda_{\min}(\Pi)}{\|C^T\| \|\varpi_p\| + \|C\| \|z(t)\|}.
\]  

**Proof:** Consider the Lyapunov function
\[
V = \frac{1}{2} s^T(t) s(t).
\]
By taking its derivative along the solutions of (5), we obtain
\[
\dot{V}(t) = s^T(t) \left( C \Delta \Phi(t) z(t) - \Pi s(t) \right)
- \frac{\mu s(t)}{\|s(t)\| + \varepsilon} + C^T \omega(t)
\leq -\lambda_{\min}(\Pi) \|s(t)\|^2 + \beta \|C\| \|z(t)\| \|s(t)\|
- \frac{\mu s(t)}{\|s(t)\| + \varepsilon} + \|C^T\| \|s(t)\| \|\varpi_p\|.
\]
From (16), we obtain
\[
\mu \geq \beta \|C\| \|z(t)\| + \|C^T\| \|\varpi_p\|.
\]
Thus,
\[
\dot{V}(t) \leq -\lambda_{\min}(\Pi) \|s(t)\|^2 + \left( \beta \|C\| \|z(t)\| + \|C^T\| \|\varpi_p\| \right) \frac{\|s(t)\| \varepsilon}{\|s(t)\| + \varepsilon}.
\]
Consequently, by using inequalities \( \frac{ab}{a+b} \leq b, \forall a, b > 0 \), we obtain the following inequality
\[
\dot{V}(t) \leq -2\lambda_{\min}(\Pi) V(t) + \left( \beta \|C\| \|z(t)\| + \|C^T\| \|\varpi_p\| \right) \varepsilon.
\]
Thus, from Definition 1 and Lemma 1, we get
\[
V(t) \leq r_0 + k_1 \exp(-\gamma t), \quad \forall t \geq 0,
\]
where
\[
r_0 = \frac{\left( \beta \|C\| \|z(t)\| + \|C^T\| \|\varpi_p\| \right) \varepsilon}{2\lambda_{\min}(\Pi)}.
\]
and \( \gamma = 2\lambda_{\min}(\Pi) \). The proof is completed.

4 \ Results and Discussion

In this study, numerical values of the offshore container crane system parameters are listed as \( m_c = 6 \times 10^3 \) kg, \( m_p = 20 \times 10^3 \) kg, \( h = 10 \) m, \( K_{cy} = 600 \) N/m/s\(^{-1}\), \( K_{cl} = 200 \) N/m/s\(^{-1}\), \( K_{ch} = 100 \) N/m/rad/s\(^{-1}\) and \( g = 9.81 \) m/s\(^{-2}\). The nominal state vector is chosen as \( x_0 = [10, 8, 0, 0, 0, 0, 0]^T \), which provides \( u_0 = [0, -196.14]^T \) N. For the sake of illustration, the following parameters are provided as \( l_c = 10 \) m, \( l_U = 4 \) m, \( v_R = 3 \) m/s, \( v_y = 0.63 \) m/s, \( y_F = 10 \) m. The values of the constants in \( \omega(t) \) are listed as \( F_{cy} = 5 \) kN, \( F_{cl} = 2 \) kN, \( F_{ch} = 2 \) kN, \( \rho_c = 1.225 \) kg/m\(^3\), \( v_w = 15 \) m/s, \( c_d = 1.05 \), \( A_p = 12 \) m\(^2\), and \( L_c = 1.2 \) m. The heaving acceleration and the rolling angular displacement of the vessel to accommodate ocean waves in an allowable range are assumed to be respectively \( \tilde{c}(t) = 0.4 \sin t \) m/s\(^2\) and \( \alpha(t) = \frac{\pi}{30} \cos t \) rad [19]. The details of matrices \( A, \Delta A(t), B, D, K \) and \( L \) are then listed as follows:

\[
A = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & -3.270 & -0.41 & 0.0021 & 0 \\
0 & 0 & 0 & -5.314 & 0.0125 & 0 & -0.0034 \\
\end{bmatrix},
\]

\[
\Delta A = \begin{bmatrix}
\Psi_{13} & 0 \\
\Psi_{13} & 0 \\
\Psi_{33} & 0 \\
\Psi_{33} & 0 \\
\end{bmatrix},
\]

\[
\Psi = \begin{bmatrix}
3.33\alpha^2 & \gamma \\
0 & 0 \end{bmatrix},
\]

\[
B = \begin{bmatrix}
0 & 0 & 0 & 1.667 & 0 & -2.083 \\
0 & 0 & 0 & 0 & 0 & 5.0 \end{bmatrix},
\]

\[
D = \begin{bmatrix}
D_1 \\
D_2 \\
\end{bmatrix},
\]

\[
D_1 = \begin{bmatrix}
0 & 0 & 0 & 1.667 & 0 & -2.083 \\
0 & 0 & 0 & 0 & 0 & 5.0 \end{bmatrix},
\]

\[
D_2 = \begin{bmatrix}
-1.667 & 0 & 2.803 \\
0 & -0.5 & 0 \\
2.803 & 0 & -2.604 \end{bmatrix},
\]

Figure 2: Trajectory tracking responses of the (a) cart position; (b) rope length; and (c) swing angle.
The transformation matrix is obtained as
\[
T = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 \\
-0.625 & 0 & 0 & 0.610 & 0 & 0.488 \\
0.781 & 0 & 0 & 0.488 & 0 & 0.390 \\
0 & 0 & 0 & -0.625 & 0 & 0.781 \\
0 & 0 & 0 & 0 & -1 & 0
\end{bmatrix}
\]

Using the quadratic minimisation, the state weighting matrix \(Q\) is chosen as \(Q = TRT^{-1}\), \(R = \text{diag}(10, 10, 5, 1, 1, 1)\) which provides the following matrix \(C\):
\[
C = \begin{bmatrix}
1.798 & -0.949 & 0 & -4.809 & 1 & 0 \\
0 & 0 & 3.162 & 0 & 0 & 1
\end{bmatrix}
\]

The upper bounds of the system disturbance \(\omega_{p}\) and uncertainty \(\beta\) are \textit{a priori} selected as 0.5 and 5.8 respectively. From Theorem 2, by choosing the exponential decay rate \(\gamma = 30\), radius \(r_0 = 0.004\) for the sliding surface trajectories \(s(t)\) and \(\varepsilon = 0.0025\), we obtain \(\mu = 48\).

The displacements and velocities of the cart position and rope length are assigned to track the trajectories as defined in Section 3. Aside from that, it can be shown that if the swing angle tracks the roll angle of the vessel during the operation process, the payload will be held on the vertical plane in the frame \(\{O_0x_0y_0z_0\}\). Therefore, the desired payload swing trajectory is chosen as \(\theta_d(t) = \alpha(t)\). Figure 2 depicts the trajectory tracking responses of cart position, rope length and swing angle and Figure 3 depicts the trajectory tracking responses of the corresponding velocities. Continuous small oscillations are occurred in the cart position and its corresponding velocity due to the persistent rolling vibration-induced motion of the vessel. This continuous small oscillation is apparent in the cart velocity response of Figure 3(a) during the container placement stage (Stage 2). However, the rope length response is less affected by the presence of heaving motion of the vessel owing to robustness of the control system.

Figure 4 shows the plot of switching functions and Figure 5 shows the control forces of the system. The effect of persistent rolling vibration-induced motion of the vessel can also be seen in the responses of both \(s_1(t)\) and \(u_1(t)\). Based on the quadratic minimisation approach, \(s_1(t)\) consists of the coupled motion of the cart and the swing angle whereas \(s_2(t)\) corresponds to the rope length dynamics. Overall, based on Figure 2(a)-(b) and Figure 3(a)-(b), an excellent decoupling of the cart position and swing angle
is exhibited.

5 Conclusion

In this paper, the problem of robust sliding mode control for offshore container crane systems with bounded disturbances and uncertainties has been addressed. By taking the effects of payload mass and length of rope changes as well as vibrations induced by ocean waves and gusty winds into account, the mathematical model of offshore container crane systems is derived for the first time. An LQR-based design approach is developed to obtain the sliding surface to achieve the optimal performance of the equivalent dynamics. To track the crane’s desired trajectory, robust sliding mode control law is then designed to drive the state variables of the system towards the sliding surface in finite time and maintain them on that surface after subsequent time. Extensive simulation results are provided to demonstrate good tracking performance of the proposed controller for offshore crane systems in dealing with the harsh open-sea conditions.

References


Force Control of Cleaning Tool System for Building Wall Maintenance Robot on Built-in Guide Rail

C.Y. Shin¹, S.M. Moon¹, J.H. Kwon¹, J. Huh¹ and D. Hong²

¹Graduate School of Mechanical Engineering, Korea University, Korea
²School of Mechanical Engineering, Korea University, Korea

E-mail: dhhong@korea.ac.kr

Abstract -
In modern days, there has been an increasing trend to build high-rise buildings. Until now, the methods to maintain building wall have relied on human efforts. However, there exist dangers and difficulties caused by height of buildings. Due to this problem, the need for system that can maintain building wall has been increasing and various systems have been developed. On this, the building maintenance robot that is used for built-in guide rail and tool system for cleaning wall of building was developed. Cleaning tool system equipped to building maintenance robot system has squeezing and injection functions. This system operates with spreading the water through injection part and eliminating dust and contaminants using squeezing module. Throughout this process, flow control is needed to spread water evenly on window at the changing velocity of robot. Also, the angle of the squeezing module and the force which is applied to window should be controlled to prevent damage of wall and to remove contaminants completely. This paper suggests algorithms for control of position and force of the squeezing module to implement cleaning building window composed of segmented façade. In addition, flow control algorithm for uniform injection is proposed. After that, the results of experiment in consistent horizontal robot velocity and specific contact angles between squeezing part and façade is presented and results are shown with graph and analyzed.

Keywords –
Force control; Flow control; Building maintenance robot; Façade cleaning

1 Introduction

As the advent of high-rise buildings, managing building wall using conventional methods with human efforts become dangerous and cost. To resolve this, robotic systems for maintaining building wall have been developed [1-4]. One of them is based on built-in guide rail [1]. This system follows the rail horizontally which is already installed in buildings and implements the work. After finishing the work in one floor, it moves to other floor by vertical robot which moves via wire winch system. When robot conducts cleaning process during moving horizontally, it needs tool system for removing contaminants efficiently. Most of the wall of buildings constructed today are made up of windows. Thus, previously developed tool system uses brush and water to clean. However, this method could not remove water perfectly by brush and there remains stain of water in window. To overcome this problem, water and squeezing rubber are used to clean window in this tool system. It is composed of the injection part and the squeezing part. Injection part sprays the water evenly on the wall and squeezing part removes contaminants and water by contacting to wall. During the process, algorithm for flow control of water injection and force control of squeezing part are used.

This paper introduces the structure of tool system and applied algorithms covering flow control, position and force control of squeezing part. Experiment in the interior parallel façade with consistent horizontal robot velocity and specific contact angles between squeezing part and façade is presented and results are shown with graph and analyzed.

2 Injection Part

2.1 Structure

Injection part has 4 PWM nozzles, a water pump, an air tank, and an air compressor. Figure [1] shows the installation of nozzles on robot and magnified picture of a nozzle with tubes connected to air compressor and water pump. Water pump provides water to nozzles through water filter and pressure of water is measured using water pressure gauge. Air compressor injects compressed air to air tank and compressed air is delivered to nozzle when it satisfies the desired air pressure. By using these water and compressed air, injection part spreads out water evenly on the wall based on two-fluid spray process [5].
2.2 Flow Control Algorithm

Flow control algorithm is presented in Figure [2]. Before operation of injection, pressure of air and water which are used in nozzle needs to be maintained at desired pressure. After process of pressure maintaining, injection should be controlled according to robot’s horizontal moving speed. There can exist water overlap region if injection speed is high comparing with robot velocity or dry region on vice versa. Efficiency of cleaning process would be diminished in those cases. Also, water would be wasted and it could not implement cleaning process to wide area because of the lack of water. Thus, by getting robot horizontal speed as input, nozzle regulates the speed of injection.

\[
\text{Injection Speed} \propto \text{Frequency} \tag{1}
\]

\[
\text{Duty ratio} = \frac{T_{ON}}{T_{ON} + T_{OFF}} \tag{2}
\]

\[
\text{Water Amount} \propto \text{Duty ratio} \tag{3}
\]

Frequency of injection and amount of water during one injection period can be handled through PWM signal of nozzle. Figure [3] shows PWM signal of nozzle in one second. Nozzle spreads water during ON-period. Thus, injection frequency is changed with signal frequency, and amount of water used in one cycle relates with time ratio of ON respect to one cycle. Based on this, amount of water used in cleaning process can be minimized and injection frequency is controlled.

3 Squeezing Part

3.1 Structure

Squeezing part is separated into upper and lower parts. Structure of them are equivalent, and it is divided for covering the specific area of wall, which is split, by controlling differently. Each part consists of two motors, a screw structure, 4 load-cells, 3 switch sensors, and 2 squeezing modules. One of motor, named as pressing motor in Figure [4] takes charge of moving forward to wall and pressing force. Another motor named as angle motor rotates the squeezing modules and changes the contact angle between squeezing module and wall. As it can be seen from figure, pressing motor is linked to a screw and this changes the rotating movement of motor to linear movement of squeezing module. Rubber is
placed on the edge of module to remove water and contaminants. Among four load-cells, two of them are installed in linkage between screw and squeezing module to measure the pressing force which is \( F_y \) in the Figure [5] and others are placed between the rotate joint and squeezing module to measure the friction force \( F_x \) occurred in the wall. Three of switch sensors are located at each of the beginning and last position of the screw and initialization position of squeezing module. These switches are used for position initialization of squeezing part when it starts and getting reference point during operation.

### 3.2 Position and Force Control

When cleaning the wall, the angle between the rubber and the wall should be maintained for efficient performance. Also, consistent and sufficient force should be applied to window. It leads to enough friction force to remove contaminants. Contact angle is sustained by position control of angle motor and force applied to wall is controlled by pressing motor using force control. Force control is categorized into impedance control and hybrid control. Predetermined relationship of end-effector force and position is maintained in the impedance control [6]. Explicit and implicit control are used in hybrid control with respect to model. Explicit control needs well analyzed model and applies motor torque directly [7]. Implicit control uses velocity and position of robot without exact model [8]. On the other view, whether reacting to force sensor directly (active compliance) or not (passive compliance) is also issue [9]. This paper uses implicit force control based on velocity and active compliance using force sensor input. Reasons for selecting implicit force control are that model is not needed and approach velocity is limited during no contact and well performance of regulating force during contact. Approach velocity is determined by equation (4). Velocity is proportionate to error between desired force \( (F_D) \) and current force sensing data \( (F_S) \). Coefficient \( (K_c) \) is determined by empirical method in this paper. Environment is assumed as rigid because it is composed of façades.

\[
v = K_c \times (F_D - F_S)
\]

### 3.3 Control Algorithm

Control algorithm of squeezer module is separated into initialization mode, contact moving mode, force control mode. As shown in Figure [6], in the initialization mode, controller moves the squeezing module to initial point and using that point as reference when controlling position. When controller finds the initial point, it uses switch sensors. At the starting point, if bottom limit switch is pushed, it moves forward until bottom limit switch is off. In the case limit switch is not pushed, it operates with opposite process. These processes locate the robot to edge of limit switch. Before the beginning of force control, difference between desired force and applied force is large because there is space between squeezing module and wall. This space leads to not sufficient contact force. If robot moves horizontally during this situation, area which is covered during start of cleaning will not be cleaned. To
prevent this, at the beginning, robot waits until desired contact force in the contact moving mode. Next to these processes, controller gets force data from load cells and compares it with desired force. If there is difference, apply the velocity calculated by equation (4) to regulate contact force. This process is repeated until robot reaches the end point.

4 Experiment

4.1 Environment

Experiment environment is shown in Figure [7]. There are 4 façades which are 1.8m height and 1.2m width. Two built-in guide rails are placed on top and bottom of facades. Wheels of robot is fixed to facades and robot moves horizontally with rotating wheels on the top. Horizontal velocity of robot was fixed to 5cm/s and contact was maintained on one façade.

4.2 Experiment Strategy

In this experiment, consistent y-axis force which is 2kgf was applied with two angles of squeezing module. Coefficient of equation (4) was 1.25mm/s/kgf. Angles were adjusted to make contact angle between rubber and façade 35, 45degrees. Data from load cells were recorded to get the information about friction force and contact force applied to wall. Linear velocity of squeezing module was also recorded to judge how force control algorithm works. Sampling period of data was 120milli seconds. Left picture of Figure [8] shows the process of experiment and right picture presents the difference between contact area and no contact area.
after process. It can be seen that water was removed well by contact of module.

### 4.3 Results

Figure [9] presents experiment result data when the contact angle is 45 degrees. Sections are separated into three which are approaching, contact, and force control. During the approaching section, squeezing module approaches to façade with limited velocity 2.5mm/s and there is no contact force. After that, contact begins in contact section and y-axis force increases. Horizontal robot does not move in these two sections according to algorithm because there is not enough contact. When y-axis force satisfies 2kgf, robot moves to horizontally and this is the beginning of force control. With slight oscillation, force regulation is well performed for 24 seconds in force control section. However, there is steep decrease of y-axis force at the start of horizontal moving. The reason for this is that friction force pushes squeezing module and it leads to sudden contact angle change. Temporary angle change structurally makes the

---

**Figure 9.** Experiment Results when 45 degrees

**Figure 10.** Experiment Results when 35 degrees
loss of contact. After contact loss, it can be seen from the data that angle modification and force control algorithm restore the contact force. Average of x-axis force is 2.78kgf. Figure [10] shows the data when contact angle is 35degrees. Flow of force data is similar with 35degrees. There is also steep decrease when robot starts to move horizontally. Y-axis force is well regulated to 2kgf during force control section. However, x-axis force is different with that of 45degrees. Average x-axis force is 3.14kgf and it is higher than that of 45degrees. It is because contact area of rubber and façade increases when angle becomes smaller. Increase of contact area leads to more friction force and x-axis data shows relationship with the friction force and contact angle. This system has independent energy sources. Thus, energy conservation is important to operate longer time. In this respect, 45degrees contact angle has better energy efficiency about 11.5% than 35degrees because that amount of torque is more needed to withstand friction force in 35degrees and results of both cleaning processes are well done which can be seen in Figure [8].

5 Conclusion

This paper deals with structure of tool system of building maintenance, flow control algorithm of water injection and velocity based force control algorithm to control that. Experiment of force control and results are shown and analyzed. Results show that velocity based force control well performs the force regulation between squeezing module and façade. This study will contribute to development of cleaning system of wall with contact.

Further study is convergence control of tool system when robot moves with variable horizontal velocity and irregular surface. Improvement of force control algorithm to solve the steep decrease of contact problem when robot starts to move horizontally will also be studied.

Acknowledgment

The work presented in this paper was funded by BMRC(Building-Façade Maintenance Robot Research Center), supported by Korea Agency for Infrastructure Technology Advancement(KAIA) under the Ministry of Land, Infrastructure and Transport(MOLIT).

This work was supported by the Human Resources Program in Energy Technology of the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea. (No. 20124010203250)

References


Towards Autonomous Robotic In-Situ Assembly on Unstructured Construction Sites Using Monocular Vision

C. Feng\textsuperscript{a}, Y. Xiao\textsuperscript{a}, A. Willette\textsuperscript{b}, W. McGee\textsuperscript{b}, and V.R. Kamat\textsuperscript{a}

\textsuperscript{a}Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, USA

\textsuperscript{b}College of Architecture and Urban Planning, University of Michigan, Ann Arbor, USA

E-mail: cforrest@umich.edu, yongxiao@umich.edu, willetta@umich.edu, wesmcgee@umich.edu, vkamat@umich.edu

Abstract -

Unlike robotics in the manufacturing industry, on-site construction robotics has to consider and address two unique challenges: 1) the rugged, evolving, and unstructured environment of typical work sites; 2) the reversed spatial relationship between the product and the manipulator, i.e. the manipulator has to travel to and localize itself at the work face, rather than a partially complete product arriving at an anchored manipulator. The presented research designed and implemented algorithms that address these challenges and enable autonomous robotic assembly of freeform modular structures on construction sites. Building on the authors’ previous work in computer-vision-based pose estimation, the designed algorithms enable a mobile robotic manipulator to: 1) autonomously identify and grasp prismatic building components (e.g., bricks, blocks) that are typically non-unique and arbitrarily stored on-site; and 2) assemble these components into pre-designed modular structures. The algorithms use a single camera and a visual marker-based metrology to rapidly establish local reference frames and to detect staged building components. Based on the design of the structure being assembled, the algorithms automatically determine the assembly sequence. Implemented using a 7-axis KUKA KR100 robotic manipulator, the presented robotic system has successfully assembled various structures autonomously as shown in Figure 1, demonstrating the designed algorithms’ effectiveness in autonomous on-site construction robotics applications.

Keywords -

On-Site Construction Robotics; Autonomous Assembly; Modular Construction; Pose Estimation

1 Introduction

Several studies have argued that among all industries, construction has seen a significant productivity decrease over the last several decades compared to other industries [1]. Construction has also been documented to have some of the highest rates of workspace injuries and fatalities [2]. Automation and robotics in construction (ARC) has the potential to relieve human workers from repetitive and dangerous tasks, and has been extensively promoted in the literature as a means of improving construction productivity and safety [3].

Figure 1 Curved wall assembled on-site by the designed autonomous robotic system

Compared to the tangible benefits of automation and robotics identified by the manufacturing industry, the construction industry is still exploring feasible and broadly deployable ARC applications [3]. This can be attributed to several commercial and technical challenges. From the commercial perspective, the fragmented and risk-averse nature of the construction industry leads to little investment in ARC research causing construction to lag behind other industries [4]. On the other hand, as described next, there are several technical complexities inherent in construction that have contributed to hindering the successful development and widespread use of field construction robots.

1) Unstructured Construction Environments

Automated and robotized manufacturing facilities are typically considered as structured environments, since both the machines and evolving products either stay in their predefined locations or move on predesigned and typically fixed paths. In general, such
environments do not change shape or configuration during the performance of manufacturing tasks, making the enforcement of tight tolerances possible [5]. In contrast, construction sites can typically be considered unstructured since they are constantly evolving, and dramatically changing shape and form in response to construction tasks. Building components are moved around without fixed paths or laydown/staging areas. Various physical connections are established through improvisation in response to in-situ conditions, making tight tolerances hard to maintain [6].

2) Mobility of Construction Manipulators

In manufacturing, factory robotics typically involves robotic platforms that are generally stationary (or have limited linear mobility) and partially complete products that arrive at robot workstations and precisely localize themselves in the robots’ reference frames. Precision is achieved by controlling the pose of the moving (and evolving) product, and the robots themselves are programmed to manipulate the products through fixed trajectories. Thus, from a mobility and cognitive perspective, a factory robot has little responsibility and autonomy. Control is achieved by enforcing tight tolerances in moving and securing the product in the manipulator’s vicinity. However, this spatial relationship is reversed in construction. A construction robot has to travel to its next workface (or be manually set up there), perceive its environment, account for the lack of tight tolerances, and then perform manipulation activities in that environment. This places a significant mobility and cognitive burden on a robot intended for construction tasks even if the task itself is repetitive.

This discussion highlights that factory-style automation on construction sites requires development of robots that are significantly more mobile and perceptive when compared to typical industrial robots. Such on-site construction robots have to be able to semantically sense and adjust to their unstructured surroundings and the resulting loose tolerances. This paper proposes a new high-accuracy 3D machine vision metrology for mobile construction robots. The developed method uses fiducial markers to rapidly establish a local high-accuracy control environment for autonomous robot manipulation on construction sites. Using this method, it is possible to rapidly convert a portion of a large unstructured environment into a high-accuracy, controllable reference frame that can allow a robot to operate autonomously (Figure 1).

The rest of the paper is organized as follows: Related work is reviewed in section 2. The authors’ technical approach is discussed next in detail in section 3. The experimental results are shown in section 4. Finally, in section 5, the conclusions are drawn and the authors’ future work is summarized.

2 Previous Work

The construction community has pursued ARC research for several decades. Various robotic platforms were prototyped focusing on specific construction activities (e.g., interior finishing robot [7], infrastructure inspection and maintenance robot [8], assembly robot [9], masonry robot [10]). During this research, it was realized that increasing the level of autonomy for construction robots requires high accuracy localization of the robot: from 3-5 cm indoor positional accuracy for contactless construction tasks such as spray-painting, to 2-3 mm accuracy for more precise tasks demanding direct contact between manipulator and building components [11]. This requirement has posed a significant challenge for ARC because even by using current state-of-the-art simultaneous localization and mapping (SLAM) techniques, such accuracy is hard to achieve at large scales [12]. In order to address this issue, the authors chose to use computer-vision-based pose estimation algorithms that can achieve high accuracy locally around a visual marker [13, 14].

Recently the architectural design community has also shown an increased interest in industrial robotics, with many academic programs investing in their own robotic work cells¹, leveraging as development platforms for the exploration and refinement of novel production techniques in which material behaviour is intrinsically linked to fabrication and assembly logic. As part of the general ecosystem of industrial robotics, computer vision systems have begun to play an increasingly important role in these research initiatives.

Initially the majority of architectural robotic research utilizing computer vision has revolved around its application at the micro scale, using vision feedback systems to make incremental adjustments to a robotic strategy based upon local variations [15, 16]. While beneficial as a means to adjust for material variation and machine error, these implementations are not robust enough for the in-situ robotics due to the complexities of construction sites.

Recently, architects have also begun to explore macro scale computer vision applications. At the forefront, an eight-meter-long module wall was assembled by an ABB robot along a gestural path captured by its vision system in the ECHORD project [17]. The mobile robot also used the same system to reposition itself on the construction site and make local adjustments based on topographic variation. Without using an extensive sensor suite like the one used in ECHORD, the robotic platform described in this paper successfully built similar module walls purely based on perceived information from a single camera.

---

3 Technical Approach

3.1 System Overview

The developed robotic in-situ assembly system consists of 7 major components, as shown in Figure 2. The workflow of the system consists of an offline design process and an online building process. During the offline process, a designer models the intended structure in 3D, which is then analyzed and validated by the automatic plan generator, outputting a building plan for the online process. The online building process uses a fixed camera mounted on the base of the robot for providing images to the pose estimator to detect staged building components and estimate their pose, and more importantly localize the robot itself in the local building reference frame. Having computed this information, the plan achiever then sequentially transforms each step from an automatically generated assembly plan into an executable command, which can be interpreted by the robot controller and subsequently executed by the robot manipulator. In addition, the visualization component also receives the information generated by the pose estimator as well as the robot’s real-time pose feedback from the controller, to simultaneously represent the actual on-site assembly process into a 3D virtual environment for improved monitoring.

3.2 Calibration of Pose Estimator

Before introducing the details of other components of the system, it is important to discuss how the pose estimator is calibrated, since this is crucial to the level of accuracy that the system can achieve. This process includes two steps: intrinsic and extrinsic calibration of the camera. Intrinsic calibration involves estimating the camera’s focal length, principle point’s position on the image plane, and distortion parameters. On the other hand, extrinsic calibration aims to determine the relative 6-DOF pose of the camera in the robot’s coordinate frame. It must be noted that these two calibrations are a one-time process, as long as the camera is fixed-focus and rigidly mounted in the robot’s coordinate frame (e.g., fixed installation on the robot’s base).

3.2.1 Intrinsic Parameters

Unlike the popular plane-based camera calibration method [18] implemented in OpenCV\(^2\), Matlab Calibration Toolbox\(^3\), the authors chose to calibrate the camera using a 3D rig, which is similar to the classic calibration in photogrammetry. This 3D rig was made by attaching 18 Apriltags\(^4\) on two intersecting planes forming a 90 degree angle (as shown in the top of Figure 3) so that the 3D coordinate \(X\) of each Apriltag’s center could be readily measured.

The process of calibration was then simply taking a sequence of \(N\) images of the rig and inputting them into the author-developed camera calibration tool\(^5\), which takes advantage of the Apriltag detection algorithm to detect the 2D image coordinate \(U\) of each tag center and establish correspondence with \(X\). Then the initial camera intrinsic and extrinsic parameters can be obtained through Direct Linear Transform (DLT) \(^6\) and subsequently optimized by bundle adjustment \(^7\).

Benefitting from the high corner detection accuracy of Apriltag as well as the 3D rig, this intrinsic calibration results in a higher level of accuracy of the pose estimation process.

---


\(^{3}\) [http://www.vision.caltech.edu/bouguetj/calib_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)

\(^{4}\) Available at [https://code.google.com/p/cv2cg/#apriltag](https://code.google.com/p/cv2cg/#apriltag)
calibration produces more robust and repeatable results than the alternatives mentioned above.

### 3.2.2 Extrinsic Parameters

Once the camera’s intrinsic parameters are calibrated, the camera’s relative pose in the robot’s coordinate frame, $T_{rc}$, can be estimated using an extrinsic calibration marker containing 4 Apriltags with known size and spacing. As shown in Figure 4, $T_{rc}$ can be composed from the two other poses, $T_{mr}$, the robot’s pose in the extrinsic calibration marker’s coordinate frame, and $T_{cm}$, the marker’s pose in the camera coordinate frame, by $T_{rc} = (T_{cm}^{-1}T_{mr})^{-1}$.

This extrinsic calibration process consists of the following steps:

1. Fix the marker in the camera’s field of view;
2. Manually control the robot manipulator to pinpoint at least 4 non-collinear points on the marker and record their 3D coordinates $X$ in the robot’s coordinate frame; also measure their local 3D coordinates $X_m$ in the marker’s reference frame (by setting all Z coordinates to be zero);
3. Take an image from the camera and detect the 4 Apriltags’ 16 corners’ 2D image coordinates $U$.

With this information collected, the $T_{cm}$ can be estimated using the well-known rigid body registration method from 3D point set $X$ to $X_m$, while the $T_{mr}$ can be estimated by decomposing the homography between $X_m$ and $U$ using the previously calibrated camera intrinsic parameter $K$ [18, 13].

In order to improve the extrinsic calibration’s accuracy, a non-linear optimization of $T_{cm}$ is also performed in addition, since during the homography decomposition, a polar decomposition is performed to get a valid rotation matrix $R_m$, which causes the result to be non-optimal. This optimization, as shown in equation (1), can be done by tuning the initial $T_{cm}$ to minimize the re-projection error:

$$
\arg\min_{T_{cm}, t_m} \sum_{j=1}^{16} \|U_j - K (R_m X_j + t_m)\|^2
$$

### 3.2.3 Calibration Validation

The calibration can be validated by the following procedure:

1. Fix the extrinsic calibration marker to a new pose that is different from the one used in the calibration;
2. Measure, in robot frame, the 3D coordinates $X$ of a set of M corner points on the marker (e.g., the 16 corners used previously);
3. Take an image and find out the corresponding 3D coordinates $X_c$ in the camera reference frame using Apriltags;
4. Calculate the residual using equation (2):

$$
\sqrt{1 \over 2M} \sum_{j=1}^{M} \|X_j - (R_c X_j + t_c)\|^2
$$

This residual should ideally approach zero, as smaller residual indicates better calibration accuracy.

### 3.3 Automatic Plan Generation

#### 3.3.1 Algorithmic Architectural Design

Starting with the introduction of Ivan Sutherland’s Sketchpad at MIT in the 1963, the architectural exploration of computation has focused on the digital environment’s ability to represent an object as a system “compromised of and working with a series of interrelated systems” [22], a surprising contrast to the discrete geometric representations found in many 2D and 3D CAD applications. Initially as a domain of specialized research groups embedded in academia or commercial practices, this systems approach to digital design has become increasingly commonplace. In the 1990s many architects, unsatisfied with the capabilities presented by off-the-self software, began to develop their own software solutions through both higher-level scripting languages for CAD packages and ground-up application development. Currently, visual programming interfaces that afford designers quick
access to the potential of computation without the effort of coding syntax are available as plug-ins for popular commercial software, such as Dynamo for Autodesk’s Revit and Grasshopper for McNeel’s Rhinoceros.

The authors implemented this systems approach (and its respective tools) to automatically derive the robotic positioning data for each building block in a curved stack-unit wall from a single user-generated non-uniform rational basis spline (NURBS) surface. Working from a predetermined block size, an algorithm developed in the Grasshopper plug-in for Rhinoceros extracts latitudinal section curves from the input surface, generating a running bond pattern of the said blocks. Each block is checked for volumetric collisions with adjacent blocks, color-coding collisions in the Rhinoceros environment. Instead of applying simple heuristics to arbitrarily resolve these collisions, the color-coding can provide real-time feedback, engaging the designer to actively participate in the development of the curved stack-unit wall. Simultaneously the necessary positional information for each block is output for the generation of the building plan. Combined into a single algorithmic process, these functions “enable an explicit and bidirectional traversal of the modern division between design and making” [23], reinforcing the implications of William Mitchell’s statement that “architects tend to draw what they can build, and build what they can draw” [24].

3.3.2 Building Plan Generation and Simulation

Given the final positions and orientations of all the building blocks in the design, the building plan is generated and written into a text file stored for the plan achiever to process later during the online building phase.

This building plan file contains a list of sequential instructions for the robot manipulator to build the designed structure. Each line in the file corresponds to an instruction. For example, the following plan file segment will instruct the robot manipulator to first grab a building component named “block0” directly from above (line 1-4), then lift it vertically up for 500 mm (line 5) and finally place it at its destination in another reference frame named “building” (line 6-8):

\[
\begin{align*}
\text{Gripper 0} \\
\text{Goto block0 0 0 500 0 0 0} \\
\text{Goto block0 -12 -10 -10 0 0 0} \\
\text{Gripper 1} \\
\text{Shift 0 0 500 0 0 0} \\
\text{Goto building 200.00 -300.00 500.00 -63.92 0.00 0.00} \\
\text{Goto building 200.00 -300.00 19.05 -63.92 0.00 0.00} \\
\text{Gripper 0}
\end{align*}
\]

Currently, 3 types of instructions are implemented in the system:

1) **Gripper 0/1**
   Control the manipulator’s gripper to open (0) or close (1);

2) **Goto reference_frame x y z a b c**
   Control the manipulator to move to a new pose (x, y, z, a, b, c) in the reference frame, in which the (x, y, z) is the new position and (a, b, c) specifies the new orientation as three Euler angles in “ZYX” order;

3) **Shift x y z a b c**
   Control the manipulator to incrementally move by (x, y, z, a, b, c).

Figure 5 Building plan simulation

This building plan can also be simulated in Rhinoceros to check if there exists any self-collision between the robot manipulator and the wall during the building process, as shown in Figure 5.

3.4 Vision-based Plan Achiever

3.4.1 Rapid Setup of Building Reference Frame

As previously mentioned, the reversed spatial relationship of product and manipulator on construction sites poses a significant challenge for autonomous mobile robots. This is notably different from typical autonomous manufacturing spatial configurations, where robots’ bases are either stationary or have finite mobility, and materials/components can be readily staged at fixed locations within the manipulators’ static workspaces. In contract, for mobile robots to autonomously perform building tasks on unstructured construction sites, their bases require significant mobility, and consequently their manipulators’ workspaces are not fixed with respect to the construction site. In order to complete building tasks at the correct locations and assemble materials into their intended poses, a robotic system must be able to establish the accurate 6-DOF transformation between the robot’s base and the building reference frame at all times. As pointed out in [11], this requires the localization accuracy to be at least at centimeter level, which is far from achievable using state-of-the-art SLAM style techniques for mobile robots.

---

5 http://autodeskvasari.com/dynamo
6 http://www.grasshopper3d.com/
In order to address this challenge, the authors propose a convenient and accurate solution using planar marker-based pose estimation \cite{14, 13}, as shown in Figure 6. By 1) attaching fiducial markers at appropriate locations on-site where building tasks are to be performed, 2) surveying their poses $m^bT_b$ in the building reference frame using a total station, and 3) storing these poses inside the system’s database, a mobile robot can readily estimate its base’s pose $b^rT_r$ inside the building reference frame using equation (3) whenever its on-board camera detects such a marker, based on previous calibration results:

$$b^rT_r = (t^cT_c^mT_m^b)^{-1}$$

### 3.4.2 Conversion from Plans to Commands

With the information input from the pose estimator, the vision-based plan achiever starts to execute the building plan generated beforehand, according to the following procedure:

1. Read a single plan step (i.e. one line) from the building plan file;
2. Wait until all the poses needed to convert this step into a building command become available;
3. Convert this step into a command that is executable by the robot’s base $b^rT_r$ inside the building reference frame using equation (3) whenever its on-board camera detects such a marker, based on previous calibration results;
4. Send the command to the robot controller;
5. Wait for the controller to complete the command;
6. Repeat this process unless all plan steps are completed, i.e. the plan is achieved.

It must be noted that the core step of this procedure is the conversion from a plan step to an executable command. This is because the poses stored in the previously generated building plan are not completely specified in the robot’s base reference frame. Recall that every pose in the “Goto” step is specified in a “reference_frame” relatively. Specifying all steps in the building plan in the robot base reference frame is not possible because: a) during the design phase the designer conceives all component locations in the building reference frame; b) the robot’s base is expected to be mobile during the building phase; and c) more importantly, the building components will be arbitrarily transported and staged in the building reference frame in the vicinity of the robot manipulator’s workspace. This, in fact, is one of the core differences between on-site construction automation and manufacturing automation.

In the authors’ approach, this conversion is facilitated by the aforementioned rapid setup of the building reference frame using markers. As long as the pose estimator can detect and report the transformation between the on-board camera and the marker used to specify the building reference frame, the poses in the plan steps can be readily converted to the robot base frame using equations similar to (3). Similarly, by attaching markers on the building components, the robot manipulator can detect and clasp them autonomously after the corresponding plan steps are converted.

### 4 Experimental Results

#### 4.1 Design and Goal

The authors implemented the designed algorithms into a robotic system using a 7-axis KUKA KR100 robotic arm, a Point Grey Firefly MV camera, and a laptop with an Intel i7 CPU, connected through the Robot Operating System (ROS)\footnote{http://www.ros.org/}. Each component in Figure 2 except for the plan generator and robot manipulator is a process corresponding to a ROS node. The camera node sends images of size 1280 pixels by 960 pixels to the pose estimator that implements the AprilTag detection algorithm in C/C++. The plan achiever was implemented in Python. The robot controller was also developed using Python to send and receive control signals via Ethernet through KUKA’s native Robot Sensor Interface (RSI). Inside this controller, a 6-DOF PID control algorithm was employed to drive the robot manipulator (with a two-finger gripper) to its destination pose when executing commands from the plan achiever. The involved inverse kinematics computations are performed inside the KUKA manipulator’s controlling middleware.

In the first phase of experiments, the robot was tasked with assembling a section of a curved wall, as designed in section 3.3.1. The design was shown in Figure 2. The building components used were a set of 170x70x20mm\textsuperscript{3} medium-density fibreboard (MDF) blocks, each affixed with two different 56x56mm\textsuperscript{2} Apriltags. The building reference frame was setup by 4 different 276x276mm\textsuperscript{2} Apriltags. The overall goal of the experiment was to test the robot’s ability to autonomously build the designed wall.
4.2 Results and Discussion

The system was first calibrated and validated using the methods discussed earlier. The validation residual calculated using equation (2) was found to be less than 1mm. With the accurately calibrated intrinsic and extrinsic parameters, the online building process proceeded smoothly. The building blocks were affixed with smaller markers, which decreased the building block localization accuracy to centimetre level (2cm) during the clasping process. The error was however compensated by the tolerance of the gripper.

A working cycle of the autonomous building process was shown in Figure 7(a)-(d). Since the pose estimator was constantly monitoring and updating the poses of each marker, the system was naturally capable of automatically adapting to pose changes on-site. As shown in Figure 7(e)-(h), when a building block’s pose changed, the robot manipulator was automatically able to pick it up at its newest location. A video recording of the experiment can be found online at the following URL: http://youtu.be/fj7AXRpj97o. A fully assembled three-layer curved wall approximately 1.5m in length was previously shown in Figure 1. Although not closely related to the core contribution of this paper, it’s worth noting that the block detection and grasp can also be achieved without markers using Kinect sensor and the author’s newly developed fast plane extraction algorithm [25], as shown in Figure 8 and demonstrated at the following URL: http://youtu.be/CyX4Pr_xly0.

5 Conclusion and Future Work

This paper reported algorithms and an implemented robotic system that is able to automatically generate building plans from computational architectural designs and achieve these plans autonomously on construction sites. In order to address the localization accuracy challenge, the authors proposed a computer-vision-based sub-centimetre-level metrology that enables pose estimation using planar markers. The conducted evaluation experiments used the designed robotic system to autonomously build a curved wall of MDF blocks, proving the algorithms and the system’s ability to meet the accuracy requirement when building computational architecture designs.

The authors’ current and planned work in this research direction is focused on continuously improving this system in aspects such as 3D perception for efficient object grasping, autonomous navigation for implementation on mobile platform, and improved and stable control algorithms.
References


Automated Dynamic Management of Road Construction Sites

R. Heikkilä\textsuperscript{a}, E. Viljamaa\textsuperscript{b}, A. Kaaranka\textsuperscript{a}, T. Makkonen\textsuperscript{a} and I. Peltomaa\textsuperscript{b}

\textsuperscript{a}Construction Technology Research Center, University of Oulu, Finland
\textsuperscript{b}VTT Technical Research Centre of Finland, Finland

E-mail: rauno.heikkila@oulu.fi, esa.viljamaa@vtt.fi, annemari.kaaranka@oulu.fi, tomi.makkonen@oulu.fi, irina.peltomaa@vtt.fi

Abstract - Automated dynamic management of road construction sites has been studied within the DigiINFRA research project in Finland. A review of currently available commercial solution, including most popular commercial softwares and systems in global markets (Topcon Sitelink3D, Trimble applications, Novatron Xsite, Leica iCONTelematics, Infrakit), is drawn. Test results of the last campus site experiments using the special DSCC (Dynamic Site Control Center) prototype system developed in the DigiINFRA research project are presented and analysed. Some site observations of the use of the commercial site control systems are presented as well. The special demands set by the dynamic site management to the development of general information modelling or national infra BIM specifications and the related open information transfer standards are introduced and evaluated.

Keywords - Automation, Road Construction, Dynamic Management, Dynamic Control

1 Introduction

Road and railway construction is mainly earth moving work, to which automation and robotics apply very well. The automation of earth moving is primarily information model based automation. Large utilization of information modelling and automation in infra sector is possible as long as the common ground rules and specifications have been created for the industry. This is one of the main aims of the active Infra FINBIM \cite{1} research in Finland.

The design information needed for construction of roads and railways is typically both place-based and time-bound. Main challenge is to perceive the linking between different type of information (place, time, task, planned schedule, resources, planned geometry of structures, material, etc.). All the design information, such as schedules or the surface models of structures, form a static information of the project, which has to be connected real-time with the information captured from site. The dynamic information gathered from site includes information about the places of machines and workmen, different statuses, work tasks, working conditions and the progression of the tasks. From these information different performance quantities can be calculated describing the function of separate work tasks as well as the whole construction process on site. Different information can be illustrated and shared case by case using suitable specified user-interfaces and licenses according to the profiles of workmen.

2 Dynamic Management of Road Construction Site – Experiments with the DSCC Prototype System

In this study, dynamic management was determined to mean flexible, fast and efficient reaction to control of different working tasks on site. In addition, dynamic management was determined to be such a system that is information modeling (BIM) based, internet browser-based, independent of different terminals, and utilizes wireless and mobile information transfer. The most important part sectors of the dynamic management in road construction are the tracking and control of schedule, costs, materials, work staff, machinery, sub contractors and site traffic.

University of Oulu and VTT Technical Research Centre of Finland studied the dynamic control of road construction site in the research project “Dynamic management of digital product process in a dynamic co-operation network (DigiINFRA)”. In the research, the functions of road construction management were cleared and prioritized by interviewing site foremen. In addition a special base for application integration and different new types of information models were developed. \cite{2}\cite{3}\cite{4}\cite{5} For the experimenting, the Dynamic Site Control System (DSCC) was implemented \cite{6}. Figure 1 presents the information sources and different visualization alternatives of DSCC prototype system.
The DSCC system was tested on the common campus test site of University of Oulu and VTT. In the tests, an excavator equipped by the VisionLink 3D machine control system of Novatron Oy and three cars representing dump trucks were used. DSCC offers different specified web-based process views, which are based on the integrated design and as-built information. For example, this kind of process views can be map or status views as well as different design or as-built views of project tasks. In Figure 2 there is a picture of construction site’s control center build in a room of University of Oulu’s premises in order to control the campus tests. In the control center the DSCC user-interfaces were used for the dynamic control of the earth-moving works on the campus test site. Figure 3 gives an example of DSCC system’s real-time schedule control view [6]. As the reference systems in the experiment, the Kuura system of Hohtolabs Oy [7] and Xsite system of Novatron Oy [8] were used.

In the campus tests, the design and modelling of road part was done by Tekla Civil software [9]. The alignment information and material quantities were transferred and utilized in the DSCC tests. The resource and schedule planning was done using PlaNet software [10]. The as-built information in the tests was produced by a special web-based material transfer application as well as using the VisionLink 3D machine control system. The design and as-built information were integrated and saved to the semantic data base used. In the experiments, the dynamic track and control was tested six work days. The passage of time was simulated, but other functions were realized on site.

3 Examining Commercial Software and Systems

The Infrakit system of Infrakit Oy is a web-based information management system for infra construction sites. Design data and different production files can be shared to site personnel and work machines. Also the quality of uploaded files can be checked automatically. Latest production models and alignment lines can bee seen by computers and tablets. The Infrakit tablet application shows the models with the digital terrain model and alignment lines. Locations of work machines are shown on the map as well. Some of the commercial machine control systems have real-time information transfer with the Infrakit system. A special as-built data measurement application is included having the feature to automatically calculate the deviations between designed and measured coordinates. Also photos and PDF files can be saved to the Infrakit data bank.
Topcon corporation has introduced and publicated new SiteLink3D [11] systems, which integrates very broadly all of the different information models needed for construction operatives. The information exchange is made real-time. Moving machines are shown on maps and in 3D information models, the surfaces measured by 3D machine control systems are updated automatically to data bank and computer views as well. There is also an integration with Dynaroad [12] mass hauling and schedule optimization design software.

Tekla Civil (Trimble Corporation) has been complemented with a new Field Mode application for site use purposes by tablets. The application can upload all the information needed directly from Tekla Civil data bank. Different information models can be examine three-dimensionally. The information transfer is done using Trimble Connected Community cloud serve application. Trimble Vision Link is the system for dynamic site control purposes, by which the progression of work tasks can be followed as well as quality control or as-built measurement data collected. It is worth noting that in this solution the geometric surveys of final road structural layers are done using 3D machine control system of compacting machines.

The main available functions of the Xsite system provided by Novatron Oy are the sharing of production models and as-built information, document searching and uploading, fleet management, status reports for each work machine used and instant messaging possibilities.
The new Xsite system developed by Novatron Oy offers functionalities for the sharing of production and as-built models and the searching and uploading of positioned documents. Machine positions can be shown on maps, different activity reports of machines can be produced, there is available remote access to operators and a special instant messaging tool.

Leica iCONTelematics includes versatile applications for remote site management, data transfer in both directions from the field to office, fleet management and reporting by a special Track tool.

Leica iCONTelematics includes versatile applications for remote site management, data transfer in both directions from the field to office, fleet management and reporting by a special Track tool.

Figure 9. The View tool of Leica iCONTelematics.

4 Conclusion

At the same time when infra industry is developing common Infra BIM guidelines, the next change and shaking is already coming from the side of technology companies. Now it could be possible to start the utilization of dynamic site control systems. DigiINFRA research project studied more extensively the challenges and potentials of dynamic management, of which some parts were tested in the campus test site environment using the programmed prototype system. The further development of this system is a real challenge and offers valuable means and potentials for the near future. The dynamic information integration includes very many possibilities, of which the commercial systems presented are already introducing. The importance of the development of open information transfer formats is amplified. In large construction sites, typically many of the different commercial systems are to be used at the same time. Therefore, open information transfer formats and standards needs to be extended to the level enabling dynamic information transfer.

References


19/02/2014.

A New Hysteretic Model for Magnetorheological Elastomer Base Isolator and Parameter Identification Based on Modified Artificial Fish Swarm Algorithm

Y. Yu, Y. Li and J. Li

Centre for Built Infrastructure Research, School of Civil and Environmental Engineering, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia
E-mail: yang.yu@uts.edu.au, yancheng.li@uts.edu.au and jianchun.li@uts.edu.au

Abstract -
Magnetorheological elastomer (MRE) base isolator is a new semi-active control device that has recently acquired more attention. This paper proposes a new model for MRE base isolator to portray the nonlinear hysteresis between generated force and the displacement. In this model, a hyperbolic expression is proposed to compare with the classical Bouc-Wen model, which includes internal dynamics represented by a nonlinear differential equation. For the identification of model parameters, a modified artificial fish swarm algorithm is adopted using the experimental force-displacement/velocity data under different testing conditions. In this algorithm, a self-adaptive method for adjusting the algorithm parameters is introduced to improve the result accuracy. Besides, the behaviours in the algorithm are simplified to descend the algorithmic complexity. Parameter identification results are included to demonstrate the accuracy of the model and the effectiveness of the identification algorithm.

Keywords -
Base isolation; Magnetorheological elastomer; Parameter identification; Artificial fish swarm

1 Introduction

Base isolation is the most commonly used technique for the seismic protection of civil structures [1-3]. When the earthquake occurs, the base isolation device will deflect external vibrations by isolating destructive frequency contents from transmitting into the main structure above thus keep the integrity and safety of the structure and contents therein. However, present base isolation devices are mainly passive and only effective in a narrow band of frequency range. Especially, the vulnerability for present base isolation practice during far-source and near-source earthquakes becomes a great concern. Recently, a novel adaptive base isolator based on the magnetorheological elastomer (MRE) has been put forward and realized [1]. The proposed MRE base isolator has the ability to adaptively protect base isolated structures in real-time against various types of earthquakes and outperforms the traditional seismic isolator in terms of effectiveness and functionality for the vibration protection [4-5].

MRE base isolator is a smart device with distinctive nonlinear and hysteretic behaviour, which may hinder its application in structural control. Therefore, it is of great importance that a traceable model is available before any realizable controller can be designed. So far, the research on modelling the nonlinear behaviour of MRE base isolator is relatively limited. The main model for portraying the hysteresis is the Bouc-Wen model [6-7]. However, due to a large number of parameters and the nonlinear equations in the model, the identification of mode parameters becomes very complicated and thus requires great computational resources. Although this problem can be solved by the direct search algorithm, it relies on the good choice of initial values to a large extent and may lead to the premature convergence. Recently, artificial fish swarm algorithm (AFSA), as a novel artificial intelligence approach, is developed for dealing with the complex problems which are difficult to be solved by other methods [8]. And it has been successfully utilized in the application of model parameter identification. Nevertheless, in virtue of the randomness of parameters and the random behaviour in basic AFSA, the algorithm has the slow convergence rate and is easy to fall into the local optimum, thus the solution accuracy is difficult to improve.

In this paper, a new computational-efficient model for MRE base isolator is proposed. This model employs a hyperbolic sine function combined with conventional viscous damping and spring stiffness to depict the hysteretic behaviour. An optimization approach based on AFSA is designed to identify the model parameters using the experimental data acquired form a practical MRE base isolator. In order to improve the result accuracy, a self-adaptive mechanism is introduced to
update the algorithm parameters. Moreover, the behaviours in the algorithm are simplified to reduce the computational complexity in the identification process. These modifications guarantee the algorithm to have the fast convergence rate as well as the high recognition accuracy.

The reminder of the paper is organized as follows. The Bouc-Wen model and the proposed model are described in Section 2 with the statement of the optimization problem. In Section 3, the modified AFSA, as applied in identifying the proposed model parameters, is developed and its benefits are highlighted. Parameter identification results are given in Section 4 together with some analysis. Finally, Section 5 draws the conclusion.

2 System Model and Problem Statement

The Bouc-Wen model is common-used model for describing the hysteresis response of MRE base isolators with some constraints. Accordingly, a novel model is designed in this part and an optimization problem is also presented for solving the model parameters.

2.1 Bouc-Wen Model

This model incorporates a Bouc-Wen component, which regenerates hysteresis loops, in parallel with a Voigt element, which depicts solid material behaviours [7]. The model can be represented by the force equation and the associated hysteretic variable, given by:

\[ F = \alpha k_0 x + (1 - \alpha)k_0 \dot{z} + c_0 \dot{x} \]  

\[ \dot{z} = \dot{x} - \beta |\dot{x}|^{n-1} z - \gamma |\dot{x}|^n \]  

Where \( k_0 \) denotes the stiffness of the spring; \( c_0 \) is the viscous coefficient indicating the damping capacity of the system; \( \alpha, A, \beta, \gamma \) and \( n \) are non-dimensional parameters which are responsible for the shape and size of the hysteresis loops; \( z \) is the hysteretic variable that represents a function of the time history of the displacement. The Bouc-Wen model is widely used in structure engineering and MR behaviour for its mathematical ability to represent a large class of hysteretic behaviour. However, because of the incorporation of internal dynamics in regard to the hysteretic variable \( z \), undesirable singularities may appear in the process of model identification.

2.2 Proposed Model

Compared with relatively complicated Bouc-Wen model, a simple model is presented in this work to model the nonlinear force-displacement characteristic of the MRE base isolator. A component-wise added approach is adopted which includes the viscous damper, spring stiffness and a hysteretic component. The structure of the proposed model is shown in Figure 1.

![Figure 1. The proposed model](image)

In terms of mathematical expressions, the model utilizes a hyperbolic sine function to describe the hysteresis and linear functions to describe the viscous and stiffness. The model expression is given as:

\[ F = c_0 \dot{x} + k_0 x + \alpha \dot{z} + F_0 \]  

\[ z = \sinh(\beta x) \]  

Where \( c_0 \) and \( k_0 \) are the viscous and stiffness coefficients; \( \alpha \) is the scale factor of the hysteresis; \( z \) denotes the hysteretic variable given by the hyperbolic sine function; \( \beta \) is the scale factor of the isolator displacement defining the hysteretic slope; \( F_0 \) is the isolator force offset.

2.3 Problem Statement

As the parameters of the proposed model are difficult to search by trials, a minimization optimization is employed to solve the problem. The critical point of the optimization is the choice of the fitness function, which has an important influence on the identification results. In this work, the fitness function \( H(X) \) is defined as follows:

\[ H(X) = \frac{1}{N_v} \sum_{i=1}^{N_v} [F_i - (c_0 \dot{x}_i + k_0 x_i + \alpha \sinh(\beta x_i) + F_0)]^2 \]  

Where \( X = [c_0, k_0, \alpha, \beta, F_0]; N_v \) is the total number of the experimental data. \( F_i, x_i \) and \( \dot{x}_i \) denote the force, displacement and velocity at the \( i \)th sampling time, respectively. If the fitness value is very close to zero,
the identification result \( X \) is regarded as the optimal solution. As a consequence, the minimum optimization problem with constraints can be formulated as:

\[
\min \ H(X) \quad \text{s.t.} \quad c_0 > 0, \ k_0 > 0
\]  

(6)

Since the proposed model is nonlinear, the gradient information is not able to be formulated by explicit expressions. Hence, the traditional gradient-based optimization method could not work well. Although this problem could be solved by the direct search algorithm, however, it relies on good choice of initial values to a large extent and may trap into the local optimum. Therefore, the global search algorithm is required. In the following section, a modified artificial fish swarm algorithm is designed to identify the model parameters.

3 Model Parameter Identification

The process of identification for model parameters is generally accomplished by the optimization algorithm. A modified artificial fish swarm algorithm (MAFSA) is devised in this part to estimate the optimal solution of model parameters with fast convergence and high accuracy.

3.1 Description of AFSA

AFSA is a novel population based evolutionary computation technique inspired by the natural social behaviour of fish school and swarm intelligence [8]. In AFSA, the AF individual state can be illustrated as \( X=(x_1, x_2, \ldots, x_n) \), where \( x_i (i=1, 2, \ldots, n) \) denotes the variable to be searched for the optimal value. The food consistence at present location can be expressed as \( Y=f(X) \), where \( Y \) is the fitness function. The distance between the individuals is defined as \( d_{ij}=||X_i-X_j|| \). The visual range of AF is denoted as visual, step is the maximum step length and \( \delta \) is crowd factor. The essence of AFSA is to search the optimal result through the iterative algorithm. In each iteration, the AF updates its state to achieve the optimum according to the different behaviours. The following part describes the main behaviours in AFSA.

3.1.1 Searching Behaviour

Suppose that \( X_i \) denotes the present state of the AF. In its sensing range, a new state \( X_j \) is randomly chosen. If \( Y_j > Y_i \), move a step in this direction; or else, choose a state randomly again and decide whether it meets the moving condition. If it still cannot meet the condition after several times \( \text{(try\_number)} \), it goes a step randomly.

3.1.2 Swarming Behaviour

The AF with the state \( X_i \) searches for the number \( n_f \) of its fellows in the present visual neighborhood \( (d_v < \text{visual}) \) and estimates the fellows’ central location \( X_c \). Here, \( Y_c \) represents the food concentration at the location of \( X_c \). When \( n_f \geq 1 \), the AF estimates the central location of its fellows. If \( Y_c/n_f > \delta Y_f \), forward a step to the fellows’ center because the food concentration at the central location is high and the environmental condition is not too crowded. On the contrary, the AF performs the searching behaviour.

3.1.3 Following Behaviour

The AF with the state \( X_i \) searches for the number \( n_f \) of its fellows in the present visual neighborhood \( (d_v < \text{visual}) \) and find the location \( X_{\text{max}} \) of its fellow with highest food concentration. \( Y_{\text{max}} \) is denoted as the value of food concentration at the location \( X_{\text{max}} \). If \( Y_{\text{max}}/n_f > \delta Y_f \), forward a step to the location \( X_{\text{max}} \) because the food concentration at the location \( X_{\text{max}} \) is high and the environmental condition is not too crowded. On the contrary, the AF performs the searching behaviour.

3.1.4 Behaviour Selection

The AF’s behaviour is determined by its hungry degree, which is signified by energy. Here, the mean energy \( \varphi \) is defined as:

\[
\varphi = \frac{1}{m} \sum_{i=1}^{m} Y_i
\]

(7)

If the food concentration at the location \( X_i \) is smaller than \( \varphi \), the AF executes the following behaviour to attain the food in the area with the high food concentration. If the food concentration at the location \( X_i \) is larger than \( \varphi \), the AF will perform the swarming behaviour to avoid the dangerous animals. If the AF does not perform the following or swarming behaviour, it will select the searching behaviour.

3.1.5 Bulletin

A bulletin board is set up to record the optimal individual’s state and the food concentration at the present location. Update the bulletin with the better state of the AF and the final value of the board is regarded as the optimal solution of the problem.
3.2 Modified AFSA

As a random searching algorithm, the AFSA has the benefits of less initial requirements, good global convergence, strong robustness, easy realization. However, it still has some disadvantages in the practical operation. First of all, since some parameters such as visual and step are usually set to constants, the algorithm has the slow convergence rate in the later stage of optimization, which makes it easy to fall into the local optimum. Then, the randomness of visual and step as well as the existence of random behaviour have a strong effect on the accuracy of optimization. Furthermore, the increase of AF number requires a large storage space, which also leads to an increase of the calculation amount. Therefore, some modifications are implemented to tackle the above problems.

3.2.1 Modification of Parameter Update

In basic AFSA, visual and step are two important parameters with respect to the algorithm performance. When visual is set to a large value, the AF has strong global searching as well as fast convergence rate. On the contrary, the AF has strong local searching ability. Similarly, the larger the value of step is, the faster the algorithm convergence will be, even though the numerical oscillation may appear sometimes. In contrast, the smaller step is, the slower the algorithm convergence will be, but the solution has the higher accuracy.

Based on the above analysis, in order to improve the global searching ability and the convergence rate of the AFSA, the larger visual and step are selected to make the AF search in a larger scope in the initial phase of the optimization. With the process of iteration, visual and step decline gradually so that the AF can carry out the local search in the adjacent domain of the optimal solution, which is aimed to increase the algorithm accuracy. In this paper, an inertia weight \( \omega \) based on the exponential function is designed to dynamically update visual and step. The update equations are given as:

\[
\begin{align*}
\omega &= \exp[-20 \cdot (iter/T_{max})^p] \\
\text{visual} &= \text{visual}_{\text{min}} + \omega \cdot \text{visual} \\
\text{step} &= \text{step}_{\text{max}} + \omega \cdot \text{step}
\end{align*}
\]

Where iter is the current iteration number; \( T_{\text{max}} \) is the maximal iteration number; \( p \) is an integer to control the change rate of parameters with the range \([1, 30]\); \( \text{visual}_{\text{min}} \) and \( \text{step}_{\text{max}} \) denote the minimal visual range and step length, respectively. Generally, both visual and step are the piecewise functions, which keep the maximal values in the early stage, descend gradually in the intermediate stage and retain the minimal values in the later stage. Figure 2 shows the inertia weight curves for different values of \( p \).

![Figure 2. The weight value](image)

3.2.2 Modification of AF’s Behaviour

In the conventional searching behaviour, the AF randomly selects a state and forwards a step to this state if it is superior to the present state. This operation would slower the search speed of the AFSA when the dimension of the AF is relatively higher. Furthermore, in the swarming and following behaviours, distances between the AF and others in its visual range are calculated to estimate the central location and optimal location, respectively. However, with the increase of the number of the AF, the calculation amount of the algorithm will become increasingly large, which may result in the long running time.

Aiming at the above drawbacks, some modifications for the AF behaviours have been conducted to improve the algorithm performance. In order to enhance the algorithm search speed, the AF directly moves to the optimal location in its visual range if it meets the condition of the searching behaviour. Besides, in order to reduce the computational complexity, the central and optimal locations in AF’s visual range are replaced by the responding ones in the whole swarm, respectively. This modification avoids the calculation for the distances between AF and others in its visual range and greatly shortens the running time.

3.3 Steps of Modified AFSA

In this part, the MAFSA is utilized to identify the parameters of the MRE base isolator model. The detailed process of parameter identification consists of the following steps:

Step1. Initialize the optimization problem and
The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

algorithm parameters: the problem is defined as Minimize \( H(X) \) subject to \( X_l \leq X \leq X_u \), where \( X = [\omega_0, c_0, \alpha, \beta, F_{01}], X_l \) and \( X_u \) denote the lower and upper bounds of the variable \( X \), the algorithm parameters include the number of the AF \( N \), \( \text{visual}_{\text{max}}, \text{step}_{\text{max}}, p, \delta, \text{try}_{\text{number}} \) and \( T_{\text{max}} \).

Step 2. Initialize the AF swarm: the initial locations of all the AF are generated randomly within the available range. Then set the initial iteration number \( \text{iter} = 0 \).

Step 3. Evaluate the fitness value: calculate and compare the food concentration of initial AF, then record and conserve the biggest one on the bulletin board.

Step 4. Update the values of \( \text{visual} \) and \( \text{step} \) according to equation (8).

Step 5. Behaviour operation: the AF performs the searching behaviour, swarming behaviour or following behaviour according to its hungry degree.

Step 6. Check the AF self-state and information on the bulletin board: if the fitness value is superior to that on the board, update the board. Or else, it remains unchanged.

Step 7. Check the termination criterion. If the iteration number is equal to the maximum iterations or the solution for the optimization problem is equal to the target value, perform Step 8. Or else, \( \text{iter} = \text{iter} + 1 \) and go to Step 4.

Step 8. Terminate the algorithm: output the optimal solution and responding the fitness value of the AF (that is the optimal model parameters and food concentration on the board).

In conclusion, the process of the MAFSA is illustrated in Figure 3.

4 Identification Results and Analysis

In order to validate the effective of the new model and the availability of the proposed MAFSA, MRE base isolator parameter identification processes will be carried out and compared in the context of model accuracy and algorithm performance.

4.1 Experiment Setup

Several groups of experimental data are acquired from a MRE base isolator and fed into the MAFSA to identify the parameters of the proposed new model. The isolator is excited by a sinusoidal displacement in several test cases with different excitation displacements and magnetization currents supplied to the MRE base isolator. The driving frequency is set as 1Hz, the current ranges from 0A to 3A whilst the displacement is varied from 2mm to 8mm, respectively. The excitation displacement has 3 settings and the current contains 4 entries. Therefore, a total of 3x4=12 experimental data sets are made up, which can be summarized in Table 1. The parameters of MAFSA are set as follows: \( N=50 \), \( \text{try}_{\text{number}}=50 \), \( \delta=0.6 \), \( T_{\text{max}}=300 \), \( \text{visual}_{\text{max}}=0.002 \), \( \text{step}_{\text{max}}=0.0003 \) and \( p=5 \). Moreover, the identification is conducted using both the Bouc-Wen model and the proposed model while the algorithms adopted contain MAFSA and other conventional ones.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current (A)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Displacement (mm)</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Frequency (Hz)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Current (A)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Displacement (mm)</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
</tbody>
</table>

4.2 Identification results

To validate the effectiveness of the new model to predict the performance of the MRE base isolator, a group of data (1Hz, 0A current and 2 mm amplitude)
are used for parameter identification. The optimal results are given in Table 2. Figures 4 and 5 show the tracking process in a sampling cycle and relative errors between the experimental and predicted forces, respectively. It can be seen that the relative error is within ±6%, which is acceptable in the modelling study. Figures 6 and 7 demonstrate the force-displacement responses and the nonlinear relationship between force and velocity. It is clearly shown that the estimated forces resemble the practical testing ones very well, especially in the regions where the strain stiffening is evident.

In order to further verify the ability of the model for depicting the hysteretic behaviours of the MRE base isolator, more groups of comparisons between the practical testing and predicted data corresponding to different loading conditions are conducted. Figures 8 and 9 reveal that the estimated data is well fitted to the experimental data under the conditions of 3A, 1 Hz frequency for 2mm, 4mm and 8mm amplitudes respectively. It is clearly seen that experimentally measured responses are reasonably modelled. The measured force-displacement pairs shown in Figure 10 are acquired by loading the isolator with a 1Hz sinusoid and a 4mm amplitude at three current levels, 0A, 1A, 2A and 3A, respectively. The four group comparisons validate the capacity of the model to portray the increasing nonlinearity of the hysteretic loops with the increasing currents. In particular, in each hysteresis loop, the estimated data resembles the unique behaviour of straining hardening perfectly.

Table 2. Parameter values of the proposed model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$c_0$</th>
<th>$k_0$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$F_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.354</td>
<td>3.572</td>
<td>4.579</td>
<td>0.302</td>
<td>0.001</td>
</tr>
</tbody>
</table>

To evaluate the accuracy of the proposed model, a root-mean-square (RMS) is adopted as the assessment criteria given by:

$$e_{rms} = \sqrt{\frac{1}{N_i} \sum_{i=1}^{N_i} (F_{i}^{pre} - F_{i}^{exp})^2}$$  \hspace{1cm} (9)
is evidently seen that because of the higher degree of nonlinearity in the Bouc-Wen model, a bigger RMS error is shown from the predicted force from the model. For another thing, the errors in the proposed model are generally less than that of the Bouc-Wen model.

Figure 8. Force-displacement response under different loading amplitudes

Figure 9. Force-velocity response under different loading amplitudes

Figure 10. Force-displacement response under different loading currents

In the meantime, the conventional Bouc-Wen model is also used to compare with the proposed one for model error analysis. Figure 11 shows the comparison result. It is evidently seen that because of the higher degree of

4.3 Identification Algorithm Analysis

The efficiency of the identification process is generally associated with the model complexity and the parameter values of the optimization algorithm. Here, the model is determined and the influence of visual and step on the optimization process has been illustrated in 3.2. So in this part, the AF population N, as one of most important parameters, is studied for its effect on the algorithm performance. Figure 12 describes the convergence of the MAFSA with different population number over 300 iterations under the loading condition of 1A current, 1Hz sinusoid and 4mm amplitude, and the optimal fitness value and calculation time are listed in Table 2. It is obvious that with the increase of the AF number, the MAFSA has the higher identification accuracy as well as more calculation time. However, when the population number exceeds 50, the fitness value varies very slightly. Thus, the selection of 50 as the AF population is reasonable.

So as to demonstrate the superiority of the MAFSA, two conventional optimization algorithms are adopted for performance comparison: (a) AFSA with fixed visual and step; (b) Particle Swarm Optimization (PSO) [9]; (c) MAFSA. For the purpose of fair assessment, all the parameter values in AFSA are similar to that of MAFSA, except visual and step. Figure 13 shows the comparison result. It is shown that although the AFSA has the fastest convergence among three algorithms, it results in the premature convergence. Compared with the PSO, MAFSA arrives at its optimum more quickly and has higher identification accuracy.
5 Conclusion

This paper has presented a new model for MRE base isolators and an efficient optimization algorithm based on AFSA for model parameter identification. The new model adopts the hyperbolic sine function to describe the hysteretic relationship between the isolator force and displacement, and requires as smallest as five parameters in contrast with eight parameters of the Bouc–Wen model. An efficient optimization algorithm based on AFSA is also designed for the model parameter identification. In order to improve the convergence rate and reduce the calculation amount of the identification process, a self-adaptive parameter update approach is introduced and some behaviours in basic AFAS are also simplified. Experimental data from a practical MRE base isolator are utilized for modelling verification. The results obtained by the new model have shown highly satisfactory coincidence with the experimental data, and also the effectiveness of the proposed identification algorithm.

References


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

BUILDING AND ARCHITECTURE
Development of Lightweight BIM Shape Format Structure to Represent Large Volume Geometry Objects Using GIS with Facility Management

T.W. Kang\textsuperscript{a} and C.H. Hong\textsuperscript{b}

\textsuperscript{a}ICT Lab, Korea Institute of Construction Technology, Republic of Korea
\textsuperscript{b}ICT Lab, Korea Institute of Construction Technology, Republic of Korea
E-mail: laputa99999@gmail.com, chhong@kict.re.kr

Abstract -
This study focuses on the development of a lightweight building information management (BIM) shape format (LBSF) structure to represent large-volume BIM geometry objects using the geographical information system (GIS) with building facility management. Recently, BIM-based facility management with GIS has been researched with regard to urban facility management. To implement this use case, the BIM geometry objects are required to be effectively visualized on the urban level using GIS. Therefore, a lightweight BIM shape format is designed with these considerations, and the prototype for the pilot test is implemented. In the pilot test phase, after developing the Industry Foundation Classes (IFC) file and LBSF file for the model data of the three areas, performance comparison with regard to the data volume and screen loading time was made.

Keywords -
BIM; Lightweight; Shape; Format; Performance

1 Introduction

Recently, studies have been conducted on the integration of the geographical information system (GIS) and the building information management (BIM) to achieve use cases such as building facility management. One of the challenges for these studies is effective visualization and representation of the geometric information of multiple BIM objects using GIS. IFC, which is the standard model generally used in the BIM field, has a complex structure and possesses parametric information; therefore, it experiences performance-related issues when used. In particular, as IFC is text-based with an uncompressed file format, processing of multiple strings and tokens is performed during the file parsing stage when the file is loaded. Consequently, to visualize the geometry of an object, mesh processing is also performed using the solid parametric information on a variety of objects included in IFC. Therefore, these IFC file processing stages require a long loading time.

With regard to urban building facility management, the performance-related issues that occur when the geometries of numerous BIM objects are visualized using GIS must be solved. For example, when rendering or selecting a BIM object, and when the performance is too low for the user to operate the system, it is difficult to perform tasks such as confirmation of the BIM object data. Thus, to effectively visualize the geometric data of large-volume BIM objects, weight reduction the large-volume geometric information is required. It is possible to efficiently connect the GIS data, visualize the large-volume data of the city unit, and provide Internet services on the data when these issues are solved. Therefore, this study proposes a method and a format structure to represent lightweight BIM geometry objects and realizes the prototype to confirm the effectiveness of visualization of the geometric information. Accordingly, studies on the weight reduction of geometric information are investigated, and necessary steps for IFC processing are performed in advance to improve the visualization performance. However, as this study focuses on the method for weight reduction of large-volume BIM objects, the level of detail (LOD) maximization of file format characteristics and spatial indexing have not been included.

2 Study Method

In comparison to the study method, as shown in Figure 1, the current trend of the relevant studies is investigated, and the limitations and implications of the previous studies are derived. The limitations are identified by analyzing the structure of the IFC files and standard BIM format, and a method and a format structure are proposed to visualize the geometry of
large-volume urban facility objects that are represented by GIS-based urban facility management. Subsequently, the structure of lightweight BIM shape format (LBSF) is defined, and the class and component structure for LBSF assistance are proposed. To confirm the performance of the proposed format structure, a program is developed, and a pilot test is conducted.

By using three samples in the pilot test, we compare the three methods including the conventional method and the two proposed methods in which weight reduction of geometric information has been performed. The verification methods by sample are tested with regard to the loading time, memory usage, and frames per second (FPS) indexes.

**Figure 1. Research flow**

### Research Objective and Methodology
- Research Trends Survey
- BIM Standard File Format IFC Limitation Analysis considering the Facility Management
- Algorithm Design to Improve the BIM Visualization Performance based on GIS
- Lightweight BIM Shape Format (LBSF) Design
- Prototype Program Development for the Pilot Test
- Pilot Test and Verification

### Component Design for LBSF
- Class Design for LBSF
- Component Design for LBSF

### Conclusion and Suggestion

#### 3 Current Status of Relevant Studies

Few studies on the visualization of large-volume object shapes have been conducted in the BIM field. In the product lifecycle management (PLM) field, which is similar to the BIM field, the method to effectively visualize the geometric information on large-volume objects has been widely studied.

Studies on promoting Web-based cooperation using large-volume shape data created in CAE/CAI were also carried out, and the methodology and data structure to reduce and visualize the created data for Web-based cooperation were investigated [1]. In the applied method, the mesh was simplified under certain conditions to reduce the weight of the parts that had an insignificant influence on the shape mesh analysis. As mesh simplification was only performed on the mesh that was output as the analytic result, the objectives of these studies are different from that of the building facility management.

Except the abovementioned studies, those on smooth data exchange through the weight reduction of analytic data using the JT format and the JT Toolkit were conducted [5]. Moreover, studies on using the Graphic Process Unit (GPU) and applying the mesh simplification method to eliminate the hidden surfaces for virtual factory simulation and three-dimensional (3D) data weight reduction were conducted. In the study on the 3D visualization using CityGML [11], the model shape was visualized through WebGL, which supports the 3D acceleration function of the GPU. However, the improvement of the visualization performance through the data format was not considered in that study. In another study related to WebGL [12], a lightweight structure considering data compression and data streaming was proposed, whereas LOD processing required to visualize BIM object shapes using GIS was not considered.

Further, in the PLM field, IGES and STEP, the international standards focusing on information exchange, were criticized for their significant size and inefficiency [7]. Therefore, lightweight visualization data formats such as JT, U3D, and 3D-XML, which are used for the common PLM, have been used.

With regard to the formats used for common PLM, JT was developed through the Jupiter Project by HP and the Engineering Animation Inc., for efficient shape visualization, in 1998 [9]. However, while providing data compression and LOD technologies, JT includes the B-Rep structure, attribute data, and complex format structure. Therefore, practical application of the format requires purchase of license and a development tool. U3D is the standard format announced in the 3D industry forum in which 24 companies on the basis of Intel participated in 2004. Although gradual viewing is provided through data streaming in the download process, speed degradation due to additional operations for mesh separation was pointed out [10]. 3D-XML proposed by Dassault has a simple structure as it only manages a vertex, a phase, and additional information.

According to the relevant studies mainly conducted in the PLM field, the method and structure required to effectively visualize large-volume BIM shape data in terms of building facility management have been rarely researched upon, and it is difficult to directly use the solutions of previous studies owing to various issues such as license.
4 Limitations and Considerations in the Case of Using the IFC Files and the BIM Standard Format with Facility Management

In March 2013, the buildingSMART alliance, a standards organization whose objective is to improve the exchange of information on architecture and construction, officially announced IFC4, the BIM standard format. It is the upgraded version of IFC 2x3, and various problems of the previous version have been solved. IFC has an object-oriented structure, and the buildingSMART alliance is continuously promoting standardization activities to exchange information on construction. The objective of the IFC4 modeling is to eliminate the ambiguity of the information structure, and the performance was enhanced, which included irregular model representation through the NURBS support, GIS connection through the coordinate system support, and improvement of the 4D/5D information models. With regard to data weight reduction, a reduction of 14% was achieved by decreasing the number of XML tags used in the ifcXML format and merging individually defined data nodes into groups in order to simplify them in the form of the data list [8].

We have attempted to improve the visualization performance of IFC4 through data weight reduction. However, the information required for visualization should be calculated because the focus is on representing parametric models and exchanging information. For example, additional mesh processing operations are required to visualize the walls and roofs that are represented as mathematical parameters.

![IFC model visualization algorithm](image)

Figure 2. IFC model visualization algorithm

The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

The following information is particularly unnecessary from the perspective of management and operation:

1. The local coordination reference system of the object
2. Parametric information for mesh generation of the object

The relevant C, D, and E stages are unnecessary unless shape modeling is required. Moreover, the stage of parsing IFC includes the processing of numerous character strings and syntactic analysis. When these parts are processed in advance, the speed of loading GIS-based BIM objects of various building facilities is increased. Accordingly, the algorithm that creates geometric information to be processed in advance in the IFC visualization stage and the format that includes this information are designed. Furthermore, the LBSF component and class structure are proposed.

5 Design of the Lightweight BIM Object Shape Model

5.1 Algorithm Design

On the basis of the considerations derived in Chapter 4, the algorithm that converts the parametric BIM shape data into an exterior surface has been designed as shown in Figure 3.
Figure 3. File format conversion algorithm from IFC

The conversion of the world coordination reference system and the creation of modeling information, which are unnecessary operations in terms of building facility object management, are calculated in advance, and the LOD is also obtained using the mesh simplification method.

To create a mesh of certain shape, the process of transforming a curved surface with a curvature into a mesh is required. The accuracy of the mesh is expressed by the number of segments in which the curves are approximated as straight lines as follows:

\[ S_N = f(r) \]

The \( f() \) function determines \( S_N \) by applying the following rules:

- if \( r < 0.03 \), then \( S_N = 4 \)
- else if \( r < 0.05 \), then \( S_N = 6 \)
- else if \( r < 0.1 \), then \( S_N = 8 \)
- else if \( r < 0.5 \), then \( S_N = 10 \)
- else \( S_N = 36 \times \frac{a}{\pi \times 2} \)

The accuracy of the curved surface expressed by the mesh varies according to the determined \( S_N \). However, the issue of shape accuracy is not considered in the research range of this study because the objective of this study is to improve the visualization performance in terms of urban facility management.

For smooth services of large-volume data, the concept of LOD processing by CityGML, which is related to the improvement of visualization performance, is considered by extracting the exterior surface information. When the BIM objects are visualized in the practical GIS, the loaded LBSF files converted from the IFC format are used.

As shown in Figure 4, the algorithm for extracting the exterior surface data is employed by placing cameras to focus on the central point of the BIM shape for data on each surface and then examining the surface data located around the BIM shape. However, this method cannot properly extract the exterior surface data when certain surface information is not observed by the camera in curved or complex surfaces, and therefore, it is necessary to improve this method.

Figure 4. Exterior surface extraction method

5.2 Design of the File Format

The file format is designed to satisfy the following requirements:

1. Visualization performance: geometric calculations should be minimized during visualization. For this, the surface information calculated in advance is used, and the file format should manage LOD.
2. Minimization: the file size should be minimized.
3. Elimination of redundancy: the data repeatedly shown in the file format should be separated and referred to as the index.
4. Compressibility: factors with similar properties in the file format are managed as groups to easily facilitate compression and decompression within the groups.

The structure of the file format comprises four main parts as shown in Figure 5.

Figure 5. File format of the overall structure

The header includes information such as the file version, location, direction, and number of objects. The object index manages the LOD chunk list to consider the visualization performance. The spatial index enables the octree nodes to manage objects only to represent the
visualized areas. The geometry comprises a list of actual coordinates of the surface data that represent the objects. The detailed structure of the header is as follows.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Name</th>
<th>Length (Byte)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>char</td>
<td>Signature</td>
<td>10</td>
<td>File format signature</td>
</tr>
<tr>
<td>char</td>
<td>Version</td>
<td>3</td>
<td>Version</td>
</tr>
<tr>
<td>string</td>
<td>File name</td>
<td>255</td>
<td>File name</td>
</tr>
<tr>
<td>integer</td>
<td>Object count</td>
<td>4</td>
<td>Object count</td>
</tr>
<tr>
<td>integer</td>
<td>LOD count</td>
<td>4</td>
<td>LOD count</td>
</tr>
<tr>
<td>bool</td>
<td>Is-Exterior</td>
<td>1</td>
<td>Exterior surface flag</td>
</tr>
<tr>
<td>vector</td>
<td>Location</td>
<td>20 = 8 + 8 + 4</td>
<td>Object location</td>
</tr>
<tr>
<td>vector</td>
<td>Orientation</td>
<td>12 = 4 × 3</td>
<td>Object orientation</td>
</tr>
<tr>
<td>vector</td>
<td>Scale</td>
<td>12 = 4 × 3</td>
<td>Object scale</td>
</tr>
<tr>
<td>rect3d</td>
<td>Bounding box</td>
<td>48 = 8 × 6</td>
<td>Object maximum bounding box</td>
</tr>
</tbody>
</table>

The structure of the object index is as follows. Considering expandability, the LOD degree of the object can vary depending on the purpose of the application.

Table 1. Header structure

Table 2 shows the structure of the head that manages general information on the LOD including the LOD degree and the start position and length of geometric information.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Name</th>
<th>Length (Byte)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>byte</td>
<td>LOD-Level</td>
<td>1</td>
<td>Level Index of LOD</td>
</tr>
<tr>
<td>long</td>
<td>Start-Position</td>
<td>4</td>
<td>Geometry Start Position</td>
</tr>
<tr>
<td>long</td>
<td>Length</td>
<td>4</td>
<td>Geometry length</td>
</tr>
</tbody>
</table>

The information structure of the objects such as LOD is summarized in Table 3. General information on LOD and the start and end positions of the object vary according to the number of objects.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Name</th>
<th>Length (Byte)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>string</td>
<td>GUID</td>
<td>22</td>
<td>Object GUID</td>
</tr>
<tr>
<td>long</td>
<td>Start position</td>
<td>4</td>
<td>Start position in file stream</td>
</tr>
<tr>
<td>long</td>
<td>End position</td>
<td>4</td>
<td>End position in file stream</td>
</tr>
<tr>
<td>LOD</td>
<td>LOD Variables</td>
<td>LOD data stream</td>
<td></td>
</tr>
</tbody>
</table>

The structure of spatial indexing is not determined in this study; however, the area for relevant information is allocated in the format structure for further expandability.

The format structure of the geometry, summarized in Table 4, comprises a list of information that represents geometric shapes by a unique ID and LOD in the form of surface data comprising that of the peak, corner, and side.
Table 4. Geometry information structure

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Name</th>
<th>Length (Byte)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>string</td>
<td>GUID</td>
<td>22</td>
<td>Object GUID</td>
</tr>
<tr>
<td>binary</td>
<td>LOD</td>
<td>list</td>
<td>LOD geometry information list</td>
</tr>
</tbody>
</table>

5.3 Class Design

On the basis of the developed format, the class structure has been designed to be object-oriented as shown in Figure 7.

Figure 7. LBSF class structure (UML)

5.4 Design of the Component Structure

The component structure is demonstrated in Figure 8.

Figure 8. LBSF component structure

The main roles of each component are as follows:
Object Property DB: It separately manages the attribute data based on object and divides the data into an object GUID. This object GUID has the same value as that of the GUID of the BIM object and is used when it is connected with the properties of the geometry model in the LBSF.
IFC Parser: This component loads and parses the IFC file format, using the function of IFC Gear, which is an open source.
LBSF Converter: This component performs the mesh processing on the geometry of each parsed object by LOD and simplifies the mesh.
LOD Mesh Processing: This component performs mesh processing by LOD and partially uses the functions of OpenCASCADE, which are open source.

6 Comparison and Verification

With regard to the visualization performance, the proposed weight lightening method has been compared with the surface information conversion method, which has been proposed to handle information such as the file loading time, data volume, and memory usage; the method to extract the exterior surface; and the conventional method to use the IFC format.

Table 5. LOD general information structure

<table>
<thead>
<tr>
<th>Data</th>
<th>Name</th>
<th>Shape Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>Cheongun University</td>
<td>Sample #1 Main and Academic Information Center</td>
</tr>
<tr>
<td>Sample</td>
<td>Family Restaurants</td>
<td>Sample #2 Family Restaurants</td>
</tr>
<tr>
<td>Sample</td>
<td>Sungkyunkwan University dormitory building</td>
<td>Sample #3 Sungkyunkwan University dormitory building</td>
</tr>
</tbody>
</table>

Table 6 lists the data characteristics of the samples used for performance comparison and verification. Although there are many ongoing BIM-based projects, it is practically difficult to use them as samples. Therefore, samples that have already been modeled and permitted for use in academic research have been selectively used in this study.

Table 6. Characteristics of Used Samples

<table>
<thead>
<tr>
<th>Type</th>
<th>Vertex Count</th>
<th>Data Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample #1</td>
<td>15,454,362 points</td>
<td>301,844 KB</td>
</tr>
</tbody>
</table>
The proposed methods are tested with samples with regard to the loading time, memory usage, and FPS indexes, which are considered important for visualizing large-volume BIM geometry. The time of loading IFC files, which is used as the performance measurement index, includes the time required for the parsing process and mesh processing of the geometry. However, the speed is expected to increase because these stages have already been processed in the lightweight BIM geometry format. The memory usage is influenced by a certain algorithm or by the use of the data format with regard to performance improvement. It is investigated to determine the influence of the proposed methods. FPS is related to the response rate; the lower the FPS value, the more difficult it is to observe the geometric information. In general, the higher the FPS value, the easier it is to observe the geometric information. Figure 9 shows a screenshot of the prototype program operation.

The results of the performance evaluation test conducted on the samples are summarized in Table 7.

Table 7. Test results of samples (data volume, loading time, memory usage, FPS)

<table>
<thead>
<tr>
<th>Format and Performance</th>
<th>Sample #1</th>
<th>Sample #2</th>
<th>Sample #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data volume (MB)</td>
<td>67</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>IFC</td>
<td>22.05</td>
<td>217.22</td>
<td>5.99</td>
</tr>
<tr>
<td>Loading time (second)</td>
<td>499</td>
<td>1,029</td>
<td>156</td>
</tr>
<tr>
<td>Memory (MB)</td>
<td>16.44</td>
<td>35.71</td>
<td>33.78</td>
</tr>
<tr>
<td>FPS</td>
<td>LBSF with exterior surface</td>
<td>42.73</td>
<td>58.14</td>
</tr>
</tbody>
</table>

According to the results of the quantitative performance evaluation, the data loading time of the LBSF format decreased by 95% in Sample #2. This implies that the performance was enhanced because the LBSF format performed parsing and the mesh processing in advance. The memory usage decreased by a maximum of 64% in Sample #1; the difference was insignificant in terms of the model data volume and FPS. This is because the amount of surface data required to be rendered for the two formats is constant. Therefore, the LBSF format that calculated the surface data in advance significantly improved the data loading time.

The LOD model data volume of the LBSF, which processed the exterior surface data at a level lower than that of LOD3, decreased by ~89% in Sample #1. The data loading time decreased by ~98% in Sample #2; the memory use decreased by ~88% in Sample #1; and FPS improved ~1.9 times in Sample #3. The general performance was enhanced significantly because the amount of exterior surface data to be rendered decreased.

The format performance of each sample has been analyzed in the form of graphs as shown in Figures 10 through 13. In general, when the BIM data of the IFC format are converted into the LBSF model, the loading time and memory usage rapidly decrease, whereas FPS improves. The result is consistent with the expectation prior to the experiment.

Figure 10. Samples data volume comparison of each format
7 Conclusion

This study proposed a method and a format structure that represents lightweight BIM geometry objects to solve performance-related problems encountered when the geometries of numerous BIM objects are visualized using GIS, which is necessary from the building facility management perspective. Accordingly, the prototype was realized, and the effectiveness of the geometric information visualization was confirmed on the basis of performance improvement using three samples and four indexes.

The comparison and verification results indicate that it is difficult to effectively represent a large number of BIM shapes that are expressed on a large area such as GIS, when the IFC format including all the information required for modeling is used from the building facility management perspective. In this case, it is beneficial to use the format that extracts the necessary shapes and properties from the perspective of facility management. Furthermore, the performance of the proposed LBSF format is better than that of the IFC format. In particular, the LBSF format that has only the exterior surface data is more effective for rendering all the BIM objects with regard to the GIS performance. Therefore, each format should be strategically used from the user and LOD perspective.

In our future work, mesh processing, in which the exterior surface data is handled, will be improved by considering the spatial index for visualization performance improvement. Moreover, the low number of samples, which is the limitation of this study, will be improved, and studies on automatically generating various LODs, which are required for visualization using GIS, will be conducted. Furthermore, the methods for data compression and virtual memory management will be studied to develop a distributed processing method for large-volume BIM objects using GIS.

Acknowledgements

This research was supported by a grant from the Strategic Research Project (Development of BIM/GIS Interoperability Open-Platform 2014) funded by the Korea Institute of Construction Technology.

References


The Adoption of Building Information Modeling in the Design Organization: An Empirical Study of Architects in Korean Design Firms

H. Son, S. Lee, N. Hwang, and C. Kim

Department of Architectural Engineering, Chung-Ang University, Seoul, Korea
E-mail: hjson0908@cau.ac.kr, leesungwook@cau.ac.kr, nahyae@cau.ac.kr, and changwan@cau.ac.kr (corresponding author)

Abstract -
Recently, Building Information Modeling (BIM) technology has attracted much attention in the Architecture, Engineering, and Construction (AEC) industry. Despite the growing interest in BIM technology, the benefits of BIM have not yet been fully realized during the course of implementation of BIM because of its low adoption rate among architects. Therefore, it is significant important in successful adoption of BIM in design organizations, understanding of the factors influencing the adoption of BIM. The aim of this study is to empirically examine the individual, organizational, social, and technical factors affecting architects’ adoption of BIM. The 162 architects with experience using BIM tools at three major design firms in South Korea were selected to participate in the face-to-face survey. This study extends the Technology Acceptance Model (TAM) by incorporating constructs such as computer self-efficacy from the individual domain, top management support and technical support from the organizational domain, subjective norm from the social domain and compatibility from the technical domain. The results strongly support the extended TAM in predicting the intention of users to adopt BIM. It also demonstrates the significant effect of computer self-efficacy, top management support, subjective norm, and compatibility on behavioral intention through perceived ease of use and perceived usefulness. This study provides academics and practitioners with the understanding of factors leading to the successful implementation of BIM in design organizations. It also provides insight into the role management plays in the adoption of BIM among architects in the AEC industry.

Keywords -
Behavioral intention; Building and architecture; Building information modeling; Technology acceptance model; Design organizations

1 Introduction

Recently, BIM technology has been attracted much attention in the AEC industry. BIM technology can be defined as the technology of generating and managing a parametric model of a building [1]. The successful implementation of BIM technology is beneficial for project stakeholders throughout the project life cycle. The following benefits for the architects who utilize BIM in the design process seem especially apparent. Those include reduced document errors and omissions, reduced rework, and reduced cycle time of design process [2]. Moreover, the successful implementation of BIM can help to improve the productivity of subsequent processes, such as construction, operation, and maintenance.

However, the benefits of BIM have yet to be fully realized even during the course of its implementation. According to the SmartMarket Report from McGraw-Hill Construction, only 3% of survey respondents stated that they experience its full benefits [3]. Such a discrepancy between expected benefits and realized benefits of BIM may be explained by its low adoption rates by architects [4]. This raises the following question: Why with significant benefit of BIM do architects hesitate to adopt it? Several previous studies identify various factors that make architects are afraid to accept the BIM. To summarize these previous studies, the main issues to BIM include management support, technical support, compatibility of BIM technology, software/computer skills and organizational culture [5].

Despite the significant importance of BIM’s successful adoption by design firms, understanding the factors that influencing this adoption by architects has yet to be seriously explored. Consequently, it is important to examine the question of how those factors affect an architect’s decision to adopt BIM. Thus, the aim of this study is to empirically examine the factors affecting architects’ adoption of BIM. This study extends the TAM by incorporating constructs such as computer self-efficacy from the individual domain, top
management support, technical support, and training from the organizational domain, subjective norm from the social domain and compatibility from the technical domain. The results of this study will provide the understanding of factors leading to the successful implementation of BIM in the organization.

2 Technology Acceptance Model

The TAM is derived from the Theory of Reasoned Action (TRA) and was developed by Davis [6]. It was widely accepted as a framework to explain users’ adoption of Information Technology (IT) by examining variables that influence users’ behavioral intention and also explaining behavioral intention by using variables [7]. TAM provides general information on users with the advantages of parsimony, robustness, and broad empirical support [8]. However, it has theoretical limitations that cannot reflect various factors [9]. Therefore, to overcome the limitations, an extended TAM has been suggested that enables the investigation of specific factors that are expected to influence the users’ behavior by using specific technology that adds additional factors as external variables into the TAM [7].

The extended TAM has enabled the determination of specific validity factors to explain users’ behavioral intention and to understand the factors when adopting the technology in each study. Furthermore, it has been helpful for in the successful adoption of technology through considering specific validity factors. Therefore, this study proposes that the extended TAM examine the factors that influence the behavioral intention of architects toward the adoption of the BIM

3 Theoretical Framework and Hypotheses

3.1 Research Model

This study suggests an extended TAM to examine the factors that influence the behavioral intention of architects in the adoption of the BIM. The proposed extended TAM comprises external variables, such as technical support, computer self-efficacy, compatibility, subjective norm, and top management support, as well as components influenced by external variables, such as perceived usefulness, perceived ease of use, and behavioral intention. Figure 1 suggests that perceived usefulness and perceived ease of use are instrumental in explaining the variance in users’ intention as based on prior research [10]. Behavioral intention was used to measure the architects’ adoption of BIM but not actual use. For some projects, the contractor or client has the right to require the use of BIM [11]. In these cases, usage of BIM is mandatory, thus requiring designers to use BIM regardless of their intention. For this reason, actual use cannot be regarded as a measure of architects’ acceptance of BIM adoption. An additional reason that the use of behavioral intention as a measure of the architects’ acceptance of BIM adoption is that it is appropriate to measure behavioral intention of technology adoption in a mandatory environment [12].

Perceived usefulness and perceived ease of use are defined by Davis [6] as positively influencing behavioral intention in the original TAM when adopting technology [13]. Perceived usefulness is the degree to which a user believes that using a particular technology will enhance his or her performance [7]. If using BIM for architects enhances their performance, then it would be considered to have a positive effect on behavioral intention. Also, there is some evidence that using BIM can improve productivity [14] in the construction industry. Thus, we hypothesize that:

H1. Perceived usefulness will have a positive effect on the behavioral intention.

H1

Figure 1. Proposed extended TAM
Specifically, perceived ease of use is the degree to which a user believes that a technology will be easy to understand and will require no effort to use [7]. If architects believe that using BIM is easy, it can affect their behavioral intention positively. So, we hypothesize that:

H2. Perceived ease of use will have a positive effect on the behavioral intention.

H3. Perceived ease of use will have a positive effect on perceived usefulness.

Technical support, computer self-efficacy, compatibility, subjective norm, and top management support were selected as external variables in the adoption of BIM. These external variables are assumed to directly affect perceived usefulness and perceived ease of use, and indirectly affect behavioral intention [15]. Details on each variable will follow.

3.2 Top Management Support

Top management support is defined by “how individuals within a firm perceive the support of management for functions such as IT as well as the willingness of management to implement specific IT functions” [16]. It has been emphasized that top management support is critical for the successful adoption of technology within an organization [17]. Specifically, the success of any technological adoption and implementation is dependent upon the support of top management because this group not only establishes priorities within the organization, but also provides funding and implements protocols [18].

Top management support will have similar effects on the acceptance of BIM among employees. Because final decisions are made by top management, top management must have an understanding of the intention of technology users within the firm, how the use of technology can benefit the employees and the costs associated with the technology.

The key to successful implementation and adoption of technology within a firm is ensuring that top management supports the implementation and that employees perceive that management supports the technology [19]. When top management supports the technology, it is also able to provide guidance and support to employees who are not comfortable with the technology or who are reluctant to change [20]. Additionally, there is an element of psychological support that goes along with top management support. Therefore, we can hypothesize that:

H4. Top management support will have a positive effect on perceived usefulness.

3.3 Subjective Norm

Subjective norm is the individual perception that others believe that the individual thinks people should or should not perform certain acts. It has been shown to be linked to individual usage of IT [21]. In this study, subjective norm refers to the architects’ belief that BIM is useful. The architects use it due to the suggestions of colleagues and important professionals in their field who have encountered BIM previously. For architects, design firms require the usage of BIM that architects are more influenced by subjective norm than the usage of other voluntary technology [22]. Also, subjective norm is more important in the early stages of technology adoption because of the limited direct experience users have in that stage; in the early stages of adoption, users have not yet formed attitudes about the technology [23].

In the construction industry, when the adoption of BIM is in the early stage [24], subjective norm is supposed to influence architects’ behavioral intention. Therefore, based on previous studies, subjective norm is expected to have a significant impact on the adoption of BIM. Thus, we hypothesize:

H5. Subjective norm will have a positive effect on the perceived usefulness.

3.4 Compatibility

Compatibility is the degree to which technology users’ feel that a technology matches their needs, values, and work practices. It is measured based on the users’ experiences [25]. High compatibility has positive effects on technology adoption [24]. Compatibility will affect users’ behavioral intention because it contributes to the ease of adoption of the technology [26]. Users will adopt the technology when they view it as being compatible with and effective for their work goals.

Based on these previous findings, compatibility is likely to influence user’s behavioral intention with respect to adopting new technology. BIM is a tool that architects will most likely find to be compatible, with their work needs and goals. It is a great tool for data management because it utilizes a constant data format. This makes it possible to retrieve information, and that, in turn, allows architects to display this information to other members of a project as well as to clients [13]. BIM is also a more accurate and suitable tool for architects than is 2-D drawing [27]. When using BIM, architects are able to show and mark their designs and materials from various angles, like construction development processing, by using computer projection that enhances the design with real buildings and materials [28]. For these reasons, we hypothesize that compatibility will affect architects’ behavioral intention. In view of these conflicted findings, we hypothesize the following:
H6. Compatibility will have a positive effect on perceived usefulness.

H7. Compatibility will have a positive effect on perceived ease of use.

3.5 Technical Support

Ralph [29] defined technical support as the assistance, offered by knowledgeable people, that technology users need when they use computer hardware or software products [30]. It has been demonstrated in previous research that technical support is one of the most influential factors in determining users’ behavioral intention in technology adoption [31].

If technology adoption is mandatory, technological support will have positive effects on individual IT usage [32]. Therefore, it can be reasonably expected that technical support can enhance users’ behavioral intention [5]. These previous studies note that, when users encounter new technology as well as being proficient with BIM, support that includes training is essential to learning how to use BIM in their work [26]. However some design firms assume that professionals who are proficient in using Computer-Aided Design (CAD) software can learn BIM quickly without any other training and therefore overlook the importance of training. BIM is fundamentally different from CAD and BIM training is essential for every professional who engages in designing or producing documents with it [26]. Additionally, new staff may be needed, such as an interoperability manager or a structural modeler. This will add to the costs of implementing such a program [2]. Based on the importance of technical support for BIM adoption, we can hypothesize that:

H8. Technical support will have a positive effect on perceived ease of use.

3.6 Computer Self-Efficacy

Computer self-efficacy is the individual belief that one is capable of using a computer competently [33]. Computer self-efficacy is based on the concept of self-efficacy by Bandura [34]. The concept refers not only to the individual’s abilities, but also to the individual’s belief about his or her abilities. These beliefs influence how the individual engages in tasks [34]. Computer self-efficacy takes Bandura’s concept and specifically applies it to the use of computers and technology. If an individual has high computer self-efficacy, he or she will be able to proficiently deal with new and difficult computer systems or software with minimal support or assistance. In contrast, if an individual has low computer self-efficacy, they maybe more easily frustrated by learning new technology [35].

In this study, computer self-efficacy is represented by the ways in which architects use BIM proficiently and their beliefs about their abilities with to use BIM. Khorrami-Arani [36] showed that computer self-efficacy corresponds with achievement of computer competency. Success in the use of BIM is related to the confidence that the user is able to use BIM in the future [37]. Therefore, computer self-efficacy will be related to architects’ behavioral intentions. Based on empirical evidence from the literature, we tested the following hypotheses:

H9. Computer self-efficacy will have a positive effect on perceived ease of use.

4 Methodology

Design firms, which were either working to introduce or were using BIM tools, were selected to participate in the survey. This included the professionals at three major design firms in South Korea. These three design firms have used a BIM, or BIM tool. According to the Building SMART Korea report [38] these three firms were among the top five in terms of the number times their cumulative BIM applied in practice applications from 2009 to 2012. Also, in 2013, these three firms were listed among the top 100 design firms by the United Kingdom magazine, Building Design [39].

The demographic information of the people who answered the survey is as follows. Senior support workers, accounted for more than 80% of total participants. Average tenure was 6.5 years. Commercial BIM tools were introduced in 2003, and the first wave of adoption in AEC occurred during the mid- to late 2000s [23]. This is consistent with participants’ tenure and the length of time they used the BIM in practice. Therefore, the participants are fully capable of responding to the effects of any factors the introduction of the BIM will have on the design professional’s behavioral intention. The participants in the survey had experience using BIM tools. The most common of these BIM tools were: Autodesk AutoCAD (of those who answered, 93.83% use it), Autodesk Revit (for 72.22% of those who answered use it), and Graphisoft ArchiCAD (17.28% of those who answered use it). As found in the survey, those participants who are capable of using the most common BIM tools are those that can be targeted.

The research model contains eight constructs to ascertain the participants’ perception. The measurement items included 22 questionnaires developed based on previous research. Each questionnaire was measured on a seven point Likert scale (1 being “strongly disagree” and 7 being “strongly agree”).

197
5 Results and Analysis

Before the test of hypotheses process which is used Structural Equation Modeling (SEM) for the data analysis, the measurement model was assessed by a Confirmatory Factor Analysis that took into account reliability, convergent validity, and discriminant validity. The statistical analysis software used was AMOS 18.0.

5.1 Reliability of Results

The reliability indicates good internal consistency and reliability of items which compose constructs. Considering the reliability, it is important for measured whether set of items are useful or not [40]. Cronbach’s alpha coefficient is widely used to assess the reliability. According to Nunnally and Bernstein [41], the coefficient above 0.7 is considered acceptable. When the Cronbach’s alpha coefficient of the construct are close to 1, then the construct is internally consistent and reliable [42]. The Cronbach’s alpha coefficients ranged from 0.768 to 0.946 and exceeded 0.7 (the recommended minimum value), so constructs are acceptable.

5.2 Convergent Validity

Through the convergent validity, it can be known whether or not the construct utilities are properly configured. Three measures were assessed to test convergent validity. Factor loading refers to how the items affect each construct that is represented. Hair et al. [43] recommended a factor loading value of more than 0.5. Composite reliability determines each constructs’ reliability [44] and should be at least 0.6 [45]. Average variance extracted measures the variance between constructs [46]. Fornell and Larcker [45] recommended 0.5 as the minimum acceptable value. The values of each measurement of convergent validity was derived: factor loading ranged from 0.640 to 0.968, composite reliability ranged from 0.773 to 0.948, and average variance extracted ranged from 0.532 to 0.858—all of these values satisfied the standards. Therefore, all showed good convergent validity.

5.3 Discriminant Validity

In order to find out whether the model consists of different constructs, the process of independent verification is applied. The independence of the construct can be proven when the self-correlation of the construct is higher than the correlation of other constructs. To achieve discriminant validity, it is necessary to compare the root square of the average variance extracted (AVE) value [47]. In Table 1, the square root of the average variance extracted and diagonal value is larger than the other values in the same row and column. Therefore, each construct’s discriminant validity is significant.

Table 1. Correlation matrix and discriminant assessment

<table>
<thead>
<tr>
<th></th>
<th>BI</th>
<th>PU</th>
<th>PEU</th>
<th>TMS</th>
<th>SN</th>
<th>CP</th>
<th>TS</th>
<th>CSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU</td>
<td>0.56</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEU</td>
<td>0.53</td>
<td>0.51</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMS</td>
<td>0.14</td>
<td>0.29</td>
<td>0.15</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN</td>
<td>0.58</td>
<td>0.51</td>
<td>0.40</td>
<td>0.37</td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>0.69</td>
<td>0.67</td>
<td>0.64</td>
<td>0.06</td>
<td>0.46</td>
<td>0.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0.03</td>
<td>-0.04</td>
<td>0.06</td>
<td>0.35</td>
<td>0.05</td>
<td>-0.08</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>CSE</td>
<td>0.47</td>
<td>0.31</td>
<td>0.69</td>
<td>0.02</td>
<td>0.26</td>
<td>0.53</td>
<td>0.06</td>
<td>0.82</td>
</tr>
</tbody>
</table>

5.4 Model Fit

The degree of consensus between the model and data is assessed by fit indices [48]. As seen in Table 2, we used five different indices: the ratio of \( \chi^2 / \text{d.f.} \), Goodness-of-Fit Index (GFI), Tucker Lewis Index (TLI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA). All the model fit indices satisfied the recommended values except the GFI which is slightly less than 0.9 but very close to it.

Table 2. Evaluation of overall fitness of model

<table>
<thead>
<tr>
<th>Fitness Index</th>
<th>Recommended Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^2 / \text{d.f.} )</td>
<td>( \leq 3.00 )</td>
<td>2.000</td>
</tr>
<tr>
<td>GFI</td>
<td>( \geq 0.90 )</td>
<td>0.825</td>
</tr>
<tr>
<td>TLI</td>
<td>( \geq 0.90 )</td>
<td>0.904</td>
</tr>
<tr>
<td>CFI</td>
<td>( \geq 0.90 )</td>
<td>0.921</td>
</tr>
<tr>
<td>RMSEA</td>
<td>( \leq 0.08 )</td>
<td>0.079</td>
</tr>
</tbody>
</table>

5.5 Tests of Hypotheses

In order to verify the statistical significance and the validity of path, this research conducted testing of the 9 hypotheses. As seen in Table 3, the relationship, the standardized path coefficient, the critical ratio (t-value) and the test result of each hypothesis are demonstrated. At this point, the standardized path coefficients (\( \beta \)) point to the statistical significance and the degree of the relationship between each construct. Also, depending on the value of \( p \), the hypothesis is either supported or not supported in this test result. Most of the hypotheses were strongly supported, excluding hypotheses H3 and H8, as shown in Table 3.
Table 3. Hypothesis testing

<table>
<thead>
<tr>
<th>Hypothesized Paths</th>
<th>β-value</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>0.33</td>
<td>3.33***</td>
<td>Supported</td>
</tr>
<tr>
<td>H2</td>
<td>0.49</td>
<td>3.87***</td>
<td>Supported</td>
</tr>
<tr>
<td>H3</td>
<td>-0.05</td>
<td>-0.36</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H4</td>
<td>0.18</td>
<td>2.41*</td>
<td>Supported</td>
</tr>
<tr>
<td>H5</td>
<td>0.20</td>
<td>2.40*</td>
<td>Supported</td>
</tr>
<tr>
<td>H6</td>
<td>0.61</td>
<td>5.09***</td>
<td>Supported</td>
</tr>
<tr>
<td>H7</td>
<td>0.53</td>
<td>5.73**</td>
<td>Supported</td>
</tr>
<tr>
<td>H8</td>
<td>0.08</td>
<td>1.20</td>
<td>Not Supported</td>
</tr>
<tr>
<td>H9</td>
<td>0.44</td>
<td>3.92***</td>
<td>Supported</td>
</tr>
</tbody>
</table>

Note: *Significant at p<0.05, **Significant at p<0.01, ***Significant at p<0.001

6 Conclusions

The extended TAM presented provides insights and better understanding regarding the factors leading to the successful adoption of BIM in the design organization in South Korea. It entailed the incorporation of several constructs, such as computer self-efficacy from the individual domain, top management support and technical support from the organizational domain, subjective norm from the social domain, and compatibility from the technical domain. Thus, an important contribution of the paper is the identification of the critical factors pertaining to the full realization of the benefits of BIM adoption in the design organization.

Organization leaders should consider architects’ behavioral intention which, in turn, would support the beneficial BIM adoption. If a design organization wants to improve business performance and to increase behavioral intention, it must improve its top management support, subjective norm, compatibility, and computer self-efficacy to make it conducive for BIM adoption. The following limitations of this study will guide future work. First, the validation of results requires a larger sample of individuals. Second, the study is cross country; a longitudinal study would be advisable in order that the different countries’ adoption of BIM be compared.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2013R1A1A2A10058175).

References


[38] Building SMART Korea report. On-line: http://www.buildingsmart.or.kr/.


A Demonstration of BIM-enabled Quantitative Circulation Analysis using BERA Language

Hyunsoo Lee, Jisoo Kim, Minkyu Shin, Inhan Kim and Jin-Kook Lee

Dept of Interior Architecture Design, Hanyang University, Seoul, Republic of Korea
Dept of Architecture, Kyung Hee University, Suwon, Republic of Korea
E-mail: hyunsoolee120@gmail.com, jicoun@gmail.com, smkzzanggood@gmail.com, ihkimm1@gmail.com, designit@hanyang.ac.kr

Abstract -
This paper describes a BIM-enabled spatial analysis approach that enables us to compute specific quantitative data using BERA (Building Environment Rule and Analysis) Language and its tool. By comparing design alternatives for an actual building model, this paper also demonstrates building circulation and its associated properties of numeric data. The conventional software-driven approaches have intrinsic limitations that include limited interfaces, restricted parameters, closed rules except pre-defined rules, etc. A script language-based approach, however, is strongly necessary for various and flexible design review tasks especially for retrieving and computing specific dataset. This paper specifically demonstrates actual design review tasks as follows: 1) comparison and visualization of the differences between design alternatives in terms of their different spatial allocations; 2) analyzing numeric data result, NDBC (Numeric Data of Building Circulation), for evaluation of building circulation factors in each department of building; and as a result, 3) overall quantitative analysis and its result will be reported on given actual BIM models in order to support design decision making among various spatial allocation scenarios. For the actual demonstration presented in this paper, we modified some part of BERA language applications and defined appropriate BERA program code and Java code that are executable on given BIM models.

Keywords -
Indoor spatial analysis, automated design review, rule-checking, BERA language, building circulation, Numeric data, NDBC

1 Introduction
Taking advantages of BIM (Building Information Modeling) is becoming common in the AEC-FM (Architecture, Engineering and Construction – Facility Management) field, and the benefits of automated analysis of building design have been reported by several pilot and actual projects [1]. The analysis-purposed dataset from conventional building models usually has intrinsic limitations. They are: 1) not well-organized and mostly arbitrary; 2) qualitative, rather than quantitative; and 3) not objective but mostly subjective. On the contrast, BIM model has been: 1) standardized according to its definition scheme or at least able readily to be converted into the standardized format; and it is 2) able to be assessed the quantity based analysis in a specific design review system. 3) It is possible to automate the process of design quality check objectively. On the basis of BIM-enabled application following this new paradigm, various kinds of automated design review tasks have been developed by several parties including quantitative spatial analysis [2]. Before introduction of BIM paradigm, design evaluation tasks have been performed manually by several domain experts. As a result, it has been cost-ineffective and time-consuming. With BIM, such simulations can be provided through automated interfaces, more quickly and credibly [3]. Also, using BIM before construction in building project introduces automated checking for code compliance and constructability and circulation checking [4]. This design quality analysis is commonly implemented in GUI (Graphical User Interfaces)-based software, however, those conventional software-driven approaches have intrinsic limitations that include limited interfaces, restricted in the way of control complex and complicated parameters, defining rule-checking code, etc. [5]. Software-driven approach to the rule-checking is based on pre-defined rules and it is effective only within the boundary of such rules. This usually hinders the flexibility of the rule definition. To the contrary, script language-based design review approach is able to resolve most part of those limitations. BERA (Building Environment Rule and Analysis) Language [6] is one of the language-driven methods for analyzing indoor spatial program and
building circulation. BERA Language and its tools are capable of making quantitative results through its flexibility such as user-defined rule checking code. Especially in building remodeling project, language-driven approach can significantly contribute to compare and analyze various design alternatives focusing on spatial program and building circulation. Therefore, this paper aims to demonstrate a BIM-enabled quantitative spatial analysis using BERA Language and its tools based on an actual building remodeling project. We can analyze specific issues of the results regarding building circulation quantitatively through returned numeric data from execution of specific BERA Language program code.

2 Research Scope and Objective

The scope of this study is about BIM-based indoor spatial quantitative analysis in the way of language-driven approach specifically using BERA Language. Another focus is on introducing and demonstrating NDBC (Numeric Data of Building Circulation), as numeric dataset from various properties of building circulation. We can apply the value of NDBC as quantitative design analysis to compare different design alternatives in actual remodeling project. The NDBC values can be a quantitative determinant to aid which design alternative would be selected.

Figure 1. Overview of the proposed approach to the BIM-enabled quantitative spatial analysis

3 Approaches

3.1 BIM-enabled quantitative spatial analysis

BIM-enabled design review projects have been conducted by several challengers especially for generating spatial program review reports. The report represents the given building model’s space program in an architect-friendly form; for example, number of spaces, gross area, usable area, building efficiency and so on [8]. Before introduction of BIM, design evaluation was performed manually by multiple domain experts and was time consuming and expensive. However, report of spatial analysis such as indoor circulation properties and spatial program can be provided through automated interfaces, more quickly and quantitatively with BIM technique [9]. In the case of GSA (General Services Administration) project for courthouse in US 2007 [2, 10], the purpose of validating a spatial program using BIM is to efficiently and accurately assess design performance, relative to GSA space program requirements. They partnered with Georgia Tech and Solibri on their circulation validation efforts. Georgia Tech research team had implemented automatic circulation analysis and security checking in the conceptual design phase of the US Courthouses. They used SMC platform and IFC from BIM architectural design tools [11]. By using BIM models, GSA could automate the spatial validation process to ensure that all designs in the final concept phase accord to the spatial requirements [10]. Therefore, the automation of analysis process could save the time and improve accuracy of the results about security issue [2, 12].

3.2 Limitations of Software-driven approach

The software-driven method, as the previous way for rule checking, is a general technique in the BIM domain area. For various purpose of automated design quality analysis, some kinds of BIM-based software based on GUI has been developing and being used. However, software tools that allow for a sophisticated and complex spatial analysis of building models are not
yet available effectively [13]. The software-driven approach has not been flexible in its application and expandability due to its pre-defined rules, limited interaction, parameter-driven approach, etc., specifically in controlling complicated rule parameters.

As shown in Figure 2, only pre-defined parameters are available in actual design review tasks. The process of problem solving for developing new functionalities is cost and time-consuming. Therefore, it is necessary to develop the method which is easy-to-use for novice users to develop and facilitate rule-checking methods.

![Figure 2](image)

**Figure 2.** An example of BIM software-driven approach to the design analysis: only pre-defined parameters in a table regulate specific design rules such as security or distance conditions. [6]

### 3.3 A script Language-driven approach and BERA Language

One of the countermeasures for limitations of software-driven method, language-driven approach has been introduced for resolving these problems. There is BERA (Building Environment Rule and Analysis) language [6] as a script language-driven method. It aims to analyze the building environment as its literal meaning, so that it doesn’t produce or modify building object model. BERA has strength in interactive automated review, user-driven rules and analysis, various execution tools in syntactic and semantic way.

On these concepts of BERA, there are 3 main points 1) it is domain-specific so that used in experts’ stand in AEC field; 2) it is based on programming language; 3) it uses meaningful symbol marks such as =, <, >, as Rule language. Using script language which deals with space objects and those properties relations belonged in IFC format model, users can analyze or compare the data they want to get for results and simulation in quantitative way. Since the BERA Language supports various commands for complex design guide, users can implement rule-checking at specific part intuitively. The approach such as BERA which is using scripts language enables to check rules in the flexible and effective way [6].

### 4 Quantitative Data of a Circulation Path and definition of NDBC

#### 4.1 Quantitative data of a circulation path

The purpose of this paper is especially focusing on analyzing the spatial circulation properties that could not be checked specifically by customized definition in previous way such as software-driven method. A fundamental issue in building circulation is specific path people take when moving from some spaces to another [9]. People using or occupying the indoor space of building usually evaluate the efficiency of space qualitatively. On the issue of Building circulation, there have not been exist numeric data reports from the way of design analysis. The analysis of spatial circulation in rule-based systems assists users to define and apply customized rules that check conditions of the building by executing on a given model and returned reports [14, 15]. BERA Language as a rule-based script language is used in this demonstration and it is able to convert qualitative contents to quantitative visualized report data. The following BERA code is an example about get Path (Lobby to Room) in building spatial program and there are its returned data properties from BERA program code. Figure 3 is a process chart getting returned quantitative data using BERA code.

![BERA Program Example](image)

**BERA language-enabled derivation of circulation properties**

**Example Returned Data**

- Number of Paths
- Walking Distances
- Number of Turns
- Sum, Average, etc.

The BERA language-enabled derivation of circulation properties extracts circulation-related data for generating NDBC.

**Figure 3.** An example of property data set from the circulation path between ‘Lobby’ and ‘Room’
Each returned properties of paths have numeric data regarding user defined BERA code. The returned data contains the property of area, height, volume, window area, number of space, number of window and so on, so that user can apply those results regarding to their specific purpose of design analysis. Figure 4 shows a visualized circulation graph from user-defined method mentioned as below using IFC test model.

```java
Path p = getPath("Lobby", "Department");
get(p);
```

Figure 4. Visualized path graph on the test model using BERA code

Figure 4 describes the all of possible circulation paths and those properties from the path of Lobby to each department A, B, and C. The result shown as Figure 4 was not possible in previous rule-checking software-driven approach. However, the study about whether those numeric information data contains extended meaning or not is out of scope in this paper. BERA and its application only produce and suggest the numeric data in quantitative way.

4.2 Definition of NDBC (Numeric Data of Building Circulation)

NDBC (Numeric Data of Building Circulation) is a quantitative building circulation properties data using language-driven approach. In this section, we provide a formal definition of NDBC derived from returned data of building circulation using BERA language and its application. As shown in Figure 3, data of number of Paths, walking distances, number of turns and summation of distances and number of turns) extracted from returned report would be NDBC. NDBC especially based on BERA Language and JAVA code that is able to calculate various and complex properties of circulation what users require for specific issues in building model. There are some meaningful formulas of NDBC as following below.

1. \((n)P(S, T)\)

   Where \(S\) is start space and \(T\) is target space, and \((n)P\) is the number of Path instances.

   \[ S = \{ s_i \mid 1 \leq i \leq n; n = \text{number of start spaces} \} \quad (1) \]

   \[ T = \{ t_i \mid 1 \leq i \leq n; n = \text{number of target spaces} \} \quad (2) \]

   \[ P(S, T) = \{ p_i \mid 1 \leq i \leq n; n = \text{number of paths} \} \quad (3) \]

   \[ (n)P(S, T) = (n)S \cdot (n)T \quad (4) \]

   \(P\) denotes a set of number of building circulation paths, \(S\) is a set of start space objects and \(T\) is a set of target space objects of building. As noticed above, start space and target space should be defined to get number of path. If there are sets of \(S = \{ (s_1), (s_2) \} \rightarrow (n)S = 2\) and \(T = \{ (t_1), (t_2), (t_3) \} \rightarrow (n)T = 3\), then set \(P\) can be derived as \(P(S, T) = \{(s_1, t_1), (s_1, t_2), (s_1, t_3), (s_2, t_1), (s_2, t_2), (s_2, t_3)\} \rightarrow (n)P(S, T) = 6\).

2. \(SW(P(S, T))\)

   Where \(SW\) is the summation of total path walking distances of Path set \(P\)

   \[ SW(P(S, T)) = \{ W_1 + W_2 + W_3 + \ldots + W_n \} = \sum_{n=1}^{N} W_n \quad (5) \]

   \(W\) denotes walking distance from start space \(S\) to target space objects \(T\). Because the each walking distance is according to the number of start and target spaces and the number of paths, analyzing comparable data of walking distance is able to be miscalculated. However, the summation of walking distance data can be compared per each different kinds of target space object.

3. \(ST(P(S, T))\)

   Where \(ST\) is the summation of total number of turns of Path set \(P\)

   \[ ST(P(S, T)) = \{ T_1 + T_2 + T_3 + \ldots + T_n \} = \sum_{n=1}^{P} T_n \quad (6) \]

   The number of turns in circulation is also significant issue in building circulation analysis. The summation of number of turns connotes that how many times people should turn their path while going to target space from start space.
4. **ASW (P(S, T))**

Where A is average of the summation of total walking space per the number of path

\[
ASW(P(S, T)) = \frac{SW(P(S, T))}{(n)P(S, T)}
\]  

(7)

The summation of total walking distance has each different value according to target spaces and those numbers. When analyzing the walking distance equally, there should be average walking distance value per the number of path of its target spaces.

5. **AST (P(S, T))**

Where A is average of the summation of total number of turns per the number of path

\[
AST(P(S, T)) = \frac{ST(P(S, T))}{(n)P(S, T)}
\]  

(8)

The summation of turn of numbers of paths also should have its average value per paths to analyze with other kinds of paths.

5  Demonstration

5.1 Application of indoor spatial analysis with NDBC

This section demonstrates a visualization of NDBC for a specific spatial program and circulation generated by BERA language and its tool. The graph of circulation path is the way of visualization and it can show not only the visualized graph but also prehensible and approximative value of walking distance or number of turn in path. Figure 5 is an example of visualized path graph in one floor of test IFC model using BERA language in SMC and rule set code. s1 is start space “Hall 4” and t1 to t12 are target spaces include “Interior” nomination. The paths from s1 to each t1 to t12 are noticed as p 1 to p12. Following code below is BERA rule-set used.

```java
Space aa {
    Space.Floor.number = 5;
    Space.name = "Hall4"
}
Space bb {
    Space.Floor.number = 5;
    Space.name = "Interior"
}
Path p = getPath(aa, bb);
get(p);
```

Through NDBC definition, we can analyze circulation path in Figure 5 as following,

- \(S = \{s1\}\)
- \(T = \{t1, t2, \ldots, t12\}\)
- \(P(S, T) = \{p1, p2, \ldots, p12\}\)
- \((n)P(S, T) = (n)S \times (n)T = 12\)
- \(SW(P(S, T)) = 488.9\)
- \(ST(P(S, T)) = 20\)
- \(ASW(P(S, T)) = 40.74\)
- \(AST(P(S, T)) = 1.6\)

In the visualization of NDBC as shown in Figure 5, one floor area that has one start space and 12 target spaces included in one department, we can figure how many paths from start space to target space and following data properties such as \(SW(P(S, T))\), \(ST(P(S, T))\), \(ASW\), \(AST\). These numeric data have differentiated method from previous rule-checking software by language-driven approach and quantitative results. This study is focusing on circulation and paths and those numeric data value, however, the application of language driven design analysis can be expanded to other issue of building design not only in domain of circulation.

5.2 Demonstration of actual remodeling project apply to NDBC

We utilized NDBC for evaluating an actual building remodeling project to compare design quantitatively between before and after design, located in Seoul, Korea. Hanyang University Human ecology building in Seoul, Korea had been remodeled since in December, 2012 and it was completed in March, 2013. There are 3 departments according to majors (Food and Nutrition, Clothing and Textile, and Interior design) in this building. Design analysis about spatial relocation should be considered whether reasonable, efficient and impartial or not. Therefore, design alternatives of remodeling project should be analyzed and compared with previous design in a quantitative way. This approach generates an objective outcome such as space
quality review reports. To prevent waste of time and expense, design quantitative evaluation is required in preconstruction stage. Through objective and numerical report of design review, the appropriate design can be selected among other design alternatives according to purpose of design.

Figure 6. An actual test case model demonstrated in this paper: a college building of Hanyang University, Seoul, Korea

Figure 7 shows spatial allocation alternatives according to 3 major departments in this building. In the view of the remodeling project, outer wall structure is not changed while department allocation has changed.

Figure 7. Planning the allocation of three different departments for the actual model.

The changed relocation of design alternative 1 and 2 is focusing on the spaces related to 3 majors in this building. For example, the classes among the same major should be located within a range at least 2 floors so that user can save unnecessary walking distance. Based on the above plans shown as figure 7, we apply example of same BERA code and JAVA to analyze NDBC value for circulation of alternative 1 and 2. Figure 8 describes the execution circulation analysis on the scope of the department ‘A’ through prepared IFC models for design 1 and 2 and Solibri Model Checker as IFC platform. The collaboration code of BERA and JAVA below is used for execution of circulation analysis, NDBC.

Figure 8. Execution of department A circulation analysis in design alternative 1(left) and 2 (right) using BERA Language and SMC

```
 Path p = getPath("Entrance", "Food");
 get(p);
 java;
 double s = 0.0;
 int st = 0;

 int i = 1;
 for (Path p : Path){
    if (p.distance != null) {
       s = s + p.distance;
       st = st + p.numberOfTurn;
       print(i + " : " + s);
       ++i;
    }
 }
```

Figure 9. Chart of NDBC result of department A in department 1 and 2.

We can analyze the circulation path and also compare NDBC as showed in Figure 9. The chart shows each data value of number of target spaces, number of path, summation of walking distance, summation number of turn, and average of SW and ST. The overall graph flow in comparison chart, Design #2
is above design 1 for each parts of NDBC contents. Department A has two more target spaces in design 2 than 1, ASW and AST value that is about numeric value per each path of design 2 is also higher than 1. Table 1 is the NDBC results of all major department spaces and circulation properties using BERA rule-set.

Table 1. Comparison between two design alternatives using spatial quantity data enabled by BERA language and its tool.

<table>
<thead>
<tr>
<th>Design</th>
<th>Path</th>
<th>Department</th>
<th>T</th>
<th>P</th>
<th>SW</th>
<th>ST</th>
<th>ASW</th>
<th>AST</th>
<th>Total ASW</th>
<th>Total AST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt #1</td>
<td>P_Start</td>
<td>Entrance 1, 2</td>
<td>A</td>
<td>17</td>
<td>34</td>
<td>2790.89</td>
<td>389</td>
<td>82.08</td>
<td>11.44</td>
<td>400.12</td>
</tr>
<tr>
<td></td>
<td>P_Target</td>
<td>Dept A</td>
<td>B</td>
<td>25</td>
<td>50</td>
<td>6533.59</td>
<td>1436</td>
<td>130.67</td>
<td>28.72</td>
<td>1080.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dept B</td>
<td>C</td>
<td>27</td>
<td>54</td>
<td>10118.21</td>
<td>2724</td>
<td>187.37</td>
<td>28.72</td>
<td>1080.24</td>
</tr>
<tr>
<td></td>
<td>Floor Level</td>
<td>B1, Level 1, Level 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Level 2, Level 3, Level 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Level 4, Level 5, Level 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Floor Level</td>
<td>Level 1, Level 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Level 3, Level 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Level 5, Level 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Area</td>
<td>9570.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alt #2</td>
<td>P_Start</td>
<td>Entrance 1, 2</td>
<td>A</td>
<td>19</td>
<td>38</td>
<td>3614.21</td>
<td>477</td>
<td>95.11</td>
<td>12.55</td>
<td>496.33</td>
</tr>
<tr>
<td></td>
<td>P_Target</td>
<td>Dept A</td>
<td>B</td>
<td>23</td>
<td>46</td>
<td>7495.24</td>
<td>1782</td>
<td>162.94</td>
<td>38.73</td>
<td>1090.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dept B</td>
<td>C</td>
<td>23</td>
<td>46</td>
<td>10960.97</td>
<td>2936</td>
<td>238.28</td>
<td>63.82</td>
<td>1090.24</td>
</tr>
</tbody>
</table>

Figure 10. Data Chart of ASW and AST for the comparison between design alternatives

We can get NDBC of circulation path properties respectively as shown in Table 1. Each design plan has different number and area of relevant spaces, so that there is difficulty for comparing between design alternatives with total summation of circulation path data. Therefore, ASW and AST are influential factors for comparing overall design alternatives irrespective of contained the number of spaces. Figure 10 is the chart of comparison ASW and AST regarding to each design alternatives #1 and #2. Both ASW and AST value in design #2 are higher than design #1 and it means that circulation path from entrance to spaces of three different departments in design #2 has longer walking distance and more number of turn than those of design #1.

6 Summary

In this paper, we proposed and demonstrated a quantitative analysis approach to the comparison between design alternatives using NDBC, as we introduced the abbreviation of the term Numeric Data of Building Circulation. This paper surveys and demonstrates that script language-driven design analysis has necessity and efficiency compared to previous software-driven approach through BERA Language. There is script language between accessibility and usability of software and sophistication and applicability of programming language. The script language-driven design analysis enables: 1) accurate query of target building object, 2) flexibility of redefining rule set regarding the complex building design.
quality, and 3) reporting about implementation of design analysis and its result. Based on those issues, we defined NDBC that quantitative returned data from BERA language and its tool and demonstrated actual remodeling project case. It is possible to suggest objective value of building circulation for helping to choose most efficient design among the other design alternatives. This study focuses on building circulation centrally from among the various building design issues, however, its expendability and applicability can be able to affect succeeding study of quantitative design analysis.

Acknowledgment

This research was supported by a grant (13AUDP-C067809-01) from Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

References

Graph-based Representation of Building Circulation
With the Most-Remote Points and Virtual Space Objects

Jisoo Kim\textsuperscript{a}, Hyunsoo Lee\textsuperscript{a}, Minkyu Shin\textsuperscript{a}, Jin Won Choi\textsuperscript{b} and Jin-Kook Lee\textsuperscript{a}

\textsuperscript{a} Dept. of Interior Architecture Design, Hanyang University, Seoul, Republic of Korea
\textsuperscript{b} Virtual Builders Co. Ltd., Seoul, Republic of Korea
E-mail: jicoun@gmail.com, hyunsoolee120@gmail.com, smkzzanggood@gmail.com, jchoi@vbuilders.co.kr, designit@hanyang.ac.kr

Abstract
This research aims to develop a graph-based representation of building circulation as an extended version of the metric graph structure named UCN (Universal Circulation Network). There are several different methods and algorithms to represent pedestrians’ indoor circulation for architectural purposes. The UCN has introduced a BIM-enabled approach to measure walking distances between different space objects using door-to-door connection network. In this paper, we focus on the extended development of the metric graph algorithm to expand its use scenarios such as measuring the exact distance of fire egress. There are two major issues we have encountered in this research and development: 1) finding the most-remote point in a given space objects to measure the distance of fire egress and 2) handling virtually subdivided space objects. This paper describes specific algorithms to resolve this problem on top of the structure of UCN. The major outcome of this development is the extended metric graph structure with the virtual space objects and the most-remote point. The implementation and demonstration depicted in this paper will show how the UCN structure can be enhanced in terms of its capability for visual representation as well as its extensibility of circulation-related analysis.

Keywords
Building Information Modeling (BIM), Building Circulation, Metric Graph, Universal Circulation Network (UCN), Virtual Space, Most-remote Point

1 Introduction
Building circulation has to be evaluated even in early phase of design because its impact is of importance in terms of the quality and the performance of building [1]. The issues of circulation sometimes involve critical aspects of building design such as complex security and spatial allocation problems in courthouse buildings [2]. To analyze the building circulation by manual used to be time-consuming and error-prone, therefore an automated and reliable way of such analysis has been strongly required. For evaluating circulation paths in automation using BIM models [3], we need an abstraction of circulation paths on top of the given BIM models to represent, calculate, and analyze topological network of building circulation, and it is the UCN (Universal Circulation Network) [4,5,6,7]. UCN [7] uses graph representation and implemented effective algorithm for circulation analysis. This paper mostly refers to the UCN [7] and aims to extend its structure of circulation paths by defining two additional aspects of circulation representation: 1) finding the most remote point, and 2) with virtual space objects. For measuring exact fire egress distances, the most-remote point based calculation is necessary [8]. We also have to take care of handling virtual space objects for generating the most-remote point edges on top of UCN structure. In this paper, therefore, we studied how to handle virtual object for checking circulations and find the most-remote point for fire egress as a way of extension of the UCN graph.

2 Review of the Universal Circulation Network
Several approaches to the BIM-enabled visual representation of building circulation have been implemented and examined [7, 15]. In terms of technical aspect, they can be classified as follows: 1) topological graph, 2) center-line-based metric graph, 3) Kannala’s metric graph [9] and 4) the UCN (Universal Circulation Network) graph [7]. As described in [7], above graph #1 represents only topological relations between space objects, and #2 and #3 do not reflect human behavioral patterns. The UCN provides a precise method to measure walking distances with appropriately visualized
metric graph structure reflecting human behavioral patterns. The UCN-based representation of building circulation, thus, has been adopted in several software tools that are dealing with circulation related issues.

General features in UCN can be summarized as follows: 1) building-object oriented graph representation, 2) application of buffer distances from the wall regarding human scale, 3) application of the shortest path finding algorithm [10], 4) the most visible path, and 5) door-to-door graph visualization using wall-bounded space objects.

Above features are key aspects of the UCN graph implementation so that it reflects realistic circulation patterns and provides precise method of measuring walking distances. In practice, the implementation of the UCN graph has contributed to BIM software developments such as Solibri model checker (SMC) and others in terms of their visualization of building circulation paths and measurement of walking distances.

The main goal of this paper is to extend the structure of the UCN graph. Therefore, we focus on the limitations found in reviewing current implementation of the UCN graph, and they can be listed as follows: 1) its graph edges can be represented only on top of the physical space objects that are bounded by wall objects. It cannot deal with virtually divided space objects that are commonly found in actual BIM models. 2) Its graph edges can be connected only between door-to-door to represent inter-space circulation. It does not generate graphs that are located inside of space objects that have only a single door object.

As described in #1, the UCN graph omitted handling virtually subdivided space objects. Virtually subdivided space objects are beyond the scope of the UCN graph which is focusing on overview of building circulation or space programming not on specific agent’s path of inside space. Dependent upon requirements of the design guide or specific needs of space circulation, some circulation graphs should be represented in virtually subdivided space object. In this paper, therefore, we notice the requirement of dealing circulation of virtually subdivided space objects.

In case of #2, due to the limitation of the UCN graph structure, it cannot represent the graph in purpose of finding the most remote point from the door object in specific spaces for measuring the egress distance defined in general Fire Code. In this perspective, the objective of this study is extending the UCN graph to virtually subdivided space objects by finding the most remote point in rooms to overcome such limitations.

3 Finding the most-remote point in a given space object

3.1 Extension of the UCN metric graph

It is controversial that the open plan office is not the best idea for office layout, but it is broadly adopted by designers. Consideration of circulation graph in this kind of spaces plays an important role in space programming and quantitative design evaluation. In case of open plan office model, many departments or teams are existed in one big open space area, separated by virtual space object boundaries. In real world, however, office areas do not have every wall and door object for separating departments or teams. Based on door-to-door connection of the UCN graph, it is not available in this kind of spaces besides any open plan designed model. In this case, more consideration is needed on circulation graph algorithm. To focus on these limitations, we aim to extend the UCN graph by handling virtual boundaries of given building model and figuring out circulation graphs in open plan designed model. The circulation graph we propose can be useful to the multi-functional space oriented model.

3.2 Definition of the most-remote point

The most-remote point (MRP) refers to the point which is located in the furthest point from the door according to the definition in Fire Code, for measuring egress distance accurately. Choi et al. [8] represented the concept of MRP as Outermost Node. This study focuses on whole structure of evacuation regulation checking system. In this study, the researchers draw the path by finding outermost node in space using UCN graph structure; however outermost node path does not focus on algorithm of the graph also not deal with virtual space object boundary. Including design guide and other building codes introduce the travel distance rules and measurement method mostly refer to the NFPA 101, and here are some selected regulations [11,12,13,14].

- The maximum travel distance from the most remote point in any room or space to the center of a door opening directly on an open exterior space shall not be greater than the limits.
- Travel distance is the total distance in a building an occupant must travel before reaching an exit. Travel distance is measured from the most remote point in a room to a point where the nearest exit begins.

In conclusion, the most remote point from the door, which is the furthest point in spaces, is used for a measurement of fire egress distance [12]. In this study, the MRP graph is based on metric graph structure.
addition to the UCN graph, and it generates the exact walking paths for checking and visualizing the fire egress related regulations.

4 The most-remote point graph in physical space objects

In the UCN graph, for visualizing the path, its algorithm uses only concave points on space object boundaries and door center points. Convex points, also the components of the space object boundaries, are useful to find the most remote point from door objects which are properties required by the N.Y. Building Code for measuring fire egress route distances. The most-remote point should be one of the convex points on given space object boundaries, and this paper suggests the algorithm for finding such points from any given BIM models.

As depicted in figure 1, the dotted lines denote to the buffer lines of space object boundaries. To consider a size of an agent itself, the MRP is on the buffer line of the space object boundary not on the edge of a space object boundary. The most-remote point graph starts form the door center point as same way as the UCN graph and finishes to one of the convex point of a buffered space object boundary which is the furthest from the start point. The distance of the graph, in contrast, should be the shortest one among many possible paths (including free curves) following the Dijkstra’s shortest path finding algorithm [10]. To find the furthest convex point, we measure the distance to each convex point of the buffered space object boundaries from the door center point of the space. In case of the rounded corner spaces like figure 1.3), which is commonly found in actual BIM models, the MRP graph represents in the same way as other models, but only different when we find out the convex points of the buffered space object boundaries. We segmentize the curve to equal angle and set of points come along would be convex points which are the materials of the MRP. The angle for segmentizing can be different dependent upon the model and the requirements of detail in evaluation. The process of finding the furthest convex point on a space object boundary is as follows.

1) Draw the buffer line of each space object boundary
2) Find concave points, convex points and a door center point of each buffered space boundary
3) Draw the metric graph of each convex point using concave point (if necessary) and a door center point
   - Rooms having corner or column: using concave points as many as necessary in this order; Door center point \( \rightarrow \) concave points (n) \( \rightarrow \) convex point
   - Rooms which doesn’t have corner or column: directly connect to a door center point and a convex point
4) The longest distance graph based on the shortest path algorithm is the most-remote point graph

Following the process above, we define an operator \( MGraph() \), which defines the MRP graph, as follows.

\[
MGraph(s_b, d_i, c_x) \rightarrow G_{M_i}
\]

One of the parameter \( s_b \) refers to space object boundary of the model, \( d_i \) is derived from a door object of the model, and \( c_x \) refers to a convex point. Output \( G_{M_i} \) is the most remote point graph which generates from those parameters above.

5 The most-remote point graph in virtual space objects

5.1 Application of the most-remote point graph in a room with virtual space boundaries

Virtually divided space objects technically refer to the areas divided by room separators, not by typical space boundary elements such as physical wall and doors. The implementation of the UCN graph omitted the representation of circulation paths for those virtually subdivided spaces such as open office areas, but they are commonly found in actual models. In this paper, we have represented how the circulation paths can be visualized between virtually subdivided spaces without physical walls by using a graph structure as an extended version of the UCN graph. This approach and implementation enabled us to represent and analyze circulation paths within an open space such as partitioned space or cubicle area that are commonly found in office buildings.
Figure 2. An example of a series of vertices that generate the most-remote point graph in case of virtually divided space object

Two cases of virtually subdivided spaces can be classified: 1) with door objects and 2) without door objects. The graph representation on the first cases can be drawn the same as physical spaces, the second cases, however, should be handled in different way. This graph representation technically require specific points acting like door center points which we decide to a mid-point of a virtual space object boundary as shown in figure 2. In virtually divided space objects, mid-point of the virtual space object boundary can be an average distance for efficient checking. Especially in cases of the space having corners, columns or rounded corner space object boundaries, are using average distance for the MRP graph for more reasonable circulation analysis.

\[
\begin{align*}
A &= m \\
D &= n \\
\mathcal{G}_M &= m \times n
\end{align*}
\]

Figure 3. Examples of the most-remote point graph in a room: 1) a single space object, and 2) three space objects that are separated by virtual space object boundaries

Figure 3 shows how the MRP graph can be varied by the existence of virtual space object boundaries. According to one of the key aspects of the MRP graph, starting from a door center point, we can notice the number of the MRP graph can be determined by the number of doors in a space object. As shown in figure 3, however, not only the number of door objects but also the number of virtual spaces can be a prime determinant of structuring the MRP graph. Thus, the number of the MRP graph can be calculated as below.

\[
\begin{align*}
A &= m \\
D &= n \\
\mathcal{G}_M &= m \times n
\end{align*}
\]

\[
\mathcal{G}_U = \frac{n(n-1)}{2} \quad (n \geq 2)
\]

\[
\mathcal{G}_S = \mathcal{G}_M + \mathcal{G}_U
\]

\[
\mathcal{G}_S = n(m + \frac{n-1}{2})
\]

where \(A\) denotes the number of the area (1), \(D\) denotes the number of the door object (2) and \(\mathcal{G}_M\) denotes the number of the MRP graph (3). In case of the model which has more than 2 doors can be drawn as the UCN graph which is based on door-to-door connection. However, the UCN graph and the MRP graph are in different sectors in perspective of utilization also the implementation. Thus, the number of circulation graph in one space is like below.

5.2 The most-remote point graph in three or more virtual space objects

The most remote point graph of two virtual areas in one space follows section 5.1. On the other hand, in case of the model which has three or more virtual areas should be handle in different way because the edges of overlapping space boundaries cannot be merged to one when touching line cope only partial space boundary. Another process is required in this case to visualize the MRP graph.

Figure 4. 1) Finding mid-points on every virtual space boundary, 2) generating a collection for all edges to define the MRP graph, and 3) determining three final MRP graphs

As shown in figure 4, the graph follows the order as
follows; convex point \(\rightarrow\) virtual door center point \(\rightarrow\) virtual door center point \(\rightarrow\)(n)…\(\rightarrow\) door center point. In this process, the metric graph could be detoured, thus, it would be the longest one as shown in figure 4.2). It represents entire set of edges can be derived from given spatial condition, and the final edges should be satisfied by the ‘most remote’ point as well as the ‘shortest’ path finding algorithm. The algorithm needs two steps: 1) Find the shortest graph in each convex point and then, 2) Compare to each graph and measure the distance for finding the longest one. Below figure 5 is the workflow of the algorithm to find the MRP Graph.

![Figure 5](image.png)

Figure 5. Overview workflow for finding the most remote point and determining the MRP graph.

In addition to the workflow shown in figure 5, a specific precondition is required: ‘The graph shouldn’t pass again same virtually subdivided area’. If the workflow does not include that precondition, then the graph should be detoured and have repetitive form, as well as it would not be the shortest path. The last part of the diagram, ‘Connect to convex point and door center point considering concave points’ and ‘Find the shortest path’ is following the UCN graph as described in this paper.

Figure 6 depicts an extended version of the UCN graph generation process with virtually subdivided space objects and the most-remote points. Two new operators, \(MGraph()\) and \(CMGraph()\), have been added on diagram of the UCN graph for handling virtual space boundaries and the MRP-based circulation paths. The operator \(MGraph()\) is in charge of finding the most remote point in each space object and determining its graphs, as well as finding the mid-points on each virtual space object boundary. As denoted in section 4, one of the parameters of \(MGraph()\), \(d_i\) can be the mid-points on virtual space object boundary in case of virtually subdivided area. \(MGraph()\), thus, applied in virtually subdivided areas recognized as physically bounded space objects. The operator \(CMGraph()\) applied to combining two graphs; the most remote point graph and space metric graph, so that the MRP-based circulation graph is the derived graph from \(CMGraph()\).

\[
CMGraph() = \{MGraph(), SGraph(), WGraph()\}
\]

![Figure 6](image.png)

Figure 6. Overview of the extended UCN generation process, with virtual space objects and the most-remote points.

### 6 Demonstration of the MRP graph

#### 6.1 Examples of the most-remote point graph in various spaces

To demonstrate more cases, figure 7 shows the MRP graph in the model with many virtual space boundaries. As shown above, we can find the 6 MRP graphs from one set door in one open plan office divided by 5 room separators and the graphs are drawn by the algorithm we described in this paper. There can be an issue in purple colored virtually subdivided area with the location of mid-point. To follow the workflow of figure 5, we consider only in purple colored area for the first step of finding the MRP graph. The purple colored area has two mid-points, so that there are 8 possibilities for the MRP graph. Among 8 possibilities, we should find the longest path with following the shortest path finding algorithm and then proceed to the next step which is the pink colored area. This process is of importance especially the area subdivided in many virtual space boundaries like the model of figure 7.
6.2 Extending UCN with in-space graph in BIM model

By combining the UCN graph and the MRP graph structure, the MRP-circulation graph can be represented. If a pair of start and end space objects have been assigned, we are able to generate the MRP circulation graph and measure a total distance of an agent’s evacuation distance. The UCN graph can be extended to rooms with the MRP graph so that the MRP graph is an add-on of the UCN graph. For representing an entire set of building circulation using graph structures, both the UCN graph and the MRP graph set are required. The UCN graph is in charge of generating door-to-door graphs, and the MRP is in charge of handling in-space graph edges especially for connecting the most remote point.

7 Summary

In this paper, we extended the metric graph structure based on the former implementation of UCN. To find out the most remote point in a specific room and to handle virtually divided rooms without physical wall objects are the major outcomes of this paper. Virtually subdivided rooms still have several space objects in BIM model, but the former study and implementation of the UCN graph only deals with physically bounded space objects. As a result, virtual space objects have been neglected in the circulation graph generation, including another important end-node: the most remote point (MRP). The MRP graph can be utilized in assessing Fire Code and measuring exact walking distances in case of fire as the code regulated. The algorithm and the implementation process depicted in this paper suggest a generic approach to the issues based on the UCN graph, and demonstrate how to facilitate them for a better representation of building circulation. In this paper, we notice that the UCN graph has much more potentials in future work and to be extended in many parts as we suggested in this paper. We tried to show the significance of the BIM-enabled assessment to be extended for further requirements such as Fire Code that used to be missing in the former implementation. We hope many studies motivated in this extension of the UCN graph and broaden the scope of the BIM applications.
Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2011-0028498)

References


Abstract -

This paper aims to develop a mechanism for visualizing indoor environment data set sensed by diverse sensors. The focus of this research is the development of an integrated visualization process using accumulated data from sensors, rather than the qualitative analysis of indoor environment. In this regard, we use a test-purposed and versatile toolkit that is cheaper, smaller and more controllable than conventional tools. We could collect indoor environmental data set composed of sequential numeric data so as to use them as a given parameters for visualization. We inspected three major issues in the process: 1) indoor temperature data of a specific room can be collected at a second interval; 2) such a data set can be varied by subdivided spots of interest using multiple toolkits; 3) as a result, the collection of data is regarded as one type of parameters for visualization on top of the room’s floor plan, e.g. a sudden change of sequential numbers. The toolkit and control mechanism described in this paper can be summarized as following modules: 1) the sensor module; 2) the data collection module; 3) the data retrieval module; 4) the spatial data module; 5) the sensor location module; 6) the data visualization module. A demonstration has been introduced in this paper for evaluating the integrated visualization approach.

Keywords -
Sensed data, Numeric data, Parameters for visualization, Indoor temperature changes

1 Introduction

People spend most of their daily lives indoor, and this simply explains why indoor temperature is one of the important factors to deal with as an environmental aspects [1, 2]. Therefore controlling comfortable indoor condition is essential for the people to keep their emotional and physical states healthy. This may be resulted in better work productivity [3, 4]. For dealing with further issues of the physical environment, this paper aims to figure out an integrated approach to visualize such conditional factors using indoor temperature changes. The baseline of this approach is to measure the temperature and analyze given conditions [5].

Conventionally, to measure and collect the indoor environmental data such as temperature data set used to be inconvenient and expensive due to several reasons [6]. The multi-use environmental sensors, however, became affordable and easy to control recently, and this enabled us to measure the broad range of indoor environments by using such low-priced and controllable sensor devices [7]. In the meantime, as the data set acquired by widely installed environmental sensors increased significantly, the need for making effective use of the data has become of importance [8, 9, 10].

An integrated approach and mechanism to visualize indoor environmental data on the floor plan is introduced as an effective method for analyzing the data of indoor environment. Consequently, to verify the feasibility of the proposed mechanism in this paper, we implemented an integrated approach to the visualization of indoor temperature changes using several software and hardware modules.

2 Scope and Approach

We reviewed several research using the sensor to manage indoor environment and found out that those focus on representing sensed data as the chart [11, 12, 13, 14].

This paper aims to demonstrate a specific visualization mechanism on top of the floor plans based
on the sensed changes of indoor temperature as one of the indoor environmental data sets. This is one of the graphical representations on the existing building geometry using sensor devices and sequential processing mechanism. The entire process from sensing indoor environment to processing and visualizing acquired data can be subdivided into three parts. Each part is composed of sub-modules. Figure 1 shows the parts and included modules of this research and development.

| Input Data Part | • Sensor Module  
| Building Model Part  |
| • Building Model Module  
| • Sensor Network Module |
| Visualization Part  |
| • Visualization Module |

Figure 1. Three major parts of the implementation and demonstration depicted in this paper.

After dealing with the three parts of visualization process and the functions of each module inside, this paper will conduct a demonstration on an actual space. The target environment element to be visualized is hourly variation of temperature change.

3 System Architecture

The system architecture for visualizing the indoor environment is composed of three parts. 1) Building model part, 2) Input data part, and 3) Visualization part.

In the building model part, as it literally means, the 2D floor plan or 3D building model which is the background for visualization is loaded. Also, all the influence factors of target environment and the reference point of the installed sensors in the building model should be included for the more accurate analysis. The input data part deals with the environment data. In this stage, the environment data are sensed by installed sensors and collected, retrieved, and processed in order to generate the data to be visualized. The outcome data generated in the Input data part are transferred to an infographic and combined with loaded building model in Visualization part.

4 Implementation

The data flow during the entire visualization process is summarized in Figure 3. Input Data Part is composed of three modules: 1) Sensor Module, 2) Data Collection Module, and 3) Data Retrieval Module. These three modules measure the target environmental element and generate the outcome data to be visualized.

The Building Model Part consists of two modules: 1) Building Model Module, and 2) Sensor Network Module. The Building Model Part loads the building model such as 2D floor plans or 3D building model and determine the reference point of installed sensors. The input data and loaded building model are combined by Visualization Module in Visualization Part.

4.1 Input Data Part

4.1.1 Sensor Module

The Sensor Module measures the target indoor environment by multiple sensors and generates sets of raw data. Each sensor senses the numeric value of target indoor environment at an interval of time, and the data are collected into a set of raw data in the Data Collection
Module. Generating enough number of datasets by installing enough number of sensors is important because the more input data makes the more precise and broader visualization outcome.

There are various kinds of target indoor environment can be measured such as temperature, humidity, brightness, air quality, movement, etc. In this paper, temperature is the target indoor environment to be visualized.

4.1.2 Data Collection Module

The Data Collection Module collects the target environment data measured by Sensor Module. In this module, the data sensed by installed sensors each time are accumulated in real time.

Depending on the compatibility between the installed sensors and visualization software, there are two possible methods to collect data. In case the sensors are compatible with Visualization s/w, the data are able to be collected directly on the Visualization s/w. Otherwise, if the sensor is incompatible with Visualization s/w, the collected data should be transferred to a suitable format which is compatible with Visualization s/w such as csv, DB, Excel file.

4.1.3 Data Retrieval Module

The Data Retrieval Module selects particular part of data among the collected raw dataset in order to generate input data. But before selecting particular part of data, defining the input data should be preceded. Depending on the definition of input data, the target and algorithm for data retrieval are decided.

Not only selecting particular part of data from dataset of measured environment data, it is also possible to decide what environment element to visualize when there are plural indoor environment elements collected by installed sensors.

4.2 Building Model Part

4.2.1 Building Model Module

The Building Model Module loads the building geometry which becomes a back-layer of the visualization. The loaded building model is used as reference for designating the boundary of sensor network layer and analyzing visualization outcome.

The use of loaded building model regarding designation of the boundary of sensor network layer is described on the Sensor Network Module. When the building model is combined with the visualization outcome, the building model becomes the basis of analysis of the visualized indoor environment data. To do that, the building model should include all the factors that influence the target indoor environment within the space.

4.2.2 Sensor Network Module

The Sensor Network Module generates the sensor network layer which becomes the basis for visualization on building model. Based on the loaded building model, the shape and the size of boundaries of the sensor network layer are defined. The generated sensor network layer will be visualized in Visualization Module, being mapped with the input data processed by Data Retrieval Module. Each sensor location, which is going to be mapped with the input data, is able to be acquired by IPS (Indoor Positioning System). There are several possible IPS system like using Wi-Fi, GPS antenna, Geo-Magnetism. Without IPS, the sensor location can be designated manually comparing given building model with the sensor location in actual space.

4.3 Visualization Part

4.3.1 Visualization Module

Visualization Module transforms the numeric input data into visual information using sensor network layer. Each input data of an installed sensor is connected to each sensor points on the sensor network layer. Connected with the input data, the sensor points can be expressed in any type of graphical image including color or geometry. In this step, the visualization method is decided depending on the building model and the definition of input data, considering how to visualize the interested environment data in the most intuitive way. When the visualization method is decided and the input data is connected to each sensor points, the sensor network layer transforms to a certain form of visual information, according to the decided visualization method and the retrieved input data.
5 Demonstration

We have demonstrated an integrated approach for visualizing the indoor temperature changes of a laboratory room located in building of College of Human Ecology, Hanyang University to examine feasibility of proposed mechanism.

Figure 5 represents the overall flow of the demonstration including the software/hardware used in each stage.

5.1 Input Data Part

5.1.1 Measurement of Temperature

Arduino Uno Board and RHT03 sensors were used as instruments measuring the temperature, which are cheap and easy to control. RHT03 is a penny-sized temperature/humidity sensor which is enable to measure from -40 to 80 degrees C temperature range and also has +/- 2% error range in humidity [15]. Arduino Uno Board is a microcontroller board controlled by Arduino Language [16]. Using the language, it is possible to control how the sensors operate and in what form the data will be collected. In this paper, the sensors operated with 60 seconds interval and temperature was collected as Celsius degree.

5.1.2 Converting and Collecting the Measured Data

The measured environment data was collected in an Excel spreadsheet due to the issue of compatibility with Rhino-Grasshopper on which the environment data will be visualized. An Excel plug-in software named PLX-DAQ (Parallax Data Acquisition tool) was used for exporting the measured data from the Sensor Module into an Excel file [17].

Each dataset measured by the sensors is collected in each spreadsheet and a dataset includes temperature, humidity and the time the sensor operated.

5.1.3 Temperature change Data Retrieval

The implementation stages between data retrieval and visualization have been implemented using Rhino-Grasshopper in this demo. The Grasshopper is compatible with Excel spreadsheet file using a plug-in program named ghExcel[18]. The datasets consist of time, humidity and temperature values in 60 seconds interval and the target environment element is temperature.

In order to generate the final input data from collected datasets, the input data need to be defined before selecting particular part of data. The input data was defined as the subtraction of two temperature value between two moments at an hour interval. As one interested moment is selected, the temperature values of the selected moment and 1 hour later are retrieved. After retrieving two temperature values to be compared, the temperature difference is generated for the final input data. This demo used Rhino-Grasshopper “slider” function for selecting interface for retrieval.
5.2 Building Model Part

5.2.1 Importing Floor Plan

As the building model for this demo, the floor plan is imported to Rhino 3D. The size and layout of the space are shown in figure 9. In addition for the sensors and all the factors affecting temperature such as heaters or openings are expressed too.

5.2.2 Generating Sensor Network Layer

Previous to the visualization stage, sensor network layer has been made above the floor plan. The sensor network layer becomes the basis when transferring the input data into visual information. Pointing at the corners of the floor plan’s outline, the boundary of sensor network layer was defined. On the boundary of layer, each sensor points were created on each sensor location. The sensor points were designated on the layer manually based on the comparison of floor plan with the actual space, because IPS, by which the sensor location is acquired, is not used in this demo.

Figure 9. An actual floor plan for the demonstration

5.3 Visualization Part

5.3.1 Visualizing Input Data

For the visualization of given input data above it, the sensor network layer is divided into grid segments with each sensor point as the center. The divided layer forms a flat surface which is able to be connected with the input data on each sensor point. The given input data are assigned to each sensor point as z value of Rhino3D. As the sensor points move to the z axis direction, the surface is transformed to an uneven surface.

At the same time, it is possible to express the temperature change as color as well as the geometry. As shown in figure 12, according to the distance to the floor plan the surface is colored with red-to-blue color gradation. As a result, input data assigned as z value are transformed into visual information like geometry, color.

Figure 12. Visualization mechanism of converting z value to color gradation

5.4 Result and Analysis

Figure 13 shows the final visualization outcome on Rhinoceros3D viewport. In top view the temperature
changes is shown as colors overlapped on the floor plan. In this way it is easier to realize the state of indoor environment and factors to influence the temperature changes. In isometric view, the degree of temperature changes appears in the form of geometry which looks like a colored uneven plane. The three-dimensional visualization of an indoor environment element is very intuitive compared to two-dimensional or numeric information.

The input data of the left visualization result is the temperature changes between 9am ~ 10am. The outcome shows that the overall temperature has been raised centrally around the heater. On the contrary, the input data of right visualization result is the temperature changes between 9pm ~ 10pm. The temperature decreased rapidly during 1 hour with the heater and window as center because the heater was turned off and the window was opened.

As a result, we find out that the visualized environmental data are more intuitive and easier to figure out state of indoor environment than numeric data. Combined with the building model, in addition, the visualized data can be used for recognition of the factors affecting to target environmental element. To visualize not only temperature changes but also other indoor environment, we need to develop the visualization mechanism, considering various kinds of available sensors, their network and visualization methods.

Acknowledgment

This research was supported by a grant (13AUDP-B068892-01) from Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

References


Application of Dijkstra’s Algorithm in the Smart Exit Sign

Jehyun Cho*, Ghang Lee*, Jongsung Won* and Eunseo Ryu*

*Dept. of Architectural Engineering, University of Yonsei, South Korea
E-mail: jkun86@naver.com, glee@yonsei.ac.kr, quietman111@gmail.com, eunseo4773@naver.com

Abstract
Previous studies on automated fire-egress guidance systems have focused on providing the shortest path information from a specific person at a certain point to the closest exit mostly using a mobile device. This study aims to develop a Smart Exit Sign system that can detect dangerous areas in real time and direct evacuees to the shortest safe evacuation path by dynamically changing the direction signs to the safe egress. The challenge was to provide the shortest safe egress to any evacuees at any point. We have developed a sensor network and algorithm that could exclude unsafe paths and calculate the shortest safe path from multiple starting points to multiple exit points based on Dijkstra’s algorithm—the most commonly used algorithm for finding the shortest path. The validity of the proposed system was tested through simulations of test cases.

Keywords -
Shortest path algorithm; Evacuation system; Exit sign

1 Introduction
Navigating the indoor spaces of large and complex buildings such as shopping malls is challenging. This makes building occupants often disoriented while they explore a building. Yet, this issue is more problematic during an emergent situation such as fires rather than general situations. During an emergent situation, exit signs may be helpful. However, traditional exit signs in buildings usually have a fixed direction sign towards exits and may direct evacuators to dangerous paths where a fire has broken out. Consequently, in this type of situation, individuals cannot dynamically change directions reflecting the fire hazard. In order to respond to this issue, we have developed a smart exit sign that can dynamically change direction signs and guide evacuators to safe evacuation paths. This paper presents an Automated Direction Setting Algorithm (ADSA), which was developed as part of the Smart Exit Sign system. The ADSA was designed to dynamically change the direction signs on a smart exit sign—reflecting the fire situations—and indicate the proper directions towards safer evacuation paths.

Previous research has studied evacuation path based on the shortest path algorithms such as Dijkstra’s algorithm, Floyd-Warshall Algorithm, and A* algorithm. The main goals of those studies were to find the shortest evacuation path and guide a certain person to the path by showing the path on a mobile phone or on other systems. However, in an emergent case, such as fire, most evacuees may not have or may not be able to download the smart device application that is required to visualize the shortest evacuation path on their mobile devices.

The eventual goal of this study is to develop a Smart Exit Sign system that detects unsafe areas and calculate the shortest safe path in real time and dynamically change its direction sign toward the shortest safe path. In the first step, this paper presents an algorithm and a simulator that can calculate and visualize a network of smart exit signs, which can guide any evacuees at any point in a building to the nearest exit via a safe route. The algorithm, developed based on Dijkstra’s algorithm, which is one of the famous shortest path algorithms, calculates safe evacuation paths from multiple starting points while directing individuals toward the nearest exits.

This study was conducted in the following order. First, through review of related literature, we determined the characteristics needed to make the algorithm for the system effective. Then, we developed an algorithm and a simulator. Finally, we tested the applicability and validity of the algorithm through simulations of test cases.

2 Literature Review
Many studies have been conducted to develop systems that could provide indoor navigation information in an unknown space during an emergent situation.

In a study conducted by Kobes et al., it examined how people determined evacuation paths, during fire related situations, through a series of tests that were...
conducted in a hotel building at night [1]. According to the research, 56.3% of people determined evacuation paths depending on exit signs when there was no smoke present; however, 81.8% determined an evacuation path depending on exit signs when their visibility was impaired due to smoke. These results show that the role of exit signs is of the utmost importance when individuals are put in adverse situations that require an exit strategy.

However, because the current exit signs installed in buildings have fixed directions, there is a possibility that the signs can lead the evacuators to dangerous area according to the study conducted by Choi [2]. Through a perception test, Choi discovered that the ambiguity of directions on exit signs could make it difficult for evacuees to recognize the correct direction towards an exit [2]. On the other hand, when animated direction images were used on exit signs, the number of people who followed the correct evacuation path doubled.

Jang et al. also argued that the evacuation guidance system with a fixed direction image could lead evacuees to dangerous areas by illustrating a simple case [3].

In regards to exit signs, Kim et al. proposed an ‘artificial intelligent directional escape light system’ based on the Floyd-Warshall algorithm [4]. This system was designed to suggest directions for the nearest exits and also to indicate dangerous points where a fire has occurred, but has not been implemented.

Similarly, another group of researchers, Kim et al. [5], proposed a similar smart exit sign system. The system was composed of wireless sensors, had used Dijkstra’s algorithm, and was tested through simulations in virtual space. However, in the Kim’s simulation test, exit signs at several points could not find the correct direction towards the nearest exit.

3 Automated Direction Setting Algorithm

We developed the Automated Direction Setting Algorithm (ADSA) for the smart exit sign system, which offers evacuation direction for a number of evacuees at any location in a building on the basis of Dijkstra’s algorithm. Dijkstra’s algorithm is a representative of the shortest path algorithm. It is a type of greedy algorithm, which is used to search for the maximum value or the minimum value.

Dijkstra’s algorithm iteratively calculates the distance between a starting point to the other points in the network in turn; and, depending on which path is used to travel from the starting point to another point, the distance will vary. The starting point has the distance value of 0. The algorithm replaces the distance value stored at each point (node) from the starting point to each point with a new value if a shorter path is found.

Dijkstra’s algorithm has been used in various fields, such as Google Maps, navigation systems for automobiles, traffic engineering, robotics, and industrial engineering. Compared to Floyd algorithm, which is another popular shortest path algorithm, Dijkstra’s algorithm is better suited for calculating the shortest path in a huge, complex building because the time complexity of Dijkstra’s algorithm is $O(n^2)$ while the time complexity of Floyd algorithm is $O(n^3)$. It means that the amount of time taken by Dijkstra’s algorithm is shorter than that of Floyd algorithm as input size goes higher.

3.1 Additional Features of ADSA

There are two main differences between plain Dijkstra’s algorithm and ADSA for the smart exit sign system. The following subsections describe these differences in detail.

3.1.1 Exits as Starting Points

Basic concept of Dijkstra’s algorithm is checking the shortest paths from the single starting point to all the other points. However, unlike general Dijkstra’s algorithm, ADSA assumes that the exits are the starting points for the calculations of the evacuation paths. The reason being is because Dijkstra’s algorithm does not distinguish the exit from other points. Thus, by specifying exits as the starting point—the exits get the distance value of 0 and the modified algorithm—ADSA automatically recognizes the point with 0 value as an exit. Moreover, ADSA points the direction sign to the exit with the smallest distance value from each point. If a new exit closer from a specific point is found, the direction sign at the point is changed towards the new exit.

For example, Figures 1 to 5 show a topological representation of a building plan, which has one exit and Dijkstra distance values at each point. In the plan, if we randomly pick point A as the starting point as shown in Figure 1, we know that the shortest distance from points A to C (D_{A,C}) is 22. However, we do not know which path is the shortest path and, thus, cannot show which direction evacuees at point A should go: from point A to point B or to point D.
On the other hand, if we assume that the exit (point C) in a plan is the starting point as shown in Figure 2, we can determine the direction to the shortest path by comparing the accumulated distance values from the exit (point C) to each neighbouring point: points B and E in this example. Since point B has a smaller accumulated distance value than point E, our algorithm leads evacuees at point A to point B.

However, normally, there exist multiple exits in buildings. So, in order to find the nearest exits from each point, the algorithm should be able to compare the distance from each point to multiple exits as shown in Figure 3. The algorithm chooses the smaller values. Figure 4 shows the final calculation results.

### 3.1.2 Reconfiguration of a Path Network

Once the one of the signs which comprise the whole system notices any fire situations by heat detectors, smoke detectors or missed signal from the adjacent sign of breaking because of the fire, Originally established networks in a building must be reflect the current situation before suggest the direction change.

The ASDA reconfigures the path network and recalculates the directions from each point to the shortest safe path by excluding dangerous paths when a fire breaks out. Some corridor will be set as impassable according to the location of the fire. Figure 5 shows an example where a fire occurs at point F. The algorithm eliminated point F and all of the paths that are connected to point F and recalculated the distance.
3.2 Implementation of ADSA in a Simulator

We developed a Smart Exit Sign simulator using the ADSA. The simulator is operated in the following three steps (Figure 6).

The first step is to extract a topological configuration of a building plan composed of nodes and lines from a drawing or from a Building Information Model (BIM). Particularly, the nodes are classified into four types: “exit,” “intersection,” “dead-end,” and “exit sign.” The purpose of distinguishing “exit” from the other node types is to run the ADSA by using exits as starting points. And, “dead-end” is defined only for the graphical purpose in the simulation software.

Secondly, users specify the actual distance between any two selected nodes so that the simulator can place a small icon that represent an exit sign at every 15m, following the Korean national emergency management law that says exit signs must be placed at every 15m or less (National Emergency Management Agency, 2012, South Korea). This step does not have to do with the calculation of the shortest safe path, but it is an important step to show where Smart Exit Signs should be placed. The distances between all the other nodes can be automatically calculated.

Finally, each smart exit sign’s distance is calculated from to the nearest “exit.” Then, each exit sign automatically changes the direction sign to show the shortest and safest path. If a fire occurs, the third step is repeated after eliminating the paths where a fire breaks out.

The simulation program was developed using C# computer programming language. Table 1 displays the icons used in the simulator.

<table>
<thead>
<tr>
<th>Menu</th>
<th>Function</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node</td>
<td>Exit</td>
<td>Specify 'Exit's</td>
</tr>
<tr>
<td></td>
<td>Intersection</td>
<td>Specify 'Intersection's</td>
</tr>
<tr>
<td></td>
<td>Dead-end</td>
<td>Specify 'Dead-end's</td>
</tr>
<tr>
<td>Line</td>
<td>Corridor</td>
<td>Specify 'Line's</td>
</tr>
<tr>
<td></td>
<td>Parameter</td>
<td>Save the actual length value</td>
</tr>
<tr>
<td></td>
<td>AP</td>
<td>Arrange the Signs</td>
</tr>
<tr>
<td></td>
<td>Fire</td>
<td>Simulate the fire</td>
</tr>
</tbody>
</table>

4 Simulation Results

We ran simulation tests to check the applicability and validity of ADSA through two test cases.

The first case was a hypothetical maze map which was a good sample to test, debug, and refine the algorithm (Figure 7).
The second test case was a subway station in Seoul, South Korea. The plan had 15 separated corridors, 6 intersections, and 8 exits.

We manually checked whether the direction of each exit sign in the simulator pointed towards the nearest and safest exit. First, it was confirmed that the simulator automatically set the initial directions of the exit signs towards the nearest and safest exit. Then, we tested three different fire scenarios (fire on a corridor, fire on an intersection, and fire on an exit) and checked to see if the signs were correctly redirected to the new safe path (Figure 8).

Figures 9 to 11 show a zoomed-in view of three cases and redirection results: fire on a corridor (Figure 9), fire at an intersection (Figure 10), and fire at an exit (Figure 11), respectively. All three cases showed the correct changes.

Through the tests, we confirmed that the simulator could correctly show the shortest safe path and redirect the exit signs avoiding impassable corridors.

5 Conclusions

This paper presents the ADSA, which was developed as a part of the Smart Exit Sign system based on Dijkstra’s algorithm. ADSA finds and leads evacuees at any point in a building to safe exits. To simplify the calculation of the direction from multiple starting points to multiple exits, ADSA set the exits as the starting point unlike the traditional application of Dijkstra’s algorithm. ADSA could find the direction from a point to the nearest exit simply by comparing the accumulated distance values of neighbouring points. The algorithm directs the exit sign towards the neighbouring point with a smallest accumulated distance value. If a fire occurs, the direction of exit signs is dynamically recalculated in real time. Through test cases, we confirmed that ADSA could be applied in an actual Smart Exit Sign system. However, our current system only considered distances. The system can be improved by considering other influential factors in a real fire situation, such as spreading direction of fire, smoke, and crowdedness. We intend to reflect these factors to our system in later studies. Also, through additional test cases, we plan to rigorously test whether they are any exceptional cases and check the validity and applicability of the Smart Exit Sign system. Development and test of an actual Smart Exit Sign system in a real building is another
future intent.

6 Acknowledgement

This research has been supported by the Ministry of Knowledge Economy, Korea, under the national HRD support program for convergence information technology supervised by the National IT Industry Promotion Agency (NIPA-2010-C6150-1001-0013).

7 References


Smart Housing for Reducing Energy Use in Single-Households, Emphasizing Effective Energy Management

S.J. Park\textsuperscript{a}, M.J. Kim\textsuperscript{b,c} and S.Y. Kim\textsuperscript{c}

\textsuperscript{a}Department of Interior and Environmental Design, Keimyung University, Republic of Korea
\textsuperscript{b}Department of Housing and Interior Design, Kyung Hee University, Republic of Korea
\textsuperscript{c}Department of Energy Engineering, Keimyung University, Republic of Korea

E-mail: sjpark@kmu.ac.kr, mijeongkim@khu.ac.kr, energy@kmu.ac.kr

Abstract

Single-households, one of the new housing types to emerge recently, seem to experience difficulties in the management of energy use due to the short duration of their residence. However, there has been little study on effective energy management in single-households. If we can find solutions that reduce energy consumption through effective management, the total energy consumption across this housing type could be decreased, even though the reduction in energy use per single-household may be small. The main purpose of the research is to provide appropriate solutions for the effective management of energy use for single-households by applying smart technologies. For the implementation of smart housing, primarily we identified single-households’ patterns of electricity use through our investigation. We categorized six representative types of single-households based on the results of a questionnaire survey, and extracted the patterns of electricity use in each type of single-household through interviews. Based on the results, smart housing solutions for single-households would be provided by emphasizing the effective management of electricity use. By maximizing smart spaces in single-households, sustainable and healthy housing would be realized in terms of energy saving. Further, this research would be a basis for enhancing the quality of housing environments for single-households.

Keywords - Smart Technology; Single Households; Energy Consumption, Energy Management

1 Introduction

One third of the energy consumed in Korea is attributable to the built environment, and households in residential buildings contribute significantly to the magnitude of energy consumption in Korea. Significantly, the number of single-households has recently increased (now 23.8% of the total number of households) and these single-households appear to experience difficulties in the management of energy use due to the short duration of their residence. However, little study has taken place on effective energy management strategies for single-households. Our research starts by questioning how the energy use of single-households could be managed effectively through smart technologies. In general, the energy consumption in households includes electricity and gas for heating, cooking and hot water. However, this research focuses on the electricity consumed by single-households because the amount of the electricity consumed in single-households has become two and half times that of four-person households \cite{1}. In particular, by analyzing lifestyle patterns in residential buildings, it was found that the single-households use more electronic appliances than the four-person households. Our research was mainly motivated by these findings on the single-households. We believe that the total amount of energy consumption could be decreased, although the reduction in energy use in each individual single-household may be small.

The main purpose of the research is to provide appropriate solutions to the effective management of energy use for single-households by applying smart technologies. Our assumption is that smart housing might have the potential to reduce energy consumption by allowing the efficient use of energy in single-households. To facilitate the implementation of smart housing, primarily the patterns of electricity consumption in single-households should be identified and analyzed in order to address the features and problems to be considered for the effective management of energy use. Thus, we investigated single-households’ patterns of electricity use with a focus on home appliances and lighting. Through the investigation, we could provide customized solutions to the different types of single-households in terms of electricity use. By emphasizing behaviors in daily life, this research would provide single-households with practical solutions that allow them to consume electricity more
efficiently, eventually leading to a reduction in their expenditure for energy.

2 Related Works

2.1 Single-Households and Energy Use

The single-household is a new housing type that should be considered as important in housing studies because its effect on the housing industry, peoples’ lives, and further, all of society might be immense. It has been reported that single-households are more interested in taking advantage of any conveniences available to support their lifestyle, and so employ more electric devices for their daily life. Above all, they try to save time and effort using smart appliances when they have to do housework [2, 3]. Thus, many researchers have paid much attention to the patterns of energy use by single-households. For example, researchers in the Energy department in the state of North Rhine-Westphalia in Germany investigated the patterns of electricity use, targeting 0.4 million people through the website “Stromcheck für Haushalte”. They found that the patterns of electricity in single-households differ from those in other types of households in terms of twelve home appliances. However, the occupant has the predominant effect on the energy use in single-households. As much as 80% of the measured variation may depend on the inhabitants’ behavior [4, 5]. There are studies that showed that energy consumption is mainly influenced by the lifestyle of their occupants, such as schedule profiles of lights and electrical devices [6]. Thus, they tried to identify residents’ energy-lifestyles because the actual conditions of energy use first need to be understood. They collected data on their energy-saving behaviors or use of electrical appliances according to individual schedules [7, 8]. Further, some researchers found lifestyle factors reflecting social and behavioral patterns associated with air conditioning, laundry, personal computer and TV usage [9].

2.2 Energy Savings and Strategies

There have been, in general, three approaches to the reduction of energy consumption in residential buildings [10]: implementing effective building performance to minimize heating and cooling loads, developing efficient home appliances to consume low amounts of energy, and promoting ‘energy-conscious’ behaviors among residents. For example, many researchers have tried to design or develop buildings and appliances and lighting. The general features of the subjects are as shown in Table 1.

<table>
<thead>
<tr>
<th>Subject (age)</th>
<th>Gender</th>
<th>Housing size</th>
<th>Job</th>
<th>time spent at home</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (65)</td>
<td>F</td>
<td>79.2 m²</td>
<td>Nil</td>
<td>16 hours</td>
</tr>
<tr>
<td>B (66)</td>
<td>M</td>
<td>66 m²</td>
<td>Nil</td>
<td>15 hours</td>
</tr>
<tr>
<td>C (36)</td>
<td>F</td>
<td>62.7 m²</td>
<td>Worker</td>
<td>12 hours</td>
</tr>
<tr>
<td>D (34)</td>
<td>M</td>
<td>42.9 m²</td>
<td>Worker</td>
<td>13 hours</td>
</tr>
<tr>
<td>E (21)</td>
<td>F</td>
<td>33 m²</td>
<td>student</td>
<td>10 hours</td>
</tr>
<tr>
<td>F (20)</td>
<td>M</td>
<td>49.5 m²</td>
<td>student</td>
<td>7 hours</td>
</tr>
</tbody>
</table>

2.3 Recent studies that showed that energy consumption is...
By investigating subjects with different ages, genders and jobs, we tried to identify the patterns of electricity use that reflected their lifestyles, to find a way to save energy. All subjects lived in multiplex housing, but the housing sizes and the time spent at home differed.

We investigated at what time and for how long they generally use home appliances and lighting on weekdays. We did not include the electricity use at weekends because it could be quite different from that of weekdays. In addition, we considered twenty-three appliances that are used often, but consume a great deal of electricity. We excluded air conditioners and electric fans from the appliance list because they are related to seasonal factors. Rather than the patterns of electricity use that vary seasonally, we tried to find the patterns of electricity use that may vary according to the number of appliances and the amount of time operated. The electricity consumption per hour was calculated by referring to data sources derived from ENERGY STAR, a joint program of the U.S. Environmental Protection Agency and the U.S. Department of Energy. The standard electricity consumption for the appliances is as shown in Table 2.

The electricity consumption for each appliance was calculated by multiplying the time used by the average electricity consumption. For example, the electricity consumption for a refrigerator was calculated according to the number of opening behaviors of the door because opening behavior consumes a lot of electricity. The time spent opening the door was regarded as ten seconds. Especially, the lighting equipment was divided into fixed lights in rooms and movable lights such as floor lamps and desk lamps. The reason for differentiating the lighting is that many residents tend to leave fixed lights on while they stay at home. It was expected that by controlling such behaviors through smart management, the electricity use could be decreased. The electricity consumption for lighting was calculated taking into consideration the time spent at home.

In this study, we suggested a solution for managing electricity use from the aspect of the elasticity of demand in electricity prices. The TOU (Time of Use) and flat electricity rate are as illustrated in Figure 1.

Table 2. Standard Electricity Consumption for the Appliances

<table>
<thead>
<tr>
<th>Appliances</th>
<th>Watt/ Hour</th>
<th>Watt/ Min. (Sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hair Dryer</td>
<td>1538.00</td>
<td>25.63</td>
</tr>
<tr>
<td>Coffee Machine</td>
<td>1500.00</td>
<td>25.00</td>
</tr>
<tr>
<td>Microwave</td>
<td>1500.00</td>
<td>25.00</td>
</tr>
<tr>
<td>Iron</td>
<td>1100.00</td>
<td>18.33</td>
</tr>
<tr>
<td>Toaster</td>
<td>1100.00</td>
<td>18.33</td>
</tr>
<tr>
<td>Vacuum Cleaner</td>
<td>650.00</td>
<td>10.83</td>
</tr>
<tr>
<td>Clothes Washer</td>
<td>425.00</td>
<td>7.08</td>
</tr>
<tr>
<td>Espresso Machine</td>
<td>360.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Blender</td>
<td>300.00</td>
<td>5.00</td>
</tr>
<tr>
<td>LCD TV</td>
<td>213.00</td>
<td>3.55</td>
</tr>
<tr>
<td>Video Game Player</td>
<td>195.00</td>
<td>3.25</td>
</tr>
<tr>
<td>Refrigerator *</td>
<td>100.00</td>
<td>1.67</td>
</tr>
<tr>
<td>Monitor</td>
<td>150.00</td>
<td>2.50</td>
</tr>
<tr>
<td>Computer</td>
<td>120.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Curling Iron</td>
<td>90.00</td>
<td>1.50</td>
</tr>
<tr>
<td>bedroom/living room lighting</td>
<td>20.00</td>
<td>0.33</td>
</tr>
<tr>
<td>bedroom stand lighting</td>
<td>60.00</td>
<td>1.00</td>
</tr>
<tr>
<td>bathroom lighting</td>
<td>10.00</td>
<td>0.17</td>
</tr>
<tr>
<td>kitchen/dining room lighting</td>
<td>10.00</td>
<td>0.17</td>
</tr>
<tr>
<td>desk lamp/floor lamp</td>
<td>25.00</td>
<td>0.42</td>
</tr>
<tr>
<td>Stereo</td>
<td>60.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Laptop</td>
<td>50.00</td>
<td>0.83</td>
</tr>
<tr>
<td>Printer</td>
<td>45.00</td>
<td>0.75</td>
</tr>
<tr>
<td>Aquarium</td>
<td>30.00</td>
<td>0.50</td>
</tr>
<tr>
<td>Clock Radio</td>
<td>10.00</td>
<td>0.17</td>
</tr>
<tr>
<td>Rice Cooker</td>
<td>1000.00</td>
<td>16.67</td>
</tr>
<tr>
<td></td>
<td>95.00</td>
<td>1.58</td>
</tr>
</tbody>
</table>

![Figure 1. TOU and Flat Electricity Rate](image)

Generally, the demand decreases as a product rises in price; the demand curve defines the elasticity around a demand price. The elasticity can be expressed as:

$$
\varepsilon = \frac{\Delta d}{\Delta p} \cdot \frac{d_p}{p_0}
$$

$\Delta d, \Delta p$ indicate the variation in demand and price, and $d_p, p_0$ indicate the price demand at the reference point. Therefore, the change in the electricity consumption caused by the change in the electricity price, at time $t$, can be expressed as:

$$
\Delta d_t = \varepsilon \cdot \Delta p_t
$$
4. Result

4.1 The Patterns of Electricity Use by Six Representative Subjects

Pattern graphs of the energy use per subject are as shown in Figure 2 to Figure 3.

The 66-year-old male mainly uses electricity in the morning and evening, using 9278.28wh on average. This electricity use is the second highest; only the 66-year-old female subject uses more. The main times that electricity is consumed are between 5:00 and 11:00 and between 17:00 and 22:00. In particular, between 6:00 and 7:00, electronic appliances such as a microwave, a LCD TV and a rice cooker are used simultaneously and intensively. Also, between 9:00 and 10:00, electronic appliances such as a clothes washer, iron, and a vacuum cleaner are used at the same time when doing household chores. To prepare dinner, he starts to use electricity in the late afternoon. Although he does not use many electronic appliances, the total electricity consumption is not small because the rice cooker is set to warm for 24 hours, similarly to the refrigerator. He does not have many lighting devices and does not turn them on for long periods of time. In particular, the lighting in the kitchen and dining room is rarely used. The stand lighting in a bedroom is an incandescent lamp, which consumes 25wh per hour, so the electricity consumption for lighting is not low.

The 34-year-old male mainly consumes electricity in the morning and late evening, using 8987.04wh on average. The main times that electricity is consumed are between 6:00 and 8:00 and between 19:00 and 23:00. Compared to the two oldest subjects, he consumes electricity intensively in the early morning and late evening because he is at work during the day. Before going to work at 7:00, he uses a toaster, a blender and a curling iron. After returning home, he uses electricity intensively for dinner and household chores. Between 20:00 and 22:00, he uses a clothes washer and a vacuum cleaner with other appliances such as a rice cooker and a microwave. In particular, he plays video games and enjoys leisure activities such as watching TV and web-surfing. Therefore, the electricity consumption is highest in the late evening. He also uses the rice cooker’s warm option continuously (24 hours). He does not use many lighting devices, but he often leaves the desk lamp on while he sleeps at night, so the amount of electricity used for lighting is rather high.

The 65-year-old female uses 9328.5wh on average, which is higher than all other subjects. The main times that electricity is consumed are between 7:00 and 11:00 and between 17:00 and 22:00. In the morning, she uses a LCD TV, in addition to a microwave used to prepare breakfast. In the afternoon, she uses a microwave and a coffee machine for lunch, and an iron and a clothes washer for chores. Compared to the other subjects, she uses high amounts of electricity for clocks, radios and an aquarium that is on 24 hours. In particular, electronic appliances such as a microwave, a LCD TV and a rice cooker are used simultaneously between 6:00 and 7:00. In addition, electronic appliances such as a clothes washer, an iron, and a vacuum cleaner are used when doing household chores between 9:00 and 10:00. On average, she uses quite a lot more electricity for lighting than the other subjects. The lighting in bedrooms or a living room is used in the afternoons due to her long-sightedness. She also has a habit of turning on stand lighting, consuming large amounts of electricity continuously at night.
The 36-year-old female mainly uses electricity in the early morning and evening, using 5452.89wh on average. The main times that electricity is consumed are between 5:00 and 8:00 and between 18:00 and 23:00. Before going to work at 6:00, she uses an espresso machine and a blender for breakfast, in addition to a hair dryer. Compared to the other subjects, she does not use many electrical appliances to cook breakfast. Rather, she consumes comparatively high amounts of electricity using a hair dryer and an iron to care for her appearance. After returning home from work, she uses electricity intensively to prepare dinner and do household chores. In particular, between 19:00 and 22:00, she uses several appliances such as a rice cooker, a clothes washer and a vacuum cleaner simultaneously. Differing from the male subjects, she cancels the rice cooker's warm option before going to bed, which implies that she intends to reduce electricity consumption. Regarding lighting devices, she leaves lighting on in bedrooms, the living room, the kitchen and the dining room continuously in the evening, and uses an incandescent light for the standard lamps in bedrooms. The amount of electricity consumption for lighting is comparatively high.

The 20-year-old male undergraduate student uses 4641.29wh electricity on average mainly in the morning and evening. The main times that electricity is consumed are between 7:00 and 9:00 and between 18:00 and 24:00. Before going to university at 8:00, he uses a coffee machine, a microwave and a toaster to prepare breakfast, in addition to a hair dryer. Compared to other subjects, the time spent using a hair dryer is longer because she does her hair twice a day, morning and evening. After school, she does activities intensively such as dinner preparation and household chores. Therefore, the electricity consumption is high in a day. In particular, from 18:00 to 19:00, she uses appliances such as a clothes washer and a vacuum cleaner, but cancels the rice cooker's warming option after she uses it for dinner. Thus, compared to subjects who leave the rice cooker on the warm option continuously, the amount of electricity used by the rice cooker is comparatively low. Regarding lighting devices, he leaves the lighting on continuously in the bedroom and living room before going to bed. She also turns the incandescent desk lamp on and leaves it on until the next morning, so electricity consumption is high.

4.2 Comparison of Patterns of Electricity Use by Six Subjects

Figure 8 illustrates the patterns of energy use of the six subjects. Largely, the electricity use of the six subjects are distributed in the morning and evening except the 65-year-old female. Five subjects use electronic appliances intensively during the time spent...
There are a few differences in the time periods when the electrical appliances are used among the six subjects. Firstly, the 65-year-old male and the 66-year-old female, who spend longer at home than the other four subjects, use electronic appliances moderately, while the other four use electronic appliances intensively and simultaneously. Secondly, the 34-year-old male and the 36-year-old female consume electricity intensively in the morning and evening. They use various electronic appliances for breakfast and dinner preparation, household chores and leisure activities simultaneously. Thirdly, the 20-year-old male person and the 21-year-old female persons show similar patterns of electricity use, consuming in the late morning and early evening compared to other subjects because they have more flexible time schedules.

The patterns of energy use are different according to gender. Firstly, electricity consumption is high for the females because the females use more electronic appliances than the males. Also, the females use electronic appliances such as hair dryers, irons and bedroom standard lamps for longer periods of time. However, there are a few other differences between males and females because females attempt to reduce energy consumption by turning off the warming option in rice cookers. Males use significant amounts of electricity by keeping the rice cooker on for 24 hours continuously, and leaving desk lamps and lighting on continuously in the living room and bedrooms. The result illustrates that the patterns of electricity use are different according to the time period. The patterns of electricity use reflect relatively different lifestyles including habitual behaviors, electronic appliances used, the times that they are used, and gender.

5. Electricity Consumption Based on Each Electricity Rate

This chapter presents how energy consumption is changed by the elasticity of demand in electricity prices when TOU is applied to the pattern of six subjects’ electricity use. According to the Korea Energy Economics Institute, the price elasticity of the demand for electricity by residence in Korea is 0.025 [22]. Thus, if the flat rate currently applied to each residence is changed to the TOU rate, the electricity consumption in the same time period will decrease or increase by responding to the changes in the price. For the six subjects, each variation applied to the electricity rate is as shown in Table 2.
6. Conclusion and Future Directions

Although significant improvements in energy efficiency have been achieved in home appliances and lighting, the electricity consumption in households has continued to increase. The reasons for this could be associated with an increased degree of comfort and level of amenities [20]. To understand energy consumption in single-households, the patterns of energy use of electronic appliances and entertainment loads were investigated with six representative subjects emphasizing different lifestyles.

Based on the patterns of the energy use, we expect that customized smart homes could be implemented to reduce energy consumption in each single-household. Rather than developing smart homes from a technology-oriented approach, residents’ lifestyles, energy-related behaviors and habits should be considered as essential for the effective management of energy consumption in households. To reduce electricity consumption in single-households, the patterns of electricity use of the electronic appliances and lighting should be understood. Based on this understanding, a smart system that monitors the residents’ patterns of electricity use could be provided to stimulate them to be conscious of potential energy savings. Through the monitoring, the residents become aware of unnecessary energy consumption in daily life, leading to the avoidance of such thoughtless behavior. For example, compared to females, males consume a great deal of electricity carelessly, for example, leaving the warm option on in a rice cooker and not turning off standard lamps; thus, there is not much difference in the total amount of electricity used even though males use less electronic appliances. Further, control systems customized to the patterns of the electricity use could be developed enabling smart time-scheduling for electricity consumption in single-households. In this study, a time-scheduling system driven by TOU is proposed for the electricity consumption of the six subjects. The system sets different prices for electricity for each time zone; thus, the amount of energy consumption by electric appliances, which are used regardless of cheaper priced time zones, will be decreased and managed more effectively. For example, with the exception of the two elderly persons, the remaining four subjects consume electricity intensively in the morning and evening. If we employ a smart clothes washer or a robot vacuum cleaner that can be operated at a customized time, the electricity use can be distributed over the time period when less electronic appliances are being operated.

Although the sizes of the dwellings are not large, and therefore the number of lighting appliances is not high, electricity consumption for lighting is high because most subjects turn on the lights and leave them on continuously in the kitchen, living room and bedrooms in the evening. It seems that several activities occur simultaneously in those spaces, thus residents carelessly leave some of the lighting on. Further, females keep lighting on in the daytime because of their long-sightedness. Appropriate illumination appears to be different in the daytime and evening. Smart lighting systems could reduce unnecessary electricity consumption by controlling lighting according to the awareness of occupancy or the desirable illumination. Manual control systems could be better for some situations, thus more desirable options should be selected by considering the patterns of electricity use by single-households. This research is a preliminary study for proposing appropriate solutions to the effective management of energy use for single-households. By maximizing smart spaces in single-households, sustainable and healthy housing would be realized in single-households in terms of energy saving. Further, this research would be a basis for enhancing the quality of housing environments for single-households.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP). (No. 2008-0061908)

References

[5] T.D.Pettersen, Variation of Energy...


1 Abstract

The solution often suggested to mitigate poor time, cost and quality in construction is to produce buildings in an automated and industrialised environment. To date, use of Off-Site Manufacturing (OSM) in Australia has enjoyed extremely limited success. Distilling the variables in the success and failure of OSM in other places and applying relevant variables to the Australian context, enables a better response for OSM. The purpose of the paper is therefore to examine whether or not OSM techniques are viable in the Australian housing market. The paper uses a detailed and critical analysis of the literature to examine OSM of housing in various countries, seeking to establish the major reasons for successful and unsuccessful models. The findings are then contrasted with the Australian context seeking criteria to inform successful introduction of OSM into Australia. Among other things, findings indicate the catalyst for the introduction of OSM of housing is almost universally a result of major events such as wars and natural disasters. Innovation has also played a role in encouraging change to construction methods. Whilst the countries addressed have diverse economies and climates compared to each other and Australia, certain common criteria have been found from those examples to assist in modelling an OSM solution in Australia. The implications of this work revolve around the provision of a more efficient, less wasteful and more responsive housing production environment which will potentially improve affordability in the market place.

Keywords

Off-site Manufacturing, housing, drivers and barriers, sustainability, mass customisation

2 Introduction

The question as to whether buildings should be produced in factories by off-site manufacture, rather than traditional on-site craft methods has been the subject of numerous papers and industry discussion over a considerable period of time [1; 2]. Today the question of OSM in Australia, particularly for housing, is being asked more often in an attempt to address issues in the construction industry of saving time, improving quality and better defining cost, as well as greater productivity. It must also noted that extraordinary events outside the industry have created a need to find ways to improve supply of buildings beyond the capability and capacity of the construction industry of the time [3-5]. Further, there are examples of building and construction designers seeking innovation to improve contemporary building construction methods and by-products of the industry. Today, for example, the aspects of sustainability and reduction of waste have evolved, important issues which Barrett and Weidman claim cannot be solved by traditional construction methods [6]. Luther suggests there is evidence that the use of factory production of buildings can solve some of the issues of time, quality and cost, and agrees OSM will enable better levels of sustainability including reduction of waste [7]. In their research for the Royal Institution of Chartered Surveyors (RICS), they go further and suggest traditional construction industry methodology will struggle to solve the challenges of finding meaningful building innovations [8]. Bengtsson nomimates a strong barrier to construction innovation is that of institutional path dependency [9]. In contrast, Thuesen and Hvam argue for modifying currently accepted construction methods citing a German example of developing product platforms, an important theoretical but practical position for OSM described later in this paper [10].

For this paper, a clear distinction needs to be drawn between a common term “Modern Methods of Construction” (MMC), used in the construction industry as a descriptive phrase for innovation in construction, and the term OSM [11]. The Building Research Establishment (BRE) defines MMC as a range of processes and technologies which include prefabrication, off-site assembly and various forms of supply chain specifications [12]. For this paper the term MMC is regarded as a broad generic term for innovation in construction techniques, both site and factory based (including OSM) and also encompasses whole-house prefabrication in a factory. Specifically for the genre of OSM, four distinct typologies have been identified by academics and industry, and these categories will inform the relevance of OSM for the Australian market. Bell and Southcombe nominate these categories as component (stick and assembly), panel (non volumetric), module (volumetric), and hybrid (module plus panel) [13].

The use of the term “off-site manufacture” (OSM) is described by Kenley et al and other respected writers on the topic, as embracing various categories for factory manufacture of buildings including off-site fabrication, off-site assembly, off-site construction and...
off-site production [14]. The literature of recent times consistently uses the term OSM to define the use of factories distant from the construction site to produce buildings and building components, and that acronym will be used in this paper [12; 13; 15].

This paper focuses on the house building construction sector in Australia. The housing sector provides an excellent example for building production to adopt alternative methods (whatever they may be) as a means to satisfy the current demand for housing, which exceeds the supply that conventional traditional methods can provide [16]. It is also relevant for this paper to review OSM of housing in countries where the methodology for manufacture has been developed, the review looks at recent examples to understand current paradigms and technology, and their relevance to the Australian model.

Historically, change from craft to manufacturing in many industries has occurred as a result of innovation. Change also occurs due to major events such as natural disasters or war. This paper will briefly discuss three industries which adopted a regime of factory production and whether that change informs the construction industry in Australia today, particularly addressing the under supply of housing stock described later in this paper.

2.1 Why consider OSM over Craft Construction?

Over time, traditional construction methods have failed to satisfactorily address the important criteria for buildings of time, cost and quality, and often for housing a methodology for supply to satisfy demand [8; 16-18].

It is often stated that the last industry to convert to significant factory production is the traditional craft construction industry [19-21]. The Egan Report titled “Rethinking Construction”, reviewed the construction industry in the UK positing that though the industry was considered by the community as excellent in execution of complex projects, the industry was under performing, unprofitable and could fall into stagnancy [17]. The Barker Review carried out for the UK government reiterated the same concerns focusing on the decline of housing construction, blaming lack of capacity by the construction industry but also noted a lack of sites for housing caused by factors often, but not always, outside control of the construction industry [22]. The same debilitative issues for construction in the UK prevail in Australia [23; 24]. Failure of the Australian construction industry to address these important issues is reported by The Built Environment Industry Innovation Council’s final report (2012), advising the Australian Government on ways to improve the performance of the construction industry and suggesting innovations for the future [25]. Their aims sought to not only improve the performance of the construction industry, but to equip the industry for the future. Importantly this advice also addressed the challenges of climate change and sustainability, and the need to address skills development and the ability to access and accept new technologies where those technologies benefit the industry.

3 The evolution of craft to manufacturing

History clearly demonstrates the changes to major craft industries over time. For this paper it is useful to briefly consider relevant examples of transition from traditional means of production to that of manufacturing, and the reasons for those changes which could be described as catalysts for change, and the benefits that flowed for that industry. The examples used in this paper which clearly demonstrate a transition to manufacturing are motor cars, textiles and ships. Gann, when comparing the manufactured housing industry to the automotive industry in Japan, notes that Henry Ford in America exploited the three main advantages of manufacturing over craft to produce motorcars. Those advantages were economies of scale where costs decrease as volumes increase, better use of capital, and better management control. Ford designed production lines which enabled use of semi-skilled and even unskilled workers to operate high-cost specialised machinery to produce motor cars [26]. Traditional vehicle manufacture at the time Ford developed his systems was unable to compete in supply or cost. The vehicles prior to the early 20C Ford paradigm, were craft constructed horse drawn buggies fitted with an engine.

Also demonstrating the new relationship between workers and machines was the textile industry, often a craft of great pride, often regional in style and requiring varying degrees of skill. Saxonghouse and Wright, in “Technological Evolution in Cotton Spinning, 1878-1933”, describe the invention by Compton of the spinning frame which was then commercialised by Arkwright, an English entrepreneur. The spinning frame enabled semi-skilled workers to operate machinery to convert raw cotton into mass-produced yarn [27].

The third example of transition from craft to manufacturing is that of shipbuilding, a craft industry right up until the Second World War changed by the advent of the “liberty ships”. These ships were produced by US shipbuilders for the UK Government seeking to address the German U Boat attacks on supply ships, seriously threatening essential supplies to the UK. These ships were constructed for a limited life, however the innovation which enabled fast assembly and supply and described by Heskett and Giorgetta, was the use of welding in lieu of rivets, and importantly, the production of modules of hulls for supply to the shipyards where they were assembled [28]. The time to produce the ships decreased from 230 days to 42 days, proving the efficacy of the systems. The ship building revolution continues today for the Korean shipbuilding industry, described by Bock et al as highly advanced and individual customised ships. Bock et al also link this vertical and horizontal evolution of ship building as providing a desirable model relevant and useful to the construction industry [29].

For this paper, the important outcomes are the further development of those industries into mass
production producing high quality, certainty of costs and variety of product. While there are many other examples of the change from craft to manufacture, the fact remains that building construction is still the last craft industry to follow a long line of predecessors in evolving into a manufacturing model [7]. For the house construction industry to develop a mass production methodology, the logical outcomes must be to replicate the outcomes for the craft to manufacturing examples described, greater supply, greater choice and high quality.

4 OSM of housing, drivers and barriers

It is of relevance to this paper to examine and contrast both positive and negative attributes of “prefabrication” of housing in recent history and in various countries. Examples of prefabrication prior to the 20\textsuperscript{th} century, whilst of interest, are not considered as relevant to a discussion of the use OSM for the provision of housing in the near future.

4.1 Scandinavian OSM of houses

In Scandinavia and Finland the use of timber construction for houses has been and continues to be their preferred medium. Use of traditional three dimensional timber framed forms and styles familiar to the Scandinavians have been readily transposed into timber-based factory built systems. For Scandinavia, adopting factory based manufacture enables continuous production all year round, thereby defying the harsh winter elements which constrains supply. Waern in “Home Delivery” describes Scandinavia as now being inits third age of prefabrication, the latest iteration that of small lumber companies which have morphed into prefabrication enterprises and are now corporate identities operating internationally [1; 30]. One group, Skanska/Ikea, is supplying houses to the Scandinavian, Russian and UK markets. The common methodology used for OSM in Sweden is essentially that of transferring traditional site tasks (except for slabs/foundations) into a factory environment, enabling lean construction and acceptable quality, realising time savings through avoiding delays due to poor weather. The growing use of timber volume elements (TVE's) in lieu of the more common assembly of large components could be a potential challenge for OSM of housing in Sweden. TVE's are viewed by the public as “one size fits all”, however, the industry has recognized the need for the introduction of mass customisation (MC) and flexibility in design [31]. Lessing et al suggests this need for MC for all typologies of OSM has been recognised in Sweden and is being addressed by housing providers through the use of information and communications technology, and the use of “agile construction”. In Sweden 80\% of detached family housing is produced using OSM [32; 33].

For Sweden, drivers are an OSM industry initiated by suppliers of timber materials in a market not dominated by a preference for brick [1] and the advantage of year round production. Success for OSM may also be attributed to the familiarity of the housing forms avoiding issues of path dependency. The size of the market creates economies of scale and therefore affordability compared with traditional site-built housing. Waern (In Home Delivery 2008) states an extreme shortage of housing, dating from 1917, created a major driver for prefabrication in Sweden [30].

4.2 Germany and OSM of houses

Linner and Bock estimate 15\% of new house production in Germany is produced by OSM [34]. The manufacture of OSM for housing production in Germany was personally experienced during a visit in 2012. Two different methodologies were observed, Huf Haus using laminated timber post and beam and pre-finished wall and glazing panels, and Massa Haus, using a system of fabricating whole walls and roofs essentially using traditional techniques and materials for later assembly on-site. Both examples when erected on site are capable of achieving weather tightness quickly, but require weeks to complete finishes and services. Huf Haus has adopted a personal face to face with an architect/client design format to produce individual solutions within the post and beam genre, whilst Massa Haus produces different styles to appeal to a broad audience from traditional to “modern”. The Huf Haus factory in Hartenfels, I was informed produces 200 houses a year while Massa Haus produces 2000 each year.

The German OSM industry has overcome earlier images of poor quality, initiated by poor perceptions of post war prefabricated house examples. Venables et al states the German OSM industry did this by developing rigorous quality standards with certification schemes to provide client assurance [35]. This negative perception of OSM due to post war prefabricated housing, is widely observed throughout the western world by both the market and the building industry [36; 37]. Venables et al in their study, noted the highly professional promotion of OSM in Germany using among other initiatives, house demonstration villages, which the author experienced and photographed on the visit to Germany in mid 2012 [35]. It is interesting to note the growing use of the concept of product platforms to rationalise production and at the same time satisfy the market need for mass customisation [38; 39]. For Germany, success has been in part due to satisfying path dependent issues through offering familiar styles and also providing security of quality and durability. Germany has provided security through the establishment of an association (The Association of German Prefabricated Building Manufacturers) (BDF) which requires the quality systems mentioned to be satisfied by its 45 members.
4.3 The United States of America and the Development of OSM for Housing

In the US the use of OSM has a history of varying success and failure. Early OSM in the first half of the 20th century comprised extremely successful kit home examples delivered to the site for assembly, produced by companies offering mail order from catalogues. Sears

Roebuck is probably the most cited company. Mostly using balloon frame construction, (typically timber wall and roof frames pre-made off-site) they were readily erected by builders or owners. Sears sold some 100,000 homes offering 447 different models over a period of 32 years [2]. When the 1930's depression engulfed the world, the housing market stalled and clients defaulted on loans made by Sears. As a result the house businesses of Sears and others failed financially. The US government, as have many governments, encouraged mass production of houses following the second world war, prefabricated houses to satisfy demand which could not be satisfied by the construction industry at the time. This was in part due to a lack of skilled workers due to absent trades serving in the war, but also due to the large number of returning service personnel promised a home by the government on their return [1]. The program was stopped in 1947, and Davies attributes the demise of the prefabricated program for post war housing to the US government [1]. The government demonstrated a preference for traditional methods of on-site production over off-site fabrication, once more demonstrating the perception of a prefabricated house as an inferior product to a site built house. The perception by government and the market, of OSM as an inferior product raises a very large barrier, regardless of the reality.

Two of the better known prefabricated companies in that post world war period encouraged by government programs were Lustron and the General Panel Corporation. Lustron developed from a company which manufactured road side restaurants and service stations using porcelain enamelled metal panels. They transferred this successful technology into a housing model, and reportedly enjoyed a level of success, however, despite significant government funding to undertake an expensive plant, Lustron was bankrupted holding 20,000 orders but having completed very few of those sales [1]. Well known architects Gropius and Wachsmann designed the Packaged House system for The General Panel Corporation. Bergdoll and Christensen describes its greatest attribute as that of comprising panels which could be used for both walls and floors, representing a considerable reduction in the number of parts reducing complexity, thereby enabling faster assembly and importantly, flexibility in design [2]. The company, although similar to Lustron in the receipt of government financial support, failed to be profitable and ceased manufacture. A significant feature of the Gropius and Wachsmann model was their development of a system with fewer parts than other systems, parts which were interchangeable which enables reduction in the amount of stock holding, and therefore improves the economies of scale. OSM is currently available in the US, albeit producing a small percentage of the housing stock, that percentage (approximately 3%) being difficult to fully quantify due to the classification of OSM as manufacturing (including mobile homes) rather than housing [24]. Of the manufactured product which can be classified as detached housing, most examples are volumetric and modular with budgets of between $200 to $250 per square metre for a completed assembly on site (Budgets were obtained by visiting relevant internet sites). Failures of OSM in the US can be attributed to manufacturing investment too high for the average market house pricing and government's mixed attitudes to prefabricated houses. Lessons from the US for manufactured houses are to resist large infrastructure investment and minimise the number of parts and design flexible systems for customisation.

4.4 The United Kingdom and its use of OSM for Housing

In the UK, the evolution of OSM of housing has followed similar precedents of other countries. Construction of housing virtually ceased during the world wars. The housing shortages thus created were exacerbated by returning service personnel from those wars, as well as government slum clearance programs and the replacement of houses demolished during the conflict. Gay and Vale discuss the UK government's action to address the shortages by their establishment of the “UK Temporary Housing Programme” [3; 4]. Under this program the government ordered production of thousands of prefabricated houses in order to satisfy demand for housing which the existing house building industry could not provide within the desired time frame. The brief for the prefabricated houses called for a maximum life span of 10-15 years after which they were to be dismantled. The government was concerned that prefabricated houses should not be construed as site built houses which were regarded as permanent and of better quality. A serious issue for the image of houses produced by prefabrication under the “Temporary Housing Programme”, was a lack of variety in their plan and three dimensional form, for all houses were briefed to have two bedrooms and maximum floor space of 800 sq ft [3]. This perception of monotony was sharpened when the prefabricated houses were placed in one area en mass. It should be noted however, that there was acceptance by many of the occupants of the “bungalows” (as they were described) resulting from inclusions of internal bathrooms, allowance for white goods and the added enjoyment of a private garden.

Both Gay and Vale suggest that the UK Government's failure to require some standardisation of components and methods by the various providers, resulted in failure to achieve economies of scale inherent in mass production. Although the plans were almost identical, the manufacturers developed their own systems seemingly to compete and offer a superior system to their competitors in order to gain a
greater market share. The prefabricated housing produced under the “temporary programme” therefore resulted in higher than necessary costs and were unable to challenge standard site constructed housing on a cost basis.

Davies in his book “The Prefabricated Home” refers to the government investing in systems of prefabricated housing during the war periods but reverting to prefer standard site based production once the war was over [1]. Prefabrication in the UK failed to evolve successfully due to public perceptions of poor quality housing, and providing temporary rather than permanent accommodation [1; 3; 4]. In assessing the percentage of OSM compared to on-site for the UK, Goodier and Gibb state they could only make assumptions in order to assess the level of OSM for housing due to lack of reliable statistics of OSM [40]. Currently, despite a large amount of literature endorsing OSM housing, Pan and Sidwell estimate the use of OSM is relatively small in the UK, quoting the entire industry (including all OSM projects) fails to gain more than 6% of the total market [41]. They place some of the reason for failure on the perception that OSM is more expensive than on-site construction, however note the fallacy of this position through their research which proved otherwise. Significantly, in the UK, house builders rely more upon land sales rather than the construction of houses to make profits, challenging the prospect of OSM further [42]. Pan and Sidwell suggest that there must be continuous exploration and commitment to refining off-site technology and importantly collaboration with the supply chain. The literature alludes to the necessity to change the construction industry to a manufacturing industry if OSM is to succeed. A major reason for this approach lies in the scarcity of skilled tradespeople, Arif et al, go further claiming the benefits of OSM will only be realised when the processes of design, manufacture and construction are completely re-engineered. Their research also recommends design flexibility to meet stakeholder needs [43]. Edge et al, in their research into market perceptions, state that the market in the UK does not discriminate in regard to methods of production of houses, but is concerned when the appearance is not traditional [36]. Davies is more critical of the UK market, citing the UK predilection for brick construction and finishes, a material and technology not easily used in OSM [1]. Craig et al address the issue of reluctance to accept OSM housing, suggesting an approach which optimises the economic, ecological and social issues, will reduce that reluctance [44]. Perhaps the prospect for use of OSM for houses in the UK can be best summed up by a visit made by the writer in 2011 to the BoKlok site at Gateshead. What was purported to be an Ikea flat-pack housing system was in fact traditional construction, “the flat pack would have been more expensive” was the response from the site foreman.

4.5 OSM housing in Japan

Japan has substantially developed the production of OSM for housing. Bergdoll et al argue tradition and traditional timber construction has given Japan an underlying philosophy encouraging the design and construction of housing using methods of prefabrication [2]. The typical Japanese concept for traditional housing is that of post and beam structure with infill panels. Use of a module was and remains a feature for housing design based on ancient traditions. Development of the post and beam system for OSM housing was described by Oshima in “Home Delivery” as modernisation rather than transformation [45]. Following the Second World War, Japan according to Oshima, had a shortfall of 4.5 million housing units and suffered from the same problems as other countries, that of a lack of skilled trades, many of whom served in the war and of whom many were lost [45]. In addition, significant quantities of their housing stock was destroyed during the war. In contrast however to other major countries, a number of Japanese industrial companies developed housing designs suitable for OSM. From 1959 companies such as Sekisui, Daiwa, National and Misawa produced simple box like houses to meet demand, these companies soon developed more complicated models in order to better compete in the market [46]. Importantly, although economies of scale were found in the size of the Japanese market, those economies were not sufficient to produce houses of less cost than site built houses by local builders. Johnson discusses the approach to offset the cost disadvantage that the factory house builders experienced, by developing solutions which were marketed as having superior performance to the on-site produced houses. In addition the factories offered long term warranties and continuing maintenance. Possibly the most important feature of the Japanese OSM housing industry according to Gann and Barlow et al was the use of a system pioneered by Toyota for manufacture of vehicles [26; 47]. That system called “Kanban” changes the supply chain conventions yielding important concepts of lean production and “just in time” which in turn reduces waste in materials and labour, and further results in less inventory due to greater flexibility and greater variety in product. Barlow et al reinforce Gann's view that the use of mass customisation by the Japanese manufacturers has enabled successful operations despite the cost challenges [47]. The Japanese systems enable house forms which according to Barlow and Oszaki satisfy issues of path dependency in shaping Japan's use of mass customisation in the house building industry [48]. Linner and Bock estimate OSM house production in Japan as 13% to 15% of new detached housing per annum [34].

4.6 OSM in Australia

For Australia, albeit a country younger and with a population far less proportionally to the countries previously described, the experience for OSM is very similar to those other countries, including various attempts to supplement housing needs using OSM. Greig for example, describes the actions following the second world war by the NSW Housing Commission (a social housing provider) [5]. The Commission obtained some 200 prefabricated house design proposals and constructed 25 test houses from those
proposals. None were put into production, the Commission claimed the prefabricated houses cost more than “orthodox” dwellings [49]. Another example of the early use of OSM housing was the “The Snowy Mountains Authority” which imported some 4,000 houses from Europe to house workers who migrated to Australia to build the Authorities scheme [5; 50]. However, the imported houses were generally regarded as unsuitable to the Australian climate having been designed to suit the standards of northern Europe. In Victoria, the State Government, after World War II, converted the Beaufort aircraft factory into a prefabricated house manufacturer. Beaufort Homes represented a ‘modern method of housing construction’, combining the skills of architect Baldwinson with the technical expertise of the Department of Aircraft Production. The Federal Government at the time ordered five thousand of those houses for production, however only twenty-three were eventually built (State Library of NSW). The reasons for rejecting the Beaufort house were that the homes were smaller in area than comparable site built houses then available, and that they were more expensive [5]. Failure to pursue OSM in the years following the world war was attributed both to a relatively small market and the lack of sufficient industrial capacity and flexibility in Australia. The industry could not overcome the economies of scale needed to produce models which could be cost competitive with site-built housing methods (State Library of NSW). Further, the image of prefabricated houses as that of a temporary dwelling in Australia was similar to that of UK, Germany and the US.

5 Current issues for Australia

The question of the use of OSM to produce housing is received with various levels of doubt and confidence world wide, however the focus of this paper is Australia. Unplanned events described earlier have eventuated in Australia due to the current and increasing deficit of the housing supply in Australia [51]. The importance of adoption of OSM for houses in some form, particularly for detached housing, which comprises 77% of residential accommodation in Australia [52].

For housing worldwide, there is an increasing recognition that the housing sector, both in construction and over their life cycle, is guilty of producing a larger carbon footprint than any other sector [6]. The size of carbon footprint together with the aspects of extreme waste in on-site construction together with low sustainability, needs to be addressed and corrective action taken. Although MMC innovation has been gradually incorporated into on-site construction methods, the methods continue to use relatively standard construction techniques which have been shown to be ineffective in addressing waste and efficient cost effective outcomes. The innovative methods as described by Dalton et al for the Australian house building context hardly qualify for great steps [21]. Firstly, they describe the use of specialist on-site equipment such as nail guns and power saws and the like, secondly, faster communications between the various contractors and sub-contractors through mobile phones, faxes and emails, and thirdly the use of information and communications technologies to enable off-site component manufacture such as roof trusses and complete wall frames. The optimal use of OSM is generally more problematic in terms of uptake. To date, the use and uptake of OSM has been variable in the design and construction of buildings. Limited but increasing success is being realised in commercial applications in Australia. Student, apartment and hospital accommodation constructed by “Quicksmart” is one example of using factory fabricated modules completely finished and craned into position on site. Also finding success in this genre is “Unitised Building”, which has recently used similar techniques of factory manufactured modules to construct a number of medium rise apartment buildings in Melbourne. These examples have, according to the two companies achieved successful outcomes for the criteria of timesavings, better quality and cost effectiveness. While there are examples of volumetric modules and whole houses constructed for the detached house market, the volumes thus far produced are incapable of satisfying the market need, and there is no evidence that the industry can substantially improve supply without further innovation [16]. According to Blismas and Wakefield the quantity of OSM detached housing supply in Australia is difficult to estimate [24]. They state there is little definitive evidence available due mainly to a lack of differentiation in the typologies of traditional and OSM approaches, however their estimate is that around 3% of housing produced in Australia could be defined as OSM. Blismas and Wakefield partially attribute unreliable statistics in Australia to the lack of a peak body to assist in quantifying OSM. This situation may have been recently redressed by the formation of “PrefabAUS”, an industry group whose “mission is to represent, showcase and advance the Industry through collaboration, innovation and education” [53]

6 Questions for the use of OSM, is there a case?

Although OSM in Australia has been identified by Hampson and Brandon as becoming a significant player in the construction industry by 2020, there is little evidence this has or will eventuate [16]. It is suggested by Kenley et al that this reluctance for the construction industry to adopt OSM is due to a lack of knowledge of the benefits or understanding of how OSM or MMC could fit into current construction practices [14] This is particularly relevant for the house building industry, not so much the volume builders, but certainly for the majority of house building companies which are classified as small to medium enterprises (SME) working on tight margins. Those SME’s are unwilling to accept the perceived additional risk which OSM and MMC creates [54].

In the countries examined for this paper, it is important to note that a severe shortage of housing was the motivation for pursuing OSM for housing. In each case the governments’ action, whether it was the US
government’s “Operation Breakthrough” (1968-1978), the UK government’s “Emergency Housing Programme” (1945-1951) or the Australian government's seeking prefabricated housing solutions, their actions failed. Failure to evolve a successful OSM program was attributed to costs of OSM housing being higher than traditional on-site production. The other major constraint for an ongoing use of OSM were the decisions made by governments to revert to site-built housing using the conventional “craft” skills, based on perceptions of doubtful quality and longevity of OSM. Today, once again, there is a critical shortage of housing in Australia partly due to skills shortages for construction, but mostly due to a reluctance by the construction industry itself to change, whether that genre is MMC or OSM. The relatively successful examples of adoption of MMC or OSM are Sweden, Germany and Japan. The entities producing OSM houses in those countries are manufacturing organisations rather than construction organisations. The indications are that manufacturing in other places will be most likely to realise suitable OSM housing outcomes. This appears to be the case for Australia, it is personally observed that most current OSM producers do not have construction arms, apart from site assembly teams. Lessons for Australia in selling OSM can be learned from Germany and Japan, both having implemented flexibility into their systems through principles of “agile” methods in order to obtain mass customisation rather than mass production. They have done this by seeking customer input at all stages, and they have also adopted quality management and certification. In Japan's case, after sales service and extended warranties also help in off-setting the often higher cost [55] of their OSM housing compared to on-site. These actions thereby address to a significant extent the issues of adverse perceptions of OSM housing the customers, authorities and importantly financial institutions. The Australian market shares with the UK a preference for brick housing solutions [5], a difficult barrier for OSM to overcome. From the literature solutions to negate this path dependency lay with offering high quality together with certainty, and addressing issues of sustainability and waste. The Australian and UK literature nominate design as an important factor, namely that OSM is regarded by industry as requiring early design freeze, and therefore the product lacks flexibility and also incurs expensive penalties for late changes. The solution for the industry is not an easy one, requiring retraining and adoption of new skills to design for manufacture and assembly [56] rather than design for traditional methodology. 

Clearly, OSM producers establishing expensive facilities such as occurred for the UK post war “Emergency Housing Programme” and the US examples of Lustron and General Panel Company, risk financial stress. The risk of financial failure is partly due to the cost of the debt, but also the necessity to amortise the debt over potentially limited sales which inflates the cost of the product in many cases to be higher than the site-built product. This is particularly applicable to the Australian market size leading to doubtful economies of scale. Australia can learn from German, Swedish and Japanese OSM producers addressing the supply chain dynamics enabling reduction in stock holdings and timely supply, thus reducing costs. Innovative use of the supply chain permits use of “pull” rather than “push” factors ensuring inventories remain minimal whilst still satisfying the market forces. From Germany and Sweden the concept of product platforms has evolved for OSM housing, a concept in principle similar to the General Panel Companies design for interchangeability of wall and floor panels. This is particularly relevant for Australia, for the concept of components having many uses and configurations enables reduction of the number of components, and therefore cost, and at the same time enables flexibility of design to suit individual briefs. Use of this product platform system reduces costs due to reduction of components required by an OSM system.

7 Conclusion

Whilst there is ample evidence in Australia that traditional construction methods cannot supply sufficient housing stock to satisfy current demand, or supply housing to address a significant deficit of some 200,000 houses in Australia [51], there is little evidence of action to innovate and solve the issue of supply. It is clear from other examples of craft industry converting to manufacturing that the products so produced are more affordable, predictable in quantity and quality, and offer offer variety and choice, no such change in the production of housing can be discerned. One factor which does assist in the affordability of manufactured products is reduction of waste. The problem of waste and sustainable practices is clearly one for the production of houses which use traditional on-site methods [6]. There is a great deal of evidence OSM can assist in solving these dilemmas.

While the role of governments is extremely important to the housing industry, particularly the social sector, past experience urges caution and suggests the industry for OSM of housing needs to develop policies to negotiate successful outcomes.

Previously mentioned in this paper, is that the use of OSM in commercial residential has been successfully adopted satisfying the important criteria of time, cost and quality, and it seems the visual form of the finished product being similar to the norm, gains acceptance by the market. For detached housing however, there is limited success in Australia, that success comprising mainly volumetric typologies using traditional methods of assembly within factories, particularly the whole of house fabrication which most often emulates the form and style of traditional housing. These examples demonstrate little innovation, but certainly demonstrate the value of a factory environments for producing houses, typically continuous working conditions unaffected by weather, safer work environments and less waste. However, these volumetric examples are not capable of meeting demand. Successful implementation of OSM in Japan, Scandinavia and Germany
demonstrates the part played by offering variety including traditional styles.

Examination of successful examples of use of OSM to produce housing has informed this paper. Cost of housing is an important factor and the means of competing with site-built production lays with concepts of lean principles, supply chain innovation and design for manufacture and assembly. Poor perceptions and path dependence create great obstacles to adoption of OSM, however the examples from Germany, Japan and Scandinavia show that use of quality systems, warranties, after sale service and customisation mitigate these negatives. Customisation is possible by the use of multi use components and the design of product platforms. These methods also enable a OSM system to more easily satisfy economies of scale. Importantly there is the question of what form of product the OSM production groups should take. In the US but particularly the UK, the providers of OSM are seen as members of the construction industry. The evidence from successful OSM industries is that the proponents are manufacturing based.

For the OSM industry there are new technologies and innovative materials such as plug and play fittings for services, and new methods of production available such as CAD/CAM systems for manufacture of components. There are smart foundation methods which do not require excavation or in-situ concrete. These innovative systems will offer the edge over on-site construction sufficient for OSM to develop into a relevant and profitable industry in Australia.

8 References


[41] Pan, W. and Sidwell, R. (2011). Demystifying the cost barriers to offsite construction in the UK,


Abstract
In this study, we developed a computational fluid dynamics (CFD)-based method to optimize the arrangement of buildings for urban ventilation potential at the conceptual design stage and demonstrated its application in a case study. For CFD-based optimization, three calculation components must be assembled: an optimizer, a geometry/mesh generator, and CFD solver. The optimizer solves the optimization problem, which comprises design variables, objective functions, and constraints functions. The geometry/mesh generator creates a building geometry (i.e., vertices, faces, and cells) that satisfies the design variables generated in the optimizer, converts them into a mesh file compatible with the CFD solver, and assigns the boundary condition. The CFD solver is used to calculate and return the value of the objective function to the optimizer. In this study, the local kinetic energy (KE) of the target area was employed as the objective function. KE is a parameter that is used to precisely evaluate the convective effects of the wind and is calculated from the turbulent components and the averaged velocity components. The results showed that genetic algorithm enables the developed method to provide options to designers that less negatively affect the urban ventilation potential of the surrounding area.

Keywords
Building shape, Urban ventilation, Urban design, Optimization, CFD simulation, Genetic algorithm

1 Introduction
At the conceptual design stage, there are many opportunities to improve the building performance in terms of the urban ventilation potential and the building energy consumption. For example, the building orientation and massing are typically determined early in the design process and have a significant impact on the environmental performance [1]. During this stage, in order to select the more valuable design, many alternatives are generated and evaluated by any indices for performance assessment. In this sense, many studies have examined building shape optimization with regard to the relationship between the building morphology and energy consumption [2-4], but little attention has yet been paid to the relationship between the building morphology and urban ventilation potential. In this study, therefore, we developed a computational fluid dynamics (CFD)-based method and demonstrated its ability to optimize the building arrangement for urban ventilation potential at the conceptual design stage.

Most current research into the shape optimization for building performance has revolved around a single building. Previous studies have considered a wide range of building shapes, from simple geometries such as a box and polygon to atypical geometries with irregularities. Horikoshi et al [2] and Tuhus-Dubrow and Krarti [5] examined polygons such as rectangles, trapezoids, crossed. Yi and Malkawi developed their own building shape representation method and enhanced a variety of building shapes [3, 4]. These papers are invaluable works of references for an engineering approach to building shape optimization. However, these researchers did not consider the fact that a designer will not accept a plan derived purely from an engineering perspective, no matter how rational. Thus, this study was based upon the premise that the building shape is the domain of the designer. We focused on building arrangement, which is practical engineering support for site-specific building planning. In this paper, we describe how our methodology allows the generation of an optimized site-specific building arrangement through CFD simulation and a genetic algorithms (GA) and present a case study.

2 Outline of Case Study
The target area for the case study covered a densely...
built-up area of approximately 500 m × 500 m in Tokyo, Japan. Figure 1 shows the current status of the target area. In the first step, a renewal plan was developed with regard to the central district of this area (target site in Figure 1). In this case study, the designer had finished the conceptual design (massing only) and was considering how to reduce the negative influence on the urban ventilation potential through the building arrangement. Table 1 lists the details of the entire target area. Table 2 lists the information on the three masses proposed by the designer. In the second step, CFD-based optimization was carried out to determine which building arrangement maximized the urban ventilation potential.

3 The Computational Method

To develop the CFD-based optimization, three calculation components must be assembled: an optimizer, a geometry/mesh generator and a CFD solver. The optimizer solves the optimization problem, which comprises design variables, objective functions, and constraint functions. The geometry/mesh generator creates a building geometry (i.e., vertices, faces, and cells) that satisfies the design variables transferred from the optimizer, converts them into a mesh file compatible with the CFD solver and prescribes the boundary condition. The CFD solver calculates and returns values of the objective function (see Figure 2). For the CFD solver, we utilized the Star-CCM+ 8.02 simulation engine. MATLAB R2012a was used to implement the representation developed for the GA optimization. The objective functions and constraint functions discussed in the following sections were programmed using m-files. To handle the geometry, we utilized a parametric design program of our own making.

Table 1. Target area properties

<table>
<thead>
<tr>
<th>Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total area</td>
<td>250,000 m²</td>
</tr>
<tr>
<td>Number of buildings</td>
<td>607</td>
</tr>
<tr>
<td>Mean building height</td>
<td>29.33 m</td>
</tr>
<tr>
<td>Mean building plan area fraction</td>
<td>35.06 %</td>
</tr>
<tr>
<td>Mean floor area ratio</td>
<td>342.66 %</td>
</tr>
<tr>
<td>Target site area</td>
<td>16,712 m²</td>
</tr>
</tbody>
</table>

Table 2. Building masses properties

<table>
<thead>
<tr>
<th>Building</th>
<th>Height</th>
<th>Floors</th>
<th>Floor height</th>
<th>Plan area</th>
<th>Total floor area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building1</td>
<td>80 m</td>
<td>20</td>
<td>4 m</td>
<td>1,600 m²</td>
<td>32,000 m²</td>
</tr>
<tr>
<td>Building2</td>
<td>24 m</td>
<td>6</td>
<td></td>
<td>3,500 m²</td>
<td>21,000 m²</td>
</tr>
<tr>
<td>Building3</td>
<td>12 m</td>
<td>3</td>
<td></td>
<td>2,700 m²</td>
<td>8,100 m²</td>
</tr>
</tbody>
</table>
a link between the design variables and a mesh file compatible with the CFD solver. In this section, we discuss the design variables that represent the building arrangement.

In building shape optimization, parameters which represent building geometries can be classified into two sets of variables: location-related and architectural form-related variables. However, because the architectural form is determined by the designer in this case study, all architectural form-related variables were fixed. Because the morphological diversity is constrained by this precondition, we can represent building morphology with shorter chromosomes in the GA. This makes the present method more practical.

Figure 3 shows design variables for building 1 on the site. The coordinates of the center of the building geometry \((C_{x1}, C_{y1})\) were used as the location-related variables, and the angle between the axes of the site and building \((\theta_1)\) was used as the architectural form-related variables. We can describe the building arrangement by using a total of nine variables (three buildings \(\times\) three variables).

![Figure 3. Location-related variables](image)

3.2 Objective Function

The mechanism of movement by the fresh and cool outside air into the target area depends on the turbulent diffusion. Therefore, the urban ventilation potential of the target area may be considered from the perspective of turbulent transport. Thus, the objective function should be an index that represents this turbulent transport performance quantitatively. In this study, we used the local kinetic energy (KE) to describe this performance. This parameter is used to precisely evaluate the convective effects of the wind and is calculated from the turbulent components and the averaged velocity components, as defined in Equation (1) [6].

\[
KE = \frac{1}{V} \iiint_{\text{space}} \left( \frac{1}{2} (\overline{U}^2 + \overline{V}^2 + \overline{W}^2) + k \right) ds
\]

where KE is the spatially averaged kinetic energy \([m^2/s^2]\), \(\overline{U}, \overline{V}\), and \(\overline{W}\) are the averaged velocity components \([m/s]\), and \(k\) is the averaged turbulent kinetic energy of the entire target space \([m^2/s^2]\).

The outdoor wind environment has previously been assessed according only to the flow in the dominant wind direction. However, this approach remains open to debate because the calculation result may change considering the fluctuations in the wind direction. In this study, therefore, we attempted to obtain a more reasonable solution by considering the annual wind distribution in an objective function calculation using Equation (2). We used the hourly measured statistical data recorded over 10 years (from 1991 to 2000) at the Tokyo Meteorological Observatory, which contained the percentage frequencies of the wind direction and the mean velocity by direction. In this study, we categorized the 16 directions into the four cardinal directions to decrease the computing load. Table 4 lists the detailed information of the four wind-direction groups.

\[
KE_{\text{annual}} = \frac{\sum KE_i \times t_i}{8760}
\]

where \(i\) is the azimuth group number (1, 2, 3, or 4), \(KE_i\) is the local kinetic energy for azimuth group \(i\) \([m^2/s^2]\), and \(t_i\) is the time for azimuth group \(i\) \([h]\).

<table>
<thead>
<tr>
<th>Group</th>
<th>Azimuth</th>
<th>Mean velocity ([m/s])</th>
<th>Time ([h])</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N, NNE, NNW</td>
<td>2.00</td>
<td>3604</td>
</tr>
<tr>
<td>2</td>
<td>E, NE, SE, ENE, ESE</td>
<td>2.24</td>
<td>1991</td>
</tr>
<tr>
<td>3</td>
<td>W, NW, SW, WNW, WSW</td>
<td>2.60</td>
<td>1728</td>
</tr>
<tr>
<td>4</td>
<td>S, SSE, SSW</td>
<td>2.25</td>
<td>1437</td>
</tr>
</tbody>
</table>

Finally, we used the relative \(KE_{\text{annual}}\) (RKE\(_{\text{annual}}\)) as objective function. In this study, we attempted to represent the urban ventilation potential of individual cases by using RKE\(_{\text{annual}}\), which is maximized by the GA process. RKE\(_{\text{annual}}\) is defined in Equation (3).
The commercial software Star-CCM+ 8.02 was used. The domain height was 500 m, and the distance downstream from the target area was set to be 500 m. Tetrahedral grid cells were generated, and a finer mesh was applied to the region close to the surface and expanded farther away from the solid surface. The minimum grid size near the building was 1 m. Although the total number of cells varied for each case, it was about 11,000,000. All solid surfaces, including ground and building surfaces, were modeled using a no-slip boundary with zero roughness. We used the Reynolds-averaged Navier-Stokes (RANS) model to compare the urban ventilation potential of many design alternatives. Large eddy simulation (LES) and direct numerical simulation (DNS) were beyond the scope of this study because of their very high computational costs, although they can provide more accurate results. As the major purpose of this study was to estimate the relative overall ventilation potential for the urban area under the influence of the new building rather than obtaining the actual information for the target area, we selected the standard k-ε model. Gao and Niu [7] also indicated that the standard k-ε model is sufficient if the emphasis is on the airflow field. Figure 5 and Table 5 present details on the conditions of the objective function calculation.

### 3.3 Constraint Functions

A constraint function specifies a condition that individual cases generated in the GA optimization process must satisfy. In this case study, two constraint functions were employed: all vertices of the building geometry must be inside the given site, and the highest building can be no higher than 100 m. These constraint functions were formulized in m-files as equations that included the design variables, and were inserted into the optimizer modules.

### 4 Results and Discussion

Figure 6 shows the evolutionary changes in RKE_{annual} as the calculation progressed. In the early calculation stages, the degree of RKE_{annual} value change was significant. As the calculation progressed, the degree of change in RKE_{annual} become insignificant. Although the GA was not completed because of time limitations, the value seemed to be converging at about 1.035. The last generation point where the results converged was composed of 12 arrangements, which was the value set as the population size of GA parameters. These arrangements can be candidates for the designer’s choice. This means that the designer will
always choose superior designs in terms of the urban ventilation potential if any design in the last generation front is selected. Figure 7 shows one arrangement from the last generation in this case study.

Figure 6. Evolutionary changes in RKE\textsubscript{annual}

Figure 7. Building arrangement for last generation

Figure 8 shows the initial case (Ini) and last case (Las) for KE in each wind direction. The differences between both cases was up to about 10% with an east wind, even KE of the initial case was larger than the case with the west wind. This led to a small deference in KE\textsubscript{annual}, of 3.5%. This may seem small when discussing the superiority of different designs. However, this small deference is caused by the fact that KE\textsubscript{annual} is the volume averaged value of the entire target area (500m × 500m × 200m). Figure 9 clearly shows the differences between both cases, it represents the area in which KE\textsubscript{annual} became bigger when the initial case was replaced with last generation case. Although the region where KE\textsubscript{annual} of the last case was larger than of the initial case (marked in black) comprised about 52% of the total area, this was concentrated around the target site.

5 Conclusion

The main purpose of this study was to develop and demonstrate a CFD-based optimization method for building arrangement. This method has three main calculation components: a mesh/geometry generator, CFD solver, and optimizer modules. In this study, we explored the use of GA as the optimization algorithm for performance-based building arrangement planning. In our case study, the results showed that the GA realized a building arrangement that suited the given environment better than the alternative design. However, a number of problems that remain to be explored. The primary issue is the calculation load. The present optimization calculation was performed in a parallel computing environment. The calculation time to complete 20 generations (population size: 12) was approximately 50 days. In the future, this study will be extended to employ a more time-efficient algorithm to optimize the building shape.
References


Wood-Frame Wall Panel Sequencing Based on Discrete-Event Simulation and Particle Swarm Optimization

Mohammed Sadiq Altaf, Mohamed Al-Hussein and Haitao Yu

Hole School of Construction Engineering, University of Alberta, Canada
Landmark Group of Builders Ltd., Canada
Corresponding author email: msaltaf@ualberta.ca

Abstract-

In recent years off-site construction has become popular in North America due to the superior quality of the product, improved productivity, and reduced environmental impact. The panelized construction approach is one of the most readily utilized off-site construction methods. In a wood-frame panelized construction plant, wall panels are customized according to various design parameters such as length; height; number of studs, windows, and doors; panel type; and number of walls. These design parameters affect the processing time at each station in the plant, while the panel sequence affects the waiting time between stations. Due to this dynamic nature of the fabrication process, it is challenging to automatically generate an optimal panel sequence, as a result this task is performed manually in current practice. This paper focuses on integrating discrete-event simulation (DES) with an optimization algorithm in order to automate the panel sequencing process. Processing time at each station is calculated based on a task time formula which is a function of the design parameters of the panel, while delay is calculated based on a distribution derived from historical data. A particle swarm optimization (PSO) algorithm is integrated with the simulation model using a central database in order to generate an optimal panel sequence. The proposed method will eliminate the manual work required for panel sequencing, and is expected to reduce production time up to 10%. The proposed method is implemented in a wood-frame panelized construction plant as a case study.

Keywords-

Panelized construction; Discrete-event simulation; Particle swarm optimization; Panel sequencing

1 Introduction

Off-site construction is increasingly being embraced as the preferred construction approach due to the superior quality of the product, improved productivity, and reduced negative environmental effects. Among the various off-site construction methods, panelized construction is becoming widely utilized due to its design flexibility and lower on-site assembly cost. In this regard, a study by the Building Systems Council (National Association of Home Builders) has revealed that panelized construction reduces waste and construction time compared to traditional stick-built construction [1].

As the majority of the activities in a panelized approach are performed in a factory environment, it is important to achieve optimal productivity in the production line. Among the various assembly line classifications, the panel production line is a mixed-model asynchronous assembly line in which \( n \) number of jobs (panels) go through \( m \) number of stations in a series configuration [2]. In a wood-frame panel fabrication plant, wall panels are customized according to various design parameters such as length; height; number of studs, windows, doors; sheets of sheathing, walls; and panel type. These design parameters affect the processing time at each station, while the panel sequence affects the waiting time between stations. As different panels require different processing times at each station, it is important to produce the wall panels in the optimal sequence in order to minimize the maximum completion time of all panels, also known as the "makespan". This kind of assembly line optimization problem is proven to be a non-deterministic polynomial-time hard (NP-hard) problem [3]. Several studies have been conducted which have sought to optimize the job sequence in a flow shop configuration using lean principles, simulation, genetic algorithm (GA), tabu search, and particle swarm optimization (PSO) [4][5][6][7]. Shewchuk and Guo (2012) have utilized lean to optimize the wall panel stacking and sequencing in residential construction [8]. However, few studies have been carried out to optimize panel production sequence within the home building industry.

Discrete-event simulation (DES) has been used extensively to model flow shop manufacturing processes. Hammad et al. (2002) have developed a simulation model for manufactured housing processes to improve both productivity and the quality of the product [9]. Garza-Reyes et al. (2012) have used DES to find the optimum line balance for every stage in the home production process [10]. DES-based optimization has also been applied to construction and...
manufacturing in order to find near-optimal solutions. Dengiz and Alabas have used a simulation model together with a tabu search to find the optimum number of kanbans in a just in time (JIT) system [11]. Rezg et al. (2004) have proposed a methodology combining simulation and GA to optimize maintenance and inventory control policies [12]. Lu et al. (2008) have developed an automated resource-constrained critical path analysis using DES and PSO [6]. More recently, Mahdavi et al. (2011) have developed a modified chaotic ant swarm simulation-based optimization model to solve the flexible manufacturing system scheduling problem [7].

As the construction industry is moving towards factory built houses, it is important to sequence the jobs at the plant to improve the productivity by implementing the knowledge of the manufacturing industry. This paper describes a methodology which begins by integrating DES with an optimization algorithm to find the optimal production sequence for a wood-frame panel production line. The formulation of the DES model of the production line and the integration of the model with a PSO algorithm is then presented. Finally, the model results are compared with the current production sequence to measure the effectiveness of this method.

2 Methodology

The methodology is developed using DES-based optimization for the panel production sequence in a flow shop configuration. Wood-frame wall panel information is uploaded into a database from the 3D model. In the optimization environment, the simulation model reads panel information such as length; width; type; and number of studs, windows, doors, and sheets of sheathing from the database and runs the simulation in order to obtain the makespan for all the panels. The panel sequence is then updated in the database based on the optimization algorithm, and the simulation model is run again based on the new sequence. This process continues until the optimal sequence is achieved. The methodology is shown in Figure 1.

The database containing the panel information and sequence is linked with both the simulation model and optimization algorithm. The simulation model reads panel information from the database and calculates the panel processing time in each station. The optimization algorithm updates the panel sequence in the database based on the makespan for all the panels. After finding the optimal panel sequence, the algorithm is terminated, with the optimum sequence and makespan outputted from the model.

Figure 1. DES-based optimization to find optimal panel sequence.

3 Simulation Model Formulation

The DES-based optimization method is implemented at Landmark Building Solutions (LBS), a wood-frame wall panel fabrication plant located in Edmonton, Alberta, Canada. A simulation model of LBS’s wood-frame wall panel production line is developed in Simphony.NET, an integrated environment effective for building simulation models of construction activities [13]. The user builds a simulation model within Simphony by creating instances of modeling elements that resemble real components of a system/process, and linking them together in ways similar to those that exist in a real system. The following sections describe Landmark wall production line system, construction of the task time formula and the simulation logic.

3.1 Landmark Wall Production Line

Landmark Group of Builders, a major production homebuilder in Alberta, Canada, has established a
wood-frame panel fabrication plant in east Edmonton, where wood-frame open-wall panels and floor panels are produced and transported to the site for on-site assembly. This manufacturing facility is equipped with state-of-the-art production lines utilizing computer numerical control (CNC) technology which are capable of producing building components for 3 homes in an 8-hour shift.

The wall panel production begins at the framing station, where exterior and interior walls are assembled using CNC machinery. To maximize the utilization of the CNC table, what will ultimately be divided into single-wall panels, which are fabricated as multi-wall panels equal in length to the maximum length of the CNC table (40 feet). From the framing station, the multi-wall panels move to the sheathing station, where the sheathing for the exterior walls is placed by workers at the station and then nailed using another CNC machine, known as the multi-function bridge. All multi-wall panels (both interior and exterior) are also marked with their panel identification number at this station. The exterior multi-wall panels then advance from the sheathing station to the spray booth for application of spray-foam insulation. Interior multi-wall panels, alternatively, advance directly from the sheathing station to the interior wall waiting line. Using a transfer cart, the interior multi-wall panels are also cut into single-wall panels, and then moved to the interior packaging area for shipment. All exterior multi-wall panels are cut into single-wall panels after spraying and moved to the exterior wall waiting line. Exterior panels without windows and doors are moved to the wall magazine line for shipment, while those requiring windows/doors are first moved to the window/door installation station and then on to the wall magazine line.

3.2 Construction of Task Time Formula

In order to simulate the wall production line, a time study is conducted at each station to develop a task time formula by which to calculate the processing time of the panel at every station. The task time formula is developed based on the time needed to perform each task for a given panel and station. Equation (1) shows the developed task time formula for the framing station.

\[
\text{Process time (sec)} = T.B. + S \times 9.92 + M \times 29.59 + L \times 20.58 + W \times 77.05 + D \times 44.43 + \text{Drill} \times 5.40 + \text{Cut} \times 7.10 + \text{Nail} \quad (1)
\]

Where T.B. is the time needed to place the top and bottom plates (sec); S is the number of single-studs; M is the number of multiple-studs; L is the number of L-studs and double-studs; W is the number of windows; D is the number of doors; Drill is the number of drill holes in the panel; Cut is the number of cut-zones in the panel; and Nail is the Time needed to refill the nails (sec).

The time needed to place top and bottom plates, single- and multiple-studs, and windows and doors at the framing station are observed and recorded. The average time to perform each task is used to generate the above equation. The delay time is also observed and categorized into different delay types. For the framing station, there are primarily four types of delays: machine breakdown, material supply delay, worker away, and error correction. The frequency of each type of delay is observed and the probability of occurrence is calculated. Each type of delay is fitted in a triangular distribution. Similarly, the task time formula is constructed for other stations in the wall production line. The task time formula is validated by comparing the calculated process time with the actual process time.

3.3 Simulation Logic

The wall panel production process is simulated in Simphony.NET, with the simulation flowchart shown in Figure 2. n number of multi-wall panels are created as the model entities. All entities (multi-wall panels) go through a “set attribute” function where all the panel attributes, i.e., panel length, width, and type; spray area; number of walls studs, windows, doors, and sheetrock panels; and panel sequence, are read from the database. Each entity (multi-wall panel) then goes through different stations (tasks) i.e., framing station, sheathing station, spray booth, transfer cart, and window installation station. The task time at every station for each panel is calculated based on the task time formula. The framing station, sheathing station, and transfer cart have one resource each and the spray booth and window installation station have two resources each. Each wall panel has to capture the resource before entering the station and can release the resource after capturing the resource of the following station. If the resource is not available, the panel will wait in the previous station. Following completion of the spray booth task, each exterior multi-panel entity in turn generates multiple entities (single-panel) based on the number of walls in each multi-wall panel. After completing all the tasks, the total production time for
each panel is stored and the entity quits the simulation model.

Figure 2. Simulation flowchart

Figure 3 shows the simulation model developed in Simphony. The model consists of several composite elements representing different stations. The framing station composite element is shown in Figure 3. The framing station model includes a capture element to capture the resource and a task element to simulate the process time. Another composite element, delay, is used to calculate the delay time. Inside the delay, the probability branch is used to capture the likelihood of each type of delay occurring. Following the delay element calculations, the entity moves out of the framing station. Also, in between tasks the model records the time using the set attribute element.

4 Particle swarm optimization

PSO is a population-based evolutionary algorithm proposed by Kennedy and Eberhart [14]. In the PSO algorithm, the search is performed by a set of particles (i), and the information is shared between all the particles in order to find the optimal solution. Each particle is considered as a point in a D-dimensional space and has a velocity and position value. The position and velocity values of the \( i \)-th particle are denoted as \( x_i = (x_{i1}, x_{i2}, ..., x_{iD}) \) and \( v_i = (v_{i1}, v_{i2}, ..., v_{iD}) \), respectively. Each particle moves towards the best solution of the entire swarm (i.e., the "global best") by updating its position and velocity values after every iteration. Initially, the position and velocity values are assigned randomly to each particle in order to start the search. Then the values are updated based on the results of all previous iterations following Equations (2) and (3).

\[
k_{id}^{k+1} = w \cdot k_{id}^k + c_2 \cdot r_2 \cdot \left( p_{id}^k - x_{id}^k \right) + c_2 \cdot r_2 \cdot \left( p_{gd}^k - x_{gd}^k \right)
\]

\[
x_{id}^{k+1} = x_{id}^k + (k+1)_{id}
\]

Each particle’s best position is represented by \( p_{id} \) and its global best position is represented by \( P_{gd} \). \( c_1 \) and \( c_2 \) are the cognitive parameter and social parameter, respectively. In this model, both values are set to 2. \( r_1 \) and \( r_2 \) are random numbers uniformly distributed from 0 to 1. \( k \) is the iteration number and \( w \) is the weight inertia required to control the impact of the previous velocity value on the current velocity. In this model the value of \( w \) is set to 0.9 at the start and then is decremented by a factor of 0.975 after every iteration (i.e., \( w^{k+1} = 0.975 \cdot w^{k} \)). The search process is terminated once the maximum number of iterations is reached.

In order to implement the PSO algorithm in a sequencing problem, a heuristic rule called Smallest Position Value (SPV) is applied [15]. In this panel sequencing problem, each particle has a continuous set of position values representing every multi-wall panel. If the model is run to optimize the sequence of 50 multi-wall panels, each particle in the PSO search will have 50 position and velocity values representing each multi-wall panel. Since the particle itself cannot represent a sequence, it is determined by the position values \( (x_{id}, x_{id}, ..., x_{id}) \) of the particle \( x_i \). According to the SPV rule, the panel with the smallest position value will be first in the production sequence; the panel with the second-lowest position value will be second in the sequence, and so on. After running the simulation, the position value is updated based on the fitness value, and a new sequence is generated following the SPV rule.
Figure 3: Simulation model in Simphony.NET.

Figure 4 illustrates the integration between the PSO algorithm and the simulation model. The PSO search algorithm’s utilization of DES for the panel sequencing problem is summarized as follows:

- **Step 1:** The user uploads multi-wall panel information (length; width; type; and number of windows, doors, studs, sheets of sheathing, and walls) into the database.

- **Step 2:** The PSO model assigns initial position and velocity values to every panel under each particle and updates them in the database. The position values are generated randomly between 0.0 and 4.0. Initial velocities are created randomly between -4.0 to 4.0.

- **Step 3:** The simulation model reads the panel information from the database. The panel is sorted based on the SPV rule, and the simulation model is run for every particle, with the total makespan stored as the personal best value \( (p_{lb}) \) for each particle.

- **Step 4:** The PSO model updates the iteration counter: \( k = k + 1 \).

- **Step 5:** The PSO model updates the inertial weight: \( w^k = w^{k-1} \cdot \alpha \), where \( \alpha \) is a decrement factor.

- **Step 6:** The PSO model updates the velocity of each particle satisfying Equation (2).

- **Step 7:** The PSO model updates the position of each particle in the database satisfying Equation (3).

- **Step 8:** The PSO model applies the SPV rule to determine the panel sequence for each particle.

- **Step 9:** The simulation model runs for each particle and stores the total makespan.

- **Step 10:** The PSO updates the personal best value \( (p_{lb}) \) and position \( (x_{lb}) \) for each particle. If the current fitness value \( (f^k) \) is less than the personal best value \( (p_{lb}) \), then \( p_{lb} = f^k \) and the personal best position, \( x_{lb} = x_{id} \).

- **Step 11:** The PSO model updates the global best value \( (p_{gb}) \) and position \( (x_{gb}) \) by taking the minimum value of the personal best and the corresponding position value. Furthermore, the global best value \( p_{gb} = \min \{p_{lb}\} \). If the fitness value for the current iteration \( (f^k) \) is less than the global best value \( (p_{gb}) \), then \( p_{gb} = f^k \) and the global best position, \( x_{gb} = x_{id} \).

- **Step 12:** If the number of iteration exceeds the maximum number of iteration, the PSO model stops the search; otherwise, it advances to step 4. Once the search stops, the optimum panel sequence is
defined as the resultant sorted global position values from smallest to largest, while the global fitness value is defined as the total production time.

5 Results and discussion

Initially, the optimization model is run for different iteration and particle numbers for 50 multi-wall panels in order to find the optimal number of particles and iterations by comparing the model result with the runtime. The model is run in Intel® Core™ i7 CPU (3.20 GHz). Table 1 presents the model runtimes and makespans corresponding to the optimal sequences for different numbers of particles and iterations. The total production time of all panels (makespan) varies from 512–550 min. The result shows that 20 particles and 20 iterations can provide satisfactory results within a reasonable model runtime of 26 min, while increasing the numbers of particles and iterations to greater than 20 does not have a significant effect on the optimization results.

Table 1 Model runtimes and makespans for different particle and iteration numbers (50 multi-wall panels).

<table>
<thead>
<tr>
<th>Particle number</th>
<th>Iteration number</th>
<th>Model runtime (hr:min)</th>
<th>Makespan for optimal sequence (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>0:02</td>
<td>550</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>0:15</td>
<td>540</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0:26</td>
<td>514</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0:26</td>
<td>522</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>0:27</td>
<td>524</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>0:40</td>
<td>523</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>0:27</td>
<td>546</td>
</tr>
<tr>
<td>30</td>
<td>15</td>
<td>0:31</td>
<td>526</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>1:03</td>
<td>514</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>1:00</td>
<td>521</td>
</tr>
<tr>
<td>30</td>
<td>30</td>
<td>1:01</td>
<td>512</td>
</tr>
<tr>
<td>30</td>
<td>50</td>
<td>1:42</td>
<td>522</td>
</tr>
</tbody>
</table>

The optimization model is run for different production dates in the LBS plant for 20 particles and 20 iterations. The simulation model, meanwhile, is run 10 times for the actual and optimal sequences, respectively, to provide the mean makespan. The result is summarized in Table 2. The model provides the mean makespan for the actual sequence, the mean makespan time for the optimal sequence, along with the maximum and minimum production times. The productivity improvement is calculated by comparing the mean makespans of the actual and optimal sequences. The results show that the productivity can be improved up to 10% by implementing this optimization model.

Table 2 Optimization results.

<table>
<thead>
<tr>
<th>No. of multi-wall panel produced</th>
<th>Mean Makespan-Actual sequence (min)</th>
<th>Min. Makespan-Optimal sequence (min)</th>
<th>Mean Makespan-Optimal sequence (min)</th>
<th>Max. Makespan-Optimal sequence (min)</th>
<th>Productivity Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>489</td>
<td>440</td>
<td>454</td>
<td>475</td>
<td>7%</td>
</tr>
<tr>
<td>66</td>
<td>691</td>
<td>665</td>
<td>681</td>
<td>705</td>
<td>1%</td>
</tr>
<tr>
<td>68</td>
<td>744</td>
<td>658</td>
<td>670</td>
<td>691</td>
<td>10%</td>
</tr>
<tr>
<td>57</td>
<td>628</td>
<td>551</td>
<td>573</td>
<td>593</td>
<td>9%</td>
</tr>
<tr>
<td>70</td>
<td>775</td>
<td>713</td>
<td>728</td>
<td>746</td>
<td>6%</td>
</tr>
</tbody>
</table>
Figure 5(a) shows the convergence chart of the PSO search for all particles (400 iterations) and Figure 5(b) shows the convergence chart of one particle (20 iterations). The optimum result was found after approximately 200 iterations of all particles. After this point the particles did not find any better solution.

6 Conclusions

This paper presents a research that integrated DES with a PSO algorithm in order to find the optimal panel sequence for a prefabricated wood-frame wall panel production line. The methodology is implemented on Landmark Group of Builders’ prefabricated multi-wall panel production line. The simulation model of the production line is developed in Simphony.NET and integrated with a PSO algorithm using a central database containing multi-wall panel information. The optimization model is run for several actual production dates, and the optimization results are compared in order to measure the productivity improvement associated with the utilization of the model. The successful implementation of the proposed method in an actual production line demonstrates the practical usefulness of this model within the evolving panelized home building industry.

In future, the proposed methodology can be further improved by implementing other search algorithms such as genetic, ant colony, and bee algorithms, and the results can be compared to find the best algorithm for this type of sequencing problem. The search algorithm can also be modified based on the given requirements of a production team and the results can be observed using the simulation model.

Acknowledgements

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC). The participation and support of our collaboration partner for this research, Landmark Group of Builders Ltd. is also appreciated.

References


Integrated Visualization and Simulation for Lifting Operations of Modules under Congested Environment

SangHyeok Han\textsuperscript{a}, Zhen Lei\textsuperscript{a}, Ahmed Bouferguène\textsuperscript{b}, Mohamed Al-Hussein\textsuperscript{a}, and Ulrich Hermann\textsuperscript{c}

\textsuperscript{a}Hole School of Construction Engineering, University of Alberta, Canada
\textsuperscript{b}Campus Saint-Jean, University of Alberta, Canada
\textsuperscript{c}PCL Industrial Management Inc., Canada

Corresponding author email: sanghyeo@ualberta.ca

Abstract-
Modular-based heavy construction projects are recognized as faster, safer, and more efficient than traditional those completed through on-site construction methods. The successful completion of these projects relies on an efficient material handling system, especially when mobile cranes are utilized to satisfy heavy lift requirements. However, engineers are faced with the challenge of planning the lifting operations of heavy modules within congested areas. Currently, mobile crane analysis is implemented manually, but requires time-consuming data input resulting in an increase of errors and a lack of proper crane productivity performance analysis; this complicates the planning process. To overcome these limitations, the research presented in this paper proposes to integrate visualization and simulation, in an approach known as post-3D visualization simulation, in order to plan collision-free crane lifts by eliminating potential errors in 3D visualization and to predict crane productivity performance in simulations. This proposed system will contribute to the successful completion of construction projects with high productivity and site-error reduction by selecting the best crane operation that includes various crane lifts. An actual industrial project which has a number of constraints, including space limitations, different types of site layouts, and various crane lifts, is used to validate the proposed framework.

Keywords -
- Mobile crane; post-3D visualization simulation; crane performance; productivity; industrial projects

1. Introduction

The Athabasca Oil Sands in Alberta, Canada is becoming one of the most attractive industrial construction markets in the world. Oil sands-related projects are built utilizing the modular-based (off-site) construction method. The successful completion of modular-based industrial construction projects, in terms of saving time and reducing associated costs, relies heavily on a well-designed material handling systems with regards to the operation of construction equipment. Among this construction equipment, mobile cranes are frequently utilized in order to lift modules to their positions for installation since they satisfy the required lift capacity requirements. This frequent utilization of mobile cranes is due to the fact that modules in modular-based industrial construction projects are becoming larger and heavier.

Crane lift studies can be divided into the following categories: (i) crane type selection and location [4, 13, 21, 22, 24, 25]; (ii) crane support system [11]; and (iii) crane lift path planning [1, 3, 7, 17, 19, 23]. When lift studies are implemented, the most critical step in terms of safety and productivity is to identify and eliminate potential collision errors. To facilitate this step efficiently and effectively, many researchers have searched for better algorithms and superior techniques such as simulation, 3D visualization, and integrated systems [5, 14, 18, 26]. Corresponding to this trend, the effectiveness of 3D visualization has been reported in actual cases [9, 19]. Crane productivity analysis which takes into account CO\textsubscript{2} emissions has been introduced for utilization of a tower crane on a high-rise building project [11]. However, previous research has not been fully adopted in practice, especially in heavy industrial construction projects where complicated and time-consuming algorithms are involved. The current mobile crane analysis is implemented manually but requires time-consuming data input, resulting in an increase of errors and a lack of proper crane productivity performance analysis; this leads to complication of the planning process. To overcome this lack of proper crane productivity performance analysis, crane performance should be evaluated among various scenarios of crane operations in order to select the best crane operation. This paper thus proposes to integrate 3D visualization and simulation in a so-called post-3D visualization.
simulation system. This system is able to design feasible motions of mobile crane operation in 3D visualization, and has the ability to evaluate the associated crane operation automatically in terms of crane performances in the simulation so that lifting engineers and project participants can actively use them to select the best crane operation. A case study constructed by PCL Industrial Management Inc. is used to validate this methodology.

2. Methodology

Among various crane types, such as hydraulic truck cranes, all-terrain cranes, rough-terrain cranes, crawler cranes, and carry-deck cranes, crawler cranes equipped with superlift (tail-swing) are selected for this research in order to accommodate the high lift capacity requirements of heavy industrial projects. According to previous research [8], mobile crane operation studies can be divided into two categories corresponding to the type of operation: (i) crane binary checking [14], which involves checking pick from fixed position (PFP) operations to identify the existence of a path and return a binary result (yes/no); and (ii) crane walking path checking, which searches for pick and walking operations (PWOs) to explore feasible walking paths. PWOs involve the crane walking from a given position for loading to a point from which the crane can reach the set positions of the lifted objects when PFPs are not possible. Since modular-based industrial construction sites consist of numerous modules in congested areas, spatial constraints such as the geometry of the mobile crane and existing obstacles (e.g., modules already installed before installation of lifted objects) are the main issues to consider when designing safe and efficient crane operations. Therefore, virtual motion planning of crane operation (VMCO) plans detailed motions of mobile crane operations by detecting and eliminating potential collision errors in order to validate the PFP or PWO for lifted objects. At this juncture, it each lifted object may have more than one possible crane operation. This is due to the fact that crane lift binary checking and crane walking path systems only search for the feasibility of crane operations on numerous locations for each individual lifted object. In analyzing crane performance, crane lift binary checking, crane walking path, and virtual motion planning systems are inadequate. As a result, these systems can lead to difficulty in selecting the best crane operation when the lifted object has more than one possible crane operation. To eliminate this inadequacy in the performance systems, this paper proposes to integrate 3D visualization and simulation, known as post-3D visualization simulation, in order to predict the crane performance when lifting objects. Industrial projects involve a large number of lifts, thereby challenging the manual system of implementation in lift studies. All the systems in the proposed methodology are therefore executed automatically.

![Figure 1. Proposed Methodology](image-url)
As shown in Figure 1, the simulation is implemented for the evaluation of crane performance after designing detailed motions of mobile crane operations in a 3D environment. The input data consists of: (i) crane information, including crane configurations, crane speeds, and capacity charts provided by manufacturers; (ii) rigging (e.g., spreader bar, sling, and hook) information, such as total weights based on different module types; (iii) module information, including module weights; and (iv) motion sequences and associated lift angles. The output of the post-3D visualization simulation is the best crane operation in terms of cycle times. The VMCO controller is an execution tool for the post-3D visualization simulation, which is developed using Maxscript, a built-in language tool in 3D Studio (3ds) Max.

2.1 Post-3D Visualization Simulation

Computer simulation is used to design a mathematical-logical model of a real world system and experiment with the model on a computer. The purpose of using a simulation is to eliminate unforeseen bottlenecks, to use resources more effectively, and to optimize system performance before an existing system is changed by the proposed design. 3D visualization, meanwhile, is recognized as a comprehensive 3D computer-aided design tool used to design the construction facility and its associated schedule; when these two aspects interact, the workability of a schedule is ensured and alternative construction strategies in terms of spatial and workflow issues are configured. However, each of these systems has its own limitations [10]. To reinforce their functionalities, some researchers have integrated simulation and 3D visualization in a system referred to as post-simulation visualization [10]. This system runs the simulation in order to analyze the productivity of a project, and then identifies spatial issues in order to check workability in 3D visualization. However, this system is not currently used for crane lift studies, which require the analysis of spatial issues in oil sands projects that involve large numbers of lifted objects, complexity, and various site layout changes. Each lifted object also may have more than one feasible crane operation, leading to challenges in selecting the best crane operation. Therefore, this paper proposes a post-3D visualization simulation which plans mobile crane operations in a simulation by calculating the required lift angles of each crane configuration and preventing potential collision errors [8]. It then simulates crane performance which is addressed in this paper as a means to analyze the process times of crane operations for lifted objects.

3ds Max, a modeling tool, and Simphony.NET [6], a simulation tool, are used to demonstrate the effectiveness of the proposed methodology. Previous studies [2, 10] have highlighted four limitations encountered in integrating these tools: (1) lack of information exchange; (2) insufficient automation; (3) only one-way information exchange from simulation to 3ds Max; and (4) inefficient information checking. These limitations result in an ASCII file type as a intermediate step whenever users try to implement information exchange from the simulation to 3ds Max. As well, outputs generated from both simulation and visualization tools are not represented in an environment, so this may lead to poorly managed decision making between project participants. To capture evolving technologies, 3ds Max provides better information exchange that connects to external databases such as Microsoft Access using the VMCO controller developed in 3ds Max. Corresponding to this technology, this paper uses a Dynamic Link Library (DLL) file type, described in Figure 2, which provides a library of executable functions that can be used by windows applications to implement information exchange quickly and easily, and to execute functionalities of Simphony.NET in 3ds Max without opening a simulation model to manually input data.

To analyze crane performance for lifted objects, the VMCO controller transfers the following information: (1) module ID; (2) tracking ID; (3) radii during each motion of crane operation; (4) total weight; (5) capacity; (6) factor; (7) motion ID; (8) lift angles; (9) lift height; and (10) speed. Since there is not enough time data (probability data) to implement a simulation model in order to predict mobile crane productivity, this paper uses the simulation as a deterministic model. The VMCO controller determines the process time of each motion of the crane operation, based on the speed and lifting capacity provided by manufacturers and the total weights satisfied in Equation (1). At this juncture, total weight may have variation since the various modules may have differing weights and be loaded by numerous types of spreader bars.

$$W_{Total} = W_{Lifted} + W_{Spreadbar} + W_{Sling} + W_{hook}$$  \(1\)

Where $W_{Total}$ is the total weight; $W_{Lifted}$ is the weight of the lifted object; $W_{spreadbar}$ is the weight of the spreader bar; $W_{sling}$ is the weight of the sling; and $W_{hook}$ is the weight of the hook.

After the total weight is calculated, the factor is
calculated by the total weight divided by the crane lift capacity setting identified in the crane lift capacity chart; this capacity is based on the working radius from center position of crane to center position of lifted object for a specific motion in 3D visualization. The factor indicates how much of the crane lift capacity is occupied by the total weight. Using this factor, speed can be also calculated.

Based on site layout and the sequences of crane operation shown in Table 1, the virtual motion planning system plots detailed motions of the mobile crane operations for lifted objects. It should be noted that the sequences of crane operation used in this paper follow practical crane operation, where the crane lifts the object as close as possible to the ground after loading (motion ID 6). The rigging system is lifted up to maintain particular clearances as defined by users (motion ID 13) whenever the distance between the lifted object and existing obstacles on-site is less than or the same as the clearance during the crane operation across animation time. Based on these sequences, a simulation model is built in Simphony.NET 4.0.

To achieve a post-3D visualization simulation system, the DLL file in 3ds Max is used to integrate the simulation and 3D visualization with automated information exchange. That is, the VMCO controller delivers the required information to the simulation model through the DLL file developed in Visual Studio.NET, which represents module ID, tracking ID, and cycle time of the mobile crane operation generated from the simulation model. The cycle times are used to identify the best crane operation (i.e., the shortest cycle time) by comparing the cycle times of the crane operations. As a result, the system can assist lift engineers and project managers to implement decision making quickly and easily by selecting the best crane operation. This is achieved by comparing the cycle times of various crane operations, and by designing the lift schedule in a manner which takes into account other material handling systems as well as the entire project schedule.

### Table 1. Sequences of crane operation

<table>
<thead>
<tr>
<th>Motion ID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Crane base Rotation</td>
</tr>
<tr>
<td>2</td>
<td>Superstructure Rotation</td>
</tr>
<tr>
<td>3</td>
<td>Boom Rotation</td>
</tr>
<tr>
<td>4</td>
<td>Rigging Rotation</td>
</tr>
<tr>
<td>5</td>
<td>Rigging Down for load</td>
</tr>
<tr>
<td>6</td>
<td>Rigging up with load</td>
</tr>
<tr>
<td>7</td>
<td>Superstructure Rotation with load</td>
</tr>
<tr>
<td>8</td>
<td>Crane walking with load</td>
</tr>
<tr>
<td>9</td>
<td>Superstructure Rotation with load</td>
</tr>
<tr>
<td>10</td>
<td>Boom Rotation with load</td>
</tr>
<tr>
<td>11</td>
<td>Rigging Rotation with load</td>
</tr>
<tr>
<td>12</td>
<td>Rigging Down with load</td>
</tr>
<tr>
<td>13</td>
<td>Rigging Up</td>
</tr>
</tbody>
</table>

3. Implementation

This paper implements the methodology in a modular-based oil sands project constructed by PCL Industrial Management Inc. in Alberta, Canada. Target areas in the case study have a total of 150 lifting cases. The PFP method is required for 140 cases, while 10 modules are installed by the PWO method. The Demag CC 2800 crawler crane equipped with superlift is selected as the crane type in this case study (Capacity: 660 tons, Boom length: 276 ft). To clearly illustrate the proposed methodology, this paper focuses on the PWO to deliver module ID 103 (Figure 2 describes the site layout). The VMCO reads information from a central database, based upon which it automatically builds 3D models such as a module, existing obstacles, and site boundary limits (ISBLs), which are inaccessible areas for the crane. The dimensions of the module are 6 m (19.68 ft) × 36.6 m (120.08 ft) × 4.1 m (13.45 ft), and the two crane walking paths (tracking IDs 12049 and 12068) are found to be available.

Based on two feasible crane walking paths (WPs), including a start point (SP) for loading and a finish point (FP) for unloading, the VMCO controller automatically develops 3D visualization of mobile crane operations by calculating lift angles of each crane configuration and maintaining the clearances defined by users. Figure 3 represents the virtual motions of crane operation for tracking ID 12049 and 12068 in module ID 103. For example, tracking ID 12049, the crane operation to lift the module, is designed based on the following sequence: (1) move and rotate by 73.9° in the counter-clockwise direction beginning at the SP of the WP; (2) rotate the superstructure by 81.1° in the clockwise direction to a pick point of the module; (3) rotate the boom upward by 6.8° to coincide with the location of the rigging system; (4) rotate the rigging system by 98.8° and lift...
it up and down for loading; (5) rotate the superstructure by 108.4° to the FP of the WP; (6) walk the crane; and (7) repeat steps (2) to (4) with other variables until the module is installed successfully on its set position. According to this sequence of crane operation, the VMCO controller generates the relevant information (see Table 2): (1) module ID; (2) tracking ID; (3) weight; (4) radius; (5) capacity; (6) factor; (7) motion ID; (8) lift angle; (9) lift height; and (10) various speeds. In the same rule described above, the 3D visualization for the other feasible crane location is developed and relevant information is generated.

(a) Tracking ID 12049

(b) Tracking ID 12068

Figure 3. Crane operations for module ID 103

To evaluate the expected performance of each crane operation, the following information is required for the simulation: (1) module ID; (2) tracking ID; (3) lift angle; (4) lift height; (5) speed; and (6) motion ID. The VMCO controller transfers this information to the simulation model efficiently and effectively through the DLL file. Once the simulation model has run completely, the cycle times of each crane operation (outputs of the simulation) are also shifted to the VMCO controller through the DLL file. This workflow of the automated post-3D visualization simulation system is successful in the information exchange. As a result, the VMCO controller represents the module ID, the tracking ID, and the expected cycle time of each crane operation (Figure 4) in 3ds Max. Based on this information, users can identify which crane WP is the best option to ensure high crane performance and project productivity. Tracking ID 12049 for module ID 103 is found to be the crane location which facilitates the best crane operation, with a cycle time of 11.8253 minutes for the lift.

Figure 4. Simulation outputs in VMCO controller

4. Conclusion

Industrial construction projects generally involve numerous objects to be lifted by a crane in congested areas. Previous research studies in industrial construction have focused on: (1) identifying feasible crane lift methods such as the pick from fixed position (PFP) and pick and walking operation (PWO); and (2) planning detailed motions of mobile crane operation using 3D visualization. However, this system lacks decision making tools by which to select the best crane lift path in terms of high crane performance (productivity) when lifted objects have more than one possible crane lift path. To overcome this deficiency, this paper has presented a system integrating 3D visualization and simulation, known as automated post-3D visualization simulation. The
The purpose of this system is to select the best crane operation by analyzing crane performance in the simulation based on the 3D visualization of crane operations. This system has been successfully implemented in a case study by representing the expected cycle times of crane operations for the lifted object. We are presently expanding the work scope such that the future system will be able to apply this system to plan an optimal material handling system, as well as project schedules to increase project productivity.

5. Acknowledgements

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development (CRD) program. The participation and support of our collaboration partner for this research, PCL Industrial Management Inc., is also appreciated.

Table 2. An example of the information for simulation

<table>
<thead>
<tr>
<th>ObjID</th>
<th>TrackingNo</th>
<th>Weight</th>
<th>Radius</th>
<th>Capacity</th>
<th>Factor</th>
<th>LiftAngle</th>
<th>LiftHeight</th>
<th>Speed</th>
<th>MotionID</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>12049</td>
<td>54762</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>-71.99176025</td>
<td>0</td>
<td>131.939665</td>
<td>1</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>54762</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>81.1828537</td>
<td>0</td>
<td>131.939665</td>
<td>2</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>54762</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>81.1828537</td>
<td>0</td>
<td>131.939665</td>
<td>3</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>54762</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>98.82134247</td>
<td>0</td>
<td>131.939665</td>
<td>4</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>54762</td>
<td>144.0997314</td>
<td>286600</td>
<td>19.1074685</td>
<td>0</td>
<td>53.67518997</td>
<td>206.2866199</td>
<td>5</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>314952</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>-71.99176025</td>
<td>0</td>
<td>131.939665</td>
<td>6</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>314952</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>81.1828537</td>
<td>0</td>
<td>131.939665</td>
<td>7</td>
</tr>
<tr>
<td>103</td>
<td>12049</td>
<td>314952</td>
<td>144.1074524</td>
<td>286600</td>
<td>19.1074685</td>
<td>98.82134247</td>
<td>0</td>
<td>131.939665</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 5. A simulation model of mobile crane operation in Simphony.NET 4
References


Post-Simulation Visualization Application for Production Improvement of Modular Construction Manufacturing

M. Moghadam1, B. Barkokebas2 and M. Al-Hussein3

1Hole School of Construction, Department of Civil and Environmental Engineering, University of Alberta, Canada
E-mail: mana.moghadam@ualberta.ca, barkokeb@ualberta.ca, malhussein@ualberta.ca

Abstract -

The modular construction manufacturing (MCM) process is a complex operation that combines line flow product movement with a complex activity precedence network. There are physical constraints related to the given facility, as well as logical constraints caused by demand variation. In order to change production line layout and make improvements within the context of Lean, a tool is needed to assist MCM to quantify, at the planning and evaluation stages, the benefits they can expect from proposed changes to their system. Simulation is a technique by which to facilitate identifying changes and benefits of future transformation and to determine where valuable resources should be applied prior to actual implementation.

Despite the benefits of simulation, project management teams typically are unwilling to make decisions based on current simulation outputs, since they are very difficult to understand and require specialized skills in interpretation of the information. Visualization is a more popular technique since it fosters better understanding of the construction process. However, to be effective for decision making purposes, such a model must be linked to project information. The visual interpretation of the simulation results constitutes a more effective approach. In this paper, a simulation model is thus generated to provide results for different production scenarios, and then the near-optimum scenarios are run to visualize the production constraints, which facilitates precise scenario comparison. The developed model capitalizes on the advantages of both simulation and visualization, whereby critical information such as the 3D model, time constraints, and resource demand are incorporated into the system. The proposed methodology is validated by a case study, a residential modular factory in Edmonton, Canada, which illustrates the effectiveness of the proposed methodology.

Keywords - Modular Construction Manufacturing; Lean; Production; Simulation; Visualization; Resource Utilization

1 Introduction

Modular construction manufacturing (MCM) provides opportunities to apply Lean for production efficiency in the plant, thereby eliminating waste and supporting the delivery of products in a shorter time and at a lower cost. The decision to implement Lean is complicated because of the substantial differences between a traditional system and a Lean system. What is needed is a tool to assist organizations considering Lean to quantify, at the planning and evaluation stage, the benefits they can expect from applying Lean. This tool should be adaptable to the specific circumstances of the organization, and should be capable of generating resource requirements and performance statistics for both the proposed Lean system and the existing system [1]. Simulation is an effective technique to facilitate identifying changes and benefits of Lean transformation. Simulation is a computer-based tool that represents real-world objects and processes in order to effectively evaluate and examine various scenarios prior to implementation and facilitate the decision-making process [2]. The impact of Lean transformation can be analyzed to determine where valuable resources should be applied before actual implementation. This increases confidence and is likely to hasten the rate of adoption of Lean, as it provides a visual and dynamic illustration to management of how the new system would work [1].

On the other hand, to make improvements on a production line within the context of Lean, the integration of simulation tools and Lean brings about a more effective approach for process management. A simulation model of the production line stream map is able to challenge the impacts of Lean application on line balancing, as well as the results [3]. Significant impact from random variation in the nature of modular manufacturing, as well as system element interactions, is expected for the future-state of the system. Currently there is no tool in Lean which contains variability data and provides performance evaluation of the system. Therefore, incorporating simulation modeling into the Lean transformation process is able to provide a quantitative conception of the potential benefits [4].
Although the simulation tool is considered to be efficient in evaluating construction processes, the project management team often views it as a “black box” that can only be understood by highly skilled managers. The gap between the specialist and the management team leads to misunderstandings regarding simulation results, and may result in incorrect interpretations [5]. The project management team is unwilling to make decisions based on current simulation outputs, such as statistics-based charts and diagrams, because they do not provide adequate information related to construction process requirements [6]. The results and analysis are difficult to understand, and translation of the data must be done by experts [7]. Still, visualization is widely accepted in construction since it promotes better understanding of process and performance of the construction system. However, the model needs to be associated with the project data in order to be effective and valuable for decision making purposes. The key aspect that is added to the output analysis of the simulation process is visualization, leading to post-simulation analysis. The produced results from a simulation model are interpreted and visualized for further analysis, which is a popular and widespread decision making technique [8].

The MCM industry is a growing industry that seeks innovative approaches to increase profitability. To improve the production efficiency in the plant, an enhanced tool is required to present the statistical results of changes in production sequencing and crew selection in a visualized model to facilitate decision making. The application of visualization to construction in general differs from its application to MCM in terms of production flow of multiple projects, product movement through the production line, work cell design, precedence network, lead time, plant physical constraints, and crew selection. Also, with MCM the production progress is not measured based on the physical project progress, but instead is determined based on the total work performed in all stations, and can be calculated as the sum total of man-hours. Numerous studies have investigated the integration of simulation and visualization for on-site construction, but only a few have accomplished this for modular manufacturing. Therefore, research in the area of on-site construction is being carried out to develop and advance an integrated tool for MCM production process analysis.

In this paper, the complicated result from the simulation model of the work process provided monitoring of both the visualization and the simulation parameters simultaneously by providing additional perception through visualization, and also provided the immediate visualized or statistical results of making changes to the model. Based on that, the simulation results of near optimum scenarios are run to simulate the process and visualize the production constraints. This integrated method is a post-simulation visualization model for MCM, with an emphasis on production efficiency. The model capitalizes on the advantages of both simulation and visualization, where critical information such as the 3D model, time constraints, and resource demand are incorporated into the system.

2 Research Methodology

The methodology process in this research is divided into five major phases. In Phase I, a simulation model of production flow is developed in order to depict the production line layout, schedule, and resource requirements. The inputs for this model consist of information extracted from a BIM-generated 3D model, resource requirements, and proposed improvements. The simulation model delivers results for different production scenarios and provides the opportunity to evaluate the proposed future-state in order to optimize the production process. In Phase II, a post-simulation visualization model of evaluated production scenarios is developed which can be used by management teams as a more efficient tool for production flow analysis. As a result, the resource requirements to complete various modules are determined, along with potential scenarios for workflow balancing.

2.1 Phase I. Developing Simulation Model

Simulation tools play an important role in the application of simulation in different industries. Simulation in construction had been limited to research applications before AbouRizk and Hajjar [9] presented a framework that customized simulation for the construction industry. They used Special Purpose Simulation (SPS) as a tool to build systems for specific purposes. Later, they improved their approach and developed a construction simulation system, called Simphony, which provides a set of predefined elements representing construction requirements [10]. Simphony.NET is a construction-oriented, general-purpose discrete event simulation (DES) software application developed at the University of Alberta using process interaction concepts to create a model. A general template with predefined element functionality enables the user to select elements based on a required function for a specific simulation model. In this research, Simphony.NET 4.0 is used to develop a simulation model for the production process. The process to generate and evaluate the simulation model is presented in Figure 1. In the simulation model, the value stream map of the case-study is created to show the product
family, information and material flow, work cells, inventory amount, daily customer demand, supplier and shipping schedule, and production volume. The generated simulation model in Simphony.NET 4.0 involves two input types: Fixed and random variable. The values of fixed variables, such as number of entities, change over time, yield rate, value-added time, and transportation time between stations, remain constant during a simulation run. These fixed inputs are defined by the user according to the factory and project specifications. The values of random variable inputs, such as process time and number of operators, change according to data distributions during the simulation run. These random variable inputs are probabilistic distributions of various activities based on the module specifications.

2.2 Phase II. Developing Post-Simulation Visualization Model

The proposed post-simulation visualization (PSV) is the 3D visualization of the DES model, representing physical working environments with 3D graphical objects. The model depicts the simulation of the production process in detail, producing and displaying production flow information simultaneously such as lead time, production rate, and resource utilization for evaluation purposes. The proposed system has the capacity to be linked to all possible production information. The model needs to be flexible in order to deal with changes in process caused by module variation and to present a complex production process in a simple way. Visualization of the simulated process is proven to be an effective tool in communicating the value and simplicity of the minute-by-minute schedule. The simulation result comparison between the initial model before applying PSV and the final version after running PSV several times shows significant improvements in terms of eliminating waste, smoothing the production flow, leveling resources, and reducing idle time. A comparison is also conducted among various scenarios, by which the management team confirms the considerable impact of the PSV model in terms of decreasing activity durations, eliminating errors and rework, and identifying the best potential production scenarios using visualization of the simulated process during the planning phase. In this research, two approaches are performed to create a PSV model. The first approach integrates animation with statistical outputs of simulation model and provides a lifelike image of the process. The second approach is more dynamic, capable of responding instantly to changes made to the process.

In the first approach, to create a real-time abstract simulation model by means of computer animation, a visually simulated process is developed. After evaluating and comparing results among various potential production scenarios, a scenario with near optimum results is selected for visualization. A 3D model of the production process is also developed in Autodesk 3ds Max, representing certain activities at all the stations in the production line. The processing times of all the stations are imported from the simulation model output in the form of an ASCII file, which is transformed to a minute-by-minute schedule. Visualization of the simulated process is then performed semi-automatically. Within the 3D animation environment the high-level simulation model is transformed to a micro-level representation in frames/minute. To develop the PSV model, 3ds Max’s scripting language, MAXScript, is used. As inputs of the PSV model, two sources are required: (1) the 3D model library of PSV; and (2) the simulation model output that stores the spatial configuration of the

![Diagram](image-url)
construction process, along with performance time. The PSV model imports 3D models from the 3D library, including models of the equipment, modules, resources, and the 3D factory, and assembles them in 3ds Max. Then the 3D animation engine uses the data from the simulation output file in order to create the key frames. The model’s outputs are production processing time, labor utilization, safety and quality control, and evaluation of potential scenarios for construction operations.

The second approach is pursued to create a more dynamic visualized model from simulation results which can be modified by any user with no special software skills required. For this purpose Autodesk Navisworks Manage is selected, which combines clash finding analysis and interface management with 4D schedule simulation. The 3D model of the factory is developed in Revit and then, along with 3D models of modules generated in the design phase, is imported into Navisworks. The imported outputs from the simulation model for existing modules in the production line and coming modules to be visualized include the processing times at all stations, resource allocation plan, and work sequence. In order to analyze different scenarios for plant layout configuration with automatic clash detection, more information is added to the model, including overhead crane capacity, factory space limitations, and labor safe work area. In this model, the relationships among components are defined through parametric modeling rules and constraints, such that the model responds automatically to any changes immediately.

3 Case Study

This research was implemented on a modular factory in Edmonton, AB, Canada, that prefabricates wood structure residential modules under environmentally-controlled conditions and transports them to sites. Manufacturing provides opportunity for the company to offer time efficiency, cost effectiveness, and superior quality compared to traditional stick-built construction. All houses are built to building code for the given region, and are transported and installed in urban, rural, or remote areas. This factory fabricates custom-designed homes that vary in features, layout, and size between 600 sq ft and 1,600 sq ft.

3.1 Simulation of Production Process

In this phase, simulation models of the current and future states of the production process are generated in Simphony.NET 4.0. The current-state production process is generated based on the factory current-state VSM. Numbers of labor personnel are constant variables and activity durations are defined by data distributions. The future-state production process is generated based on set of proposed recommendations. The simulation model depicts the production line layout; individual and overall production schedules through the production line for 10 modules that vary in size and specifications; resource requirements based on each module’s dimensional properties; and the optimum Takt time to reach an optimum resource allocation plan. The inputs for this model are frontloaded from information in the BIM-generated 3D model of sample modules, which is extracted and sorted into a spreadsheet. Then man-hours requirements are calculated based on modules specification. For the case-study, similar to with on-site construction, a customer can choose from among existing floor plans or provide their own customized floor plan which accommodates their needs and lot size. As a result, the factory production line cannot be run at a steady pace, since the activities taking place at each station are contingent upon individual design. Therefore required man-hours at different work-stations are calculated based on modules dimension and specification, which is beyond the scope of this paper. Then required man-hours are imported into a database which is linked to the simulation model. The simulation model delivers results for different production scenarios and provides the opportunity to evaluate the proposed future-state in order to optimize the production process.

3.1.1 Current-State map

The current-state simulation model of the factory production line is shown in Figure 2. All activities and their sequences in each station are generated and proper data distributions for the processing time of each activity are defined. In this model, the current-state of the production process is simulated based on the factory current-state VSM for 10 sample modules. The numbers of assigned labor personnel are fixed at each station and there is no cross-training through the production line. Modules vary slightly in size and specifications, entailing that processing times for modules are not uniform. The variation range is limited in order to identify defects in situations in which the source of deficiency is not obvious. The results of the simulation model comprise processing time for sample modules to be fabricated at each station, total processing time, idle time, and total man-hour requirements for the current-state production process. Variations in processing time at each station for different modules are plotted in output charts as well as total processing times for all the sample modules as presented in Figure 3. The duration variation of processing hours to complete each module at different stations and overall are presented in Figure 4.
The results of the simulation model demonstrate the variation in module completion duration at each station. When a larger module enters the production line, it is returned to the bottleneck of the production line, keeping upstream stations idle. Also, modules in the downstream stations are unable to move since the work on the previous module is not complete. As a result, the production capacity is decreased and the scheduled target based on customer demand cannot be reached.

### 3.1.2 Future-State map

The future-state of the production process is generated based on a set of proposed recommendations as follows:

- The number of labor personnel at each workstation is not fixed. Takt time is defined in such a way as to move the line at a steady pace and create continuous flow. Different scenarios are therefore run in order to find the near optimum result for the Takt time at which an optimum resource allocation plan is reached with the least fluctuation in labor requirements for different modules.

- The layouts of the off-line framing stations (floor framing, roof/ceiling framing, and wall framing) and cubing station are changed. The location of floor framing station is changed, and some activities are moved upstream including sheathing the floor and marking the walls layout. The proposed layout eliminates three waste activities from the process: lifting the floor with the overhead crane; undoing the air-jacks before lifting and replacing them afterwards; and carrying necessary tools and material to the next station to complete sheathing.

- The process was improved through utilization of more advanced tools and equipment. The required time to manually square the walls is eliminated by means of jigs at the wall framing table. The required time to cut wall components is reduced by half by means of a radial arm saw with measuring ability to cut several pieces to size at once. The required time to prepare a platform for roof/ceiling is reduced by 60% by means of an adjustable work platform to set up the roof for various layouts.

- The idle time associated with material delivery delay is eliminated through the use of just-in-time delivery along with Kanban cards for material replenishment and handling.

The future-state simulation model, as shown in Figure 5, determines resource allocation plan scenarios of the future-state production process. This model runs the simulation for a series of Takt times and calculates man-hour requirements for a number of sample modules at all the stations through the production line. Then, the best match for number of labor personnel at each station is defined, and, based on this the man-hour fluctuation caused by module variation at individual stations is measured. The model then calculates for each station and for the entire production line (1) total labor idle time due to earlier completion of a module; and (2) additional man-hour requirements due to late completion of a module. The total required time that is covered by idle labors defines the required number of labor personnel in the multi-skill worker crew that is cross-trained through the production line to increase production rate at stations which are behind the scheduled Takt time. The model is run to find the scenario with the minimum man-hours not covered by...
the multi-skill worker crew. Table 1 presents the required number of labor personnel and labor fluctuations at the floor station for sample modules for different scenarios, with Takt times ranging from 6 to 11 hours.

Figure 5 Future-state VSM simulation model

Table 1 Labor requirements and labor fluctuation at floor station for different scenarios

<table>
<thead>
<tr>
<th>Run</th>
<th>Labor personnel requirements</th>
<th>Labor fluctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>67</td>
<td>78</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 2 Scenario analysis for future-state production process

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Takt time (hour)</th>
<th>Total fixed labor personnel number</th>
<th>Cross-training hours at stations</th>
<th>Remaining hours of all shifts</th>
<th>Uncovered hours</th>
<th>Multi-skill labor personnel number</th>
<th>Total labor personnel number</th>
<th>Production rate (module/month)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>71</td>
<td>4</td>
<td>17</td>
<td>4</td>
<td>71</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>59</td>
<td>3</td>
<td>28</td>
<td>3</td>
<td>59</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>52</td>
<td>2</td>
<td>15</td>
<td>2</td>
<td>52</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>46</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>46</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>41</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>41</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>11</td>
<td>37</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>37</td>
<td>21</td>
<td>21</td>
</tr>
</tbody>
</table>

3.2 Post-Simulation Visualization of Production Process

In this phase, a post-simulation visualization model of the evaluated production scenarios is developed which facilitates production flow analysis and decision making for the management team. This model visualizes the results of the simulation model and provides the opportunity to compare near-optimum production scenarios, not only through charts and statistical data, but also by means of visualizing the ideal process prior to actual implementation of the proposed changes.

3.2.1 Animation and Process Evaluation

The proposed method integrates 3D visualization of a process with the statistical results of the simulation. The generated post-simulation visualization (PSV) model visualizes the simulation of the production process in detail, presenting process information such as modules’ fabrication schedule (both individually and in total) throughout the production line for evaluation purposes, whereas the proposed technique has the capacity to be linked to all statistical production information. For this purpose, a general visual...
A complex module production process which includes various constraints such as time, predecessor activity networks, space, continuous flow, labor assignments, and product variation, requires a flexible evaluation model to deal with changes in process and present the future-state in a simple way. For this purpose, another approach is pursued to create a more dynamic visualized model of the simulation results in Autodesk Navisworks Manage. The 3D model of the
factory is developed in Revit and then, along with the 3D models of modules generated in the design phase, is imported into Navisworks. Required information pertaining to modules currently in the production line and coming modules, including the processing times at all the stations, resource allocation plan, and work sequence, are imported from the simulation output for different scenarios. In order to analyze different plant layout configuration scenarios with automatic clash detection, more information is added to the model, including overhead crane capacity, factory space limitations, and labor safe work area.

In this model, a minute-by-minute schedule of the production process is generated and various layout configurations are analyzed for optimum scenarios. The developed PSV model capitalizes both on the advantages of the simulation as well as on the simplicity of the visualized minute-by-minute schedule. Figure 8 shows a snapshot from the model’s output for the case study that presents the process for scenario 3 with 8-hour Takt time. The model reveals a number of clashes due to space limitation for module transportation. In order to reach the production capacity, the number of produced modules is not changed, but the layout of the workstation arrangement changes until no clash is detected. Furthermore, an investigation is carried out considering various production process scenarios in order to compare the resource allocation plans and identify the best potential scenario for the given number of labor personnel at each station. The result of comparing the initial simulation model before generating PSV and the final version after running PSV indicates significant improvements in the production schedule and resource allocation.

Figure 8 PSV model of scenario 3 with 8-hour Takt time

Conclusion
With the capacity to predict and evaluate different scenarios for the production line, the introduction of simulation tools to resource planning provides several advantages to user in reducing risk of failure, accurate production forecast, and Takt time prediction. Since results are based on statistical information and specific expertise on simulation, PSV models provide a better understanding of simulation outcomes for managers and workers. However, the two-presented visualization models have different purposes and scopes. The first model presents in detail the planned manufacturing process after decisions being made from simulation result. It has, therefore, a more reviewing purpose in order to demonstrate how the process is carried out. The second model is used as a tool to visualize the effect of the activity network in the schedule and predict clashes that the simulation model is not able to detect as the conflict with models and the space limitation of facility.

References
Improving Productivity of Yard Trucks in Port Container Terminal Using Computer Simulation

Essmeil Ahmed a, Tarek Zayeda, and Sabah Alkassb

a Concordia University /Department of Building Civil and Environment Engineering, Montreal, Canada
b University of Sharjah/ College of Engineering, Sharjah, United Arab Emirates
e-mail: esm_662003@yahoo.com; tarek.zayed@concordia.ca; salkass@sharjah.ac.ae

Abstract

In port container terminals, reducing vessel turnaround time improves the terminal productivity and increases the capacity of world trades across the globe. This time reduction can be achieved by improving one or more container terminal major resources. Much research has been done on improving container terminal efficiency as a means to reduce vessel turnaround time; however, room still available for improvement. Minimizing the empty journeys of yard trucks could improve container terminal efficiency with reasonable time and cost investment. The objective of this research is to maximize container terminal productivity by minimizing vessel turnaround time within reasonable expenses. A handling container strategy is introduced by employing double cycling to reduce the empty travel trips of yard trucks. This double-cycling strategy still requires the use a single-cycle one before the trucks can be incorporated into double-cycle scheduling. The strategy is based on combining the efforts of two quay cranes (loading and unloading) to work as a unit. Simulation is used to test the efficiency of the proposed strategy. Simulation results indicate an improvement in the productivity rate combined with the unit cost savings using yard truck double cycling comparing to the standard single cycle operation.

Keyword:

Container terminal; Quay crane; Yard Truck; Double cycle; Simulation

1 Introduction

In container terminals, Minimizing vessel turnaround time (the time it takes for a vessel to be unloaded and loaded at its berth) accelerates the shipping time and reduces delays in delivering trade goods. This time reduction can be achieved by improving one or more container terminal resources. in [1], it is stated that the capacity of vessels will be increased by up to 800 TEUs (twenty-foot equivalent unit containers). Vessels today have been upgraded to carry more than 15000 TEUs as one way to minimize container shipment costs. Although, much research has been done on improving container terminal efficiency as a means to reduce vessel turnaround time, it is still space to do more improvements. Minimizing the empty journeys of yard trucks could improve container terminal efficiency with reasonable time and cost investment. The objective of this research is to maximize container terminal productivity by minimizing vessel turnaround time within reasonable hourly and unit costs. A handling container strategy is introduced by employing double cycling to reduce the empty travel of yard trucks. This double-cycling strategy still requires the use a single-cycle strategy before the trucks can be incorporated into double-cycle scheduling. The strategy is based on unite the efforts of two quay cranes to unload and load the yard truck during its cycle. The EZstrobe simulation system will be used to test the efficiency of the proposed strategy. The research is expected to be of value by improving container terminal productivity using existing facilities and resources. Gains will also be realized by reducing vessel turnaround time, which could save shipping costs and accelerate the global trade transported by sea. Simulation results indicates an improvement in the productivity rate combined with the
unit cost savings using YT double cycling comparing to the standard single cycle operation.

This paper is organized as follows. The section follows is a briefly back ground and literature review relating to the topic of this research. Followed is a description in details the proposed strategy of YT double cycling comparing to the tradition single cycling of handling containers at terminals. It also shows the how to calculate the vessel turnaround time in both strategies mentioned. The section after, defines the major factors affect the container terminal productivity and formulates the involved equipment’s cycle times. Then after, the section that proposes the simulation steps and models for both YT single and double cycling. Section before the last, proposes a case study to test and validate the strategies and thus the simulation models. In the last section, conclusions of this work have been drawn.

2 Literature Review

Vessels today have been upgraded to carry more than 15000 TEUs as one way to minimize container cost shipment. These large vessels are usually used to transfer containers through large container terminals to be transshipped by smaller vessels between medium or small terminals which called transshipment. Large transshipment container terminals now operate 24 hours a day with no stops, all year long, to meet the demand of the worldwide container trade. Not less than 100,000 containers are transferred weekly between berth side and temporary SYs [2].

Quay cranes (QCs), which are huge and costly machines, are used to unload and load container from and onto sea-going vessels. Vessel can be serviced by more than one QC on time to minimize the vessel turnaround time [3]. When using multiple QCs to service one vessel, the turnaround time is equal to the maximum time one of the QCs must spend to unload and load its assigned hold. The number of QCs also depends on a container terminal’s handling equipment availability and its internal road capacity. Horizontal transporters Yard trucks (YTs) are used to transport containers between the berth and a storage yard (SY). The more trucks that move within the same time, the more likely it is that traffic congestion could arise, which would cause delays in the truck cycle time. Gantry cranes are yard cranes (YCs) that unload and load containers from and onto vehicles at a SY. Rubber tired gantry crane (RTGC) is one of types of Gantry cranes. The RTGC has a width of 7 lanes, which are each equivalent to the forty-foot container width; 6 lanes are used to store the containers and the 7th lane is customisable for the yard trucks [4]. One conclusion is that, while YCs can stack containers above each other in stacks of up to seven, the optimal stacking height was determined to be 5 levels [5]. It is not only YT congestion that needs to be solved; YC clashing can occur when more than one crane work in the same lane. The research in [6] studied both YT and YC problems. Katta improved the layout design with buffering to eliminate YT congestion and YC clashing. Any delays in the availability of these resources directly leads to a proportional delay for the other resources, and ultimately on the container terminal productivity in general. On the other hand, improving any of these resources will improve container terminal productivity. Traditionally, vessels are unloaded and then loaded (single cycle) at transhipment container terminals. Recently, a new technique has been proposed by Goodchild in [7]. This technique has been developed in [8], [9] and [10] to optimize QC productivity and minimize vessel dwell time by minimizing the empty travel of QCs. Single cycling means that the imported containers from a vessel must be unloaded first, and then the exported containers can be loaded. In contrast, double cycling means to the loading and unloading of containers is carried out at the same time, in the unloading conditions. QC double cycling is “a technique that can be used to improve the efficiency of quay cranes by eliminating some empty crane moves” [7]. A scheduling problem is presumed by [7] one that can be solved by double cycling. The [10] extended what has been done in [7] research so that it would no longer be limited to just the stacks under a hatch, but would also work for above-hatch stacks. In order to reduce YT empty trips, [11] introduced a heuristic algorithm and test the algorithm by simulation for various scenarios of QCs types (single cycle, double cycling and combination of two QCs one loading and the other unloading) in different locations. One of their conclusion is the YT efficiency is affected by the QC operation type. The [12] supports the advantages of double cycling technique as a service method for improving container terminal productivity. They enhanced the conception that doubly cycling is a cost reduction method which does not require any improvement of existing infrastructure or introduce new technology.

Improving the productivity of existing container terminals without introducing new equipment and thereby expanding and/or developing the infrastructure of a facility is the primary objective of this research. This research is focussed on implementing the double cycling of YTs based on this QC double cycling technique, thereby minimizing empty YT journeys.
A new strategy of handling containers is to be proposed in this work to be able of combining the effort of two QCs to work as a unit. Because of its complexity, container terminal productivity is commonly tested by using simulation. The effectiveness of this proposed strategy will be verified via a simulation model.

3 Study Objectives

The overall objective of this research is to minimize vessel turnaround time and optimize container fleet size and hourly costs by implementing the yard truck double cycling technique to minimize empty truck journeys. Developing a simulation model is a part of this research and will test the technique. An optimization of the simulation outcome’s group solutions will be used as an input of the sensitivity analysis to optimize fleet size and the associated hourly costs. Several sub-objectives need to be achieved in order to satisfy the main objective:

- Identify and study the various factors that affect container terminal productivity;
- Build a productivity model to improve container terminal operation; and
- Compare the YT single and double cycling in terms of productivity rates and hourly costs.

4 Methodology

The research methodology consists of three sections. Each section is divided into its sub-phases according to the priority order. The first section addresses the understanding of container terminals, and starts with a comprehensive literature review organized and detailed as in the previous section. This section includes a state of the art review of yard crane scheduling, container transporting between storage yard and berth, temporary container storage yards, quay crane and allocation problems, quay crane double cycling and Yard truck double cycling. This second section focuses on introduce a container handling strategy and how does container handling method affect the vessel turnaround time. Moreover, it defines the factors that affect container terminal operation in terms of productivity and costs. The third section is the simulation section. It starts with the simulation modeling of both single and double cycling, followed by, case study and collection of the data needed to run the simulation, simulation implementation and model validation. The methodology ends with a conclusion and recommendation based on the results from the previous sections.

The methodology of this research is constructed on introducing a strategy of container handling, called the YT double cycling technique. This strategy depends on being able to combine the effort of two QCs to work as a unit with one crane discharging the vessel while the other loads it. Both QCs will serve the same truck in its cycle. Each truck will transport containers from the storage yard to the vessel and from the vessel to the storage yard in the same cycle. Just as with the QCs, two YCs will load and discharge the trucks at the storage yard.

QCs must be located more than two rows of forty feet container apart. In the interest of safety and to prevent conflicts, the QCs in this system will be three rows of forty feet apart. A YT single cycle loading is still needed to create space on vessel to be able start loading before incorporate to double cycle. At least, two YCs will be allocated on SY to load and discharge the truck. To be able to understand the proposed strategy, it is important to know the traditional strategy (YT single cycling).

4.1 The introduced strategy of YT double cycling

In general, vessel turnaround time starts with the unloading the imported containers until loading the last exported container. When using single cycling technique, loading the vessel should not start until the vessel is fully discharged. The vessel departs after the last exported container has been loaded. The total unloading and loading time is then counted as the vessel turnaround time. The bar chart on Figure 1(a) describes the operation. Then vessel turnaround time using single cycling ($T_s$) is:

$$T_s = w \times [\sum_{i=0}^{n} \sum_{j=1}^{m} Uc (i,j) + \sum_{k=0}^{g} \sum_{\beta=1}^{\gamma} Lc (\beta, \gamma)].$$

(1)
Where:

- \( w \) = average QC unloading/loading cycle time, \( U_c \) and \( L_c \) containers to be unloaded and loaded respectively,
- \( i \) = number of containers to be unloaded per row, where \( 1 \leq i \leq n \);
- \( j \) = number of rows on the vessel, where \( 1 \leq j \leq m \);
- \( \beta \) = number of containers to be loaded per row, \( 1 \leq \beta \leq s \); and
- \( \gamma \) = number of rows to be loaded, \( 1 \leq \gamma \leq g \).

As with single cycling, vessel turnaround time starts with the unloading of the first imported container and ends with the loading of the last exported container. However, in double cycling, loading exported containers can be started at a certain time, in parallel to the container unloading. This time has specific constraints and conditions which have been discussed earlier. When it is time to convert to double cycling, two QCs will work together as a unit to serve YT's with different activities (Loading and unloading).

The overlapping of some of the QCs’ cycle time reduces the vessel turnaround time to less than it is in single cycling. This time savings is the main justification for applying the double-cycling technique. A vessel still needs to be loaded with the last exported container before departure. Turnaround time is counted as the sum of the series of single cycle unloading, double cycling and single cycling loading of the imported and exported containers.

The vessel turnaround time using the YT double cycling (\( T_D \)) then:

\[
T_D = w \times \left( \sum_{i=1}^{n} \sum_{j=1}^{2} U_c (i, j) + \max\left[ \sum_{i=0}^{n} \sum_{j=1}^{m} U_c (i, j), \sum_{\beta=0}^{n} \sum_{\gamma=1}^{(g-3)} L_c (\beta, \gamma) \right] + \sum_{\beta=0}^{s} \sum_{\gamma=1}^{(g-1)} L_c (\beta, \gamma) \right)
\]  \( (2) \)

The operation time line of double cycling can be described as in the bar chart in Figure 1(b) where the unloading activity precedes the loading activity.

4.2 Factors that affect container terminal productivity

4.2.1 Quay crane cycle times:

The quay crane cycle time starts from the movement of the trolley (empty or loaded) from the truck lane to the (discharged or loaded) container position in the bay. The trolley makes different forward and backward moves. The trolley’s vertical speed is purposely different between its loaded and empty movements.

When discharging containers, the trolley starts empty, moving forward up vertically and horizontally at the same time (diagonally) in order to save time. Then, it moves horizontally to be close to the container location. The trolley again makes a diagonal movement, downward to be able to lift the container. If the container is above the hatch, lifting the container is the next step. If not, the trolley will move toward the container vertically down and then lift the container. After lifting the container, loaded backward moves will be applied, see Figure 2 for more details. The same steps of moving forward, but in the opposite direction, begin from where the bay ends to the truck lane. If the truck is available, the QC will load the container on the truck. To load the vessel, the trolley makes the same moves, only replacing the empty with loaded and the loaded with empty moves. In order to formulate the QC cycle time referring to the QC trolley motions, let \( w_u \) represent the QC unloaded cycle time and considering that all distances are in feet and times are in minutes, where:
wu = Quay crane cycle time = \[\sum\text{(forward times when empty)} + t_{\text{lift}} + \sum\text{(backward times when loaded)} + t_{\text{load}} + t_{l} \]

\[wu = \{\max\left[\frac{dp}{dv_1}, \frac{dx}{dv_1}\right] + \max\left[\frac{dh}{dv_2}, \frac{dr}{dv_1}\right] + t_{l}\} + \frac{dx}{dv_1} \max\left[\frac{dp}{dv_3}, \frac{dh}{dv_1}\right] + t_{l} + t_{\text{load}}\]  

Where: \(t_{\text{lift}}\) Represents the late time that the QC has to wait for the truck;

\(P, S, X, b, R, H\) and \(h\) are the vertical and horizontal distances the trolley makes when loading or discharging the vessel.

\(t_{\text{load}}\) and \(t_{\text{lift}}\) are the time to lift the container from the vessel and the time to load the container on the truck, respectively.

\(v_1, v_2\) and \(v_3\) are The QC trolley horizontal, vertical empty and vertical loaded speeds respectively.

4.2.2 YC cycle time

As with the quay crane, the yard crane cycle time starts from the movement of the trolley (empty or loaded) from the truck lane to the (discharged or loaded) container position on the pre assigned storage yard. The trolley makes the same set of forward and backward moves. YC trolleys also have different vertical speeds when they are loaded than it is empty. As with QCs, diagonal movements are applied to save time. As it is detailed, almost all of the QC’s forward and backward movement steps can be implemented by YCs (except those for the hatch as storage yards do not have hatches). The truck delay will lead to YC delay time. This delay is added to the cycle time and counted as a late time.

4.2.3 YT cycle time

YT single-cycle unloading starts with a YT moving from the truck pool or storage yard to the berth side. At the truck lane on the berth side, the truck will be loaded by the QC if it is ready. Otherwise, the truck waits for the QC to be ready. After being loaded, the YT returns to the storage yard to a pre-assigned lane, where an YC discharges the truck when it is available. A waiting time will be added if the YC is not ready. The YT will repeat the process until the last imported container is fully unloaded from the vessel. A specific number of YT's is needed to do the job in order to keep the cranes busy. Yard truck single cycle loading starts at the same place as YT single cycle unloading, in the storage yard. The only difference is that the YT has to be loaded with the exported container before it departs the storage yard.

\[Ts = w_1 + \left(\frac{dx_1}{dv_2}\right) + w_2 + \left(\frac{dx_2}{dv_2}\right) + \sum Ly + \sum Lb\]  

Where: \((Ts)\) Represents the yard truck unloading cycle time;

\(w_1, w_2\) are container unloading time by QC and YC respectively

\(x_1\) Empty and loaded trucks’ travel paths between the storage yard and bay side;

\(v_1\) Truck’s speed when loaded and \(v_2\) is the truck’s speed when empty;

\(Ly, Lb\) Truck waiting times at the yard zone; and bay side.

In double cycling, the first YC starts the cycle by loading the YT. The YT then moves, loaded with its container, to the berth side to be discharged by the first QC. After discharging, the YT moves empty to the second QC to be loaded. Next, it returns to the storage yard to unload the container at an imported lane. Figure 3 explains the YT double cycling procedure. The second YC should be ready to discharge the truck, which then departs empty to the exported lane to be loaded by the first YC, thus starting a new cycle. Any delay or waiting time for a crane will be added to the cycle time as late time. A fleet of YTs will continue the work until reverting back to single cycle loading to load the remaining containers.

For YT double cycle, the YT cycle time is:

\[TD = 2 \times (w_1 + w_2) + \left(\frac{dx_1 + dx_2}{dv_1}\right) + \left(\frac{dx_2 + dx_4}{dv_2}\right) + \sum Ly + \sum Lb\]  

Where: \(TD\) = yard truck unloading and loading (double) cycle time;
\( w1 \) = container loading or unloading time by QC;
\( w2 \) = container loading or unloading time by YC;
X1 and X3 are loaded trucks’ travel path from the storage yard to the bay and vice versa.
X2 = empty trucks’ travel path between loading and unloading QCs;
X4 = empty trucks’ travel path between the export and import lanes.

5 Simulation of Container Terminal Operation

The single and double cycling simulation models are designed in accordance with pre-defined steps. The EZstrobe simulation system has been used for modeling the container terminal operation due to its simplicity and power. To apply the EZstrobe simulation system, some steps must be followed as mentioned above. The data collection stage will be clarified in depth in the next section, followed by a case study to test and validate the simulation models. Single and double cycling steps are more detailed next.

5.1 Single cycle simulation model

A single cycle simulation model is designed according to the condition of discharging the vessel first then start loading after the unloading is completely done. The truck cycle will start moving empty from the storage yard or truck pool toward the berth side. At the same time, the QC starts its cycle by empty movement toward the target container to be unloading from the vessel. Once a truck arrives at the berth, the QC loads the container on the truck. The truck then moves loaded to the storage yard to be discharged by the YC, and then it travels back to the berth side (empty) to make another cycle. Meanwhile, the YC moves the container into the lane at the storage yard. The other trucks repeat this process until the last container is unloaded. Next, the loading process starts by loading containers on the truck at the export storage yard, to be transported to the berth, where the QC loads the containers on the vessel the opposite way of unloading. The process continues loading until the last exported container is to be loaded. At this time, the vessel turnaround time is done and the vessel free to depart.

5.2 Double cycle simulation model

The double cycle simulation model is designed as a form of integration between single and double cycling. This integration begins with unloading three or more rows before starting double cycling as a pre-condition, in order to minimize the fleet size and maximize crane use. Next, the unloading QC1 will change from unloading to loading the containers on the vessel. Another QC2 will be introduced to the fleet to continue unloading the containers from the fourth row to the end. QC1 starts loading the containers from the first to the last row. The trucks, loaded by YC1 at the export storage yard, arrive at the berth Side to be discharged by QC1. After being discharged, each truck will precede empty to QC2, to be loaded with a container unloaded from the vessel. QC1 simultaneously starts its cycle to load its container on the vessel. The loaded truck will move back to the (import) storage yard to be discharged by YC2, which should be ready for this discharge. After being discharged by YC2, the truck will proceed to the export SY to be loaded by YC1 and start a new cycle. YC2 starts its cycle as soon as it lifts a container from a truck. The YT, QC, and YC continue repeating their cycles until the last container has been fully unloaded. The fleet will then be reduced to one QC and one YC and fewer trucks to complete loading the vessel as a single cycle, as described earlier in the single-cycling simulation. is the designed YT double cycling simulation model of the procedure.

6 Data Collection and Case Study

6.1 Data collection

Since it is not yet possible to collect the data directly in this research, a technique to estimate the data needed to run the simulation is utilized. To employ this technique, the QCs, the YTs, and the YCs cycle times must first be calculated. It is assumed that the times will vary according to the speed variance. Any delay or acceleration of the cycle times will relate to the movement speed. For instance, a crane operator’s skills
and or weather changes should have an impact on the vertical and horizontal speeds of the. The YT cycle time is also calculated, according to the expected distances from the vessel to the storage yard. Finally, this data is analyzed by using probability distribution to determine the mean and standard deviation of each machine’s cycle time. Commercial software, the Easy fit distribution, has been used for the statistical analysis. A simulation of container terminal productivity operation will be done in the framework of this research. The simulation requires each activity to be run over its complete duration. It is understood that a huge amount of random data for all three types of equipment must be collected, and then statistically analyzed to get the mean time (μ) and standard deviation (σ) of each type of equipment’s cycle time. The result data was collected and analysed using an EasyFit distribution to draw the histogram and to calculate the mean and the standard deviation. Lifting and loading containers into or from vessel and YT was estimated using a constant of 0.166 minutes. Because it is a very small time compared to the other durations, QC moving between rows is neglected.

6.2 Cost estimation

As previously described, the hourly costs are estimated. These estimates are not exactly ‘real’ and are not confidential. To make this data useful and easier to manipulate, the percentage saving in each category of productivity rate, hourly cost and unit cost is utilized. The estimated hourly costs are to be followed in simulation inputs.

6.3 Case study description

The proposed case study considers a hatched vessel with a 16000 TEU (8000, 40 -t containers) capacity. The vessel will totally unload and be loaded with the same number of containers. The containers are estimated to be distributed uniformly on the vessel in 20 rows and 20 stacks. The number of stacks above the hatch is equal to the number of stacks below the hatch, with 10 levels of containers each. The total number of containers per row is 400. For single cycling, only one QC and one YC will do the job of unloading and loading the vessel. However, for double cycling, two QCs and two YCs are needed to do the job. Each activity (loading and unloading) requires one QC and one YC. The same trucks will work as duel loading/unloading to serve the QCs and the YCs. The small movement of QCs between the rows is neglected due its minor time value compared to the total time of unloading each row. The YCs are the RTG type. The trolley speeds of the QCs and the YCs have been quoted from the cranes’ manufacturer publications. The hourly costs are estimated, as real data from container terminals is not yet available.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Loaded containers TEUs</th>
<th>Unloaded containers TEUs</th>
<th>QC</th>
<th>YC</th>
<th>YT</th>
<th>QC</th>
<th>YC</th>
<th>YT</th>
<th>Overhead cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single cycling</td>
<td>16,000</td>
<td>16,000</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>150/QC</td>
<td>100/YC</td>
<td>50/YT</td>
<td>110</td>
</tr>
<tr>
<td>Double cycling</td>
<td>16,000</td>
<td>16,000</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>150/QC</td>
<td>100/YC</td>
<td>50/YT</td>
<td>110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Simulation output</th>
<th>Single cycle</th>
<th>Double cycle</th>
<th>Difference</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity rate (TEU/Hour)</td>
<td>63.73</td>
<td>107.6</td>
<td>43.87</td>
<td>68.83% (increases)</td>
</tr>
<tr>
<td>Unit cost ($/TEU)</td>
<td>9.57</td>
<td>7.99</td>
<td>1.58</td>
<td>16.50% (saving)</td>
</tr>
<tr>
<td>Total cost ($)</td>
<td>30,625</td>
<td>25,574</td>
<td>5,051</td>
<td>16.49% (saving)</td>
</tr>
<tr>
<td>Total hours</td>
<td>502</td>
<td>297.3</td>
<td>205</td>
<td>40.77% (saving)</td>
</tr>
</tbody>
</table>
Additional to the durations, other inputs are needed to run the simulation model. Simulation parameters and hourly costs are identified in Table 1. The result of the simulation reveals an improvement in terms of productivity, time and cost. This improvement can be seen in Table 2. In summary, it is concluded that double cycling can improve CT productivity which leads to minimize vessel turnaround time with reasonable cost saving.

7 Conclusions

Container terminal customers (shipping companies) believe that “Vessels do not make money while berthing”, which means that minimizing vessel turnaround time is crucial to satisfy these customers. It is clear that improving the productivity of any container terminal’s resources leads to the improvement of the other elements’ productivity and of terminal productivity as a whole. QC double cycling has been introduced recently to improve container terminal productivity and minimize vessel turnaround time. This work introduces a new strategy that implements double cycling on YTIs to improve container terminal productivity. This new strategy of handling containers has been modeled, tested and verified. The simulation indicates a reasonable improvement in maximizing productivity and minimizing hourly and unit costs. The simulation models reveal a productivity improvement of about 68% in terms of (TEU/hr) or about 34% in terms of (TEU/hr/QC) and cost savings of about 16% in both unit cost and cost per vessel of 16000 TEUs capacity.

8 References:


BIM-based Integrated Framework for Detailed Cost Estimation and Schedule Planning of Construction Projects

Hexu Liu*, Ming Lu* and Mohamed Al-Hussein*

* Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada T6G 2W2
E-mail: hexu@ualberta.ca, mlu6@ualberta.ca, malhussein@ualberta.ca
(*Corresponding author: mlu6@ualberta.ca)

Abstract -

With the increasing popularity of Building Information Modeling (BIM), several research studies have focused on its application in the construction industry, and advancements such as developments of BIM-based tools for construction scheduling and cost estimating have been carried out. Nevertheless, these tools are each developed only for one particular aspect of construction management, such as scheduling or cost estimating. As such, most are still limited to the product element level, rather than the construction operation level. This paper thus presents a BIM-based integrated framework for detailed cost estimation and schedule planning of construction projects. In the proposed framework, a BIM product model developed in Autodesk Revit, is integrated with a construction process model retrieved from RS Means, with the objective of generating detailed cost estimation and a workface construction schedule simultaneously. In this paper, the literature pertaining to BIM, cost estimating and scheduling is surveyed; then, the proposed framework is presented, and a simple building project is used as a case study to facilitate understanding of and verify the applicability of the integrated framework. Finally, findings from the implementation of the proposed framework are summarized.

Keywords -

Building Information Modeling; Integrated Framework; Detailed Cost Estimation; Construction Scheduling.

1 Introduction

Detailed cost estimating is generally referred to as the process of predicting the project cost at the workface level based on detailed design drawings/documents and specified construction methods/specifications. During this process, defining the scope of the construction project and establishing its work breakdown structure (WBS) are the first steps. Then, the quantity takeoff is surveyed according to the detailed design drawings or building information modeling (BIM) model. The following task is to find applicable unit cost items from a unit cost database in accordance with the specified construction method. Furthermore, direct cost is obtained by multiplying quantities with unit cost; other costs, such as indirect costs, are calculated by taking a percentage of the direct cost. Finally, project cost is calculated by summing up direct cost with the other costs mentioned above. Similarly, detailed schedule planning is the processes - including identification of WBS for construction schedule, quantity survey, assessment of productivity, calculation of activity duration, and determination of construction sequencing logic and project duration—by which to plan the construction schedule at the operation level. Due to the fact that both detailed cost estimation and schedule are at the workface level, they pertain not only to construction planning in the planning phase, but also can be utilized to monitor and control construction progress during the project execution phase. Hence, detailed cost estimating and schedule planning are two critical tasks of construction management which directly contribute to the success of construction projects.

Currently, cost estimation and construction scheduling usually are performed separately by construction practitioners, due to the lack of a well-developed integrated framework or system for both cost estimating and project scheduling. In fact, given that cost estimation and schedule planning share some common processes, such as quantity takeoff, it is possible to combine these two processes to develop a single integrated information framework for detailed cost estimation and project scheduling. Such a framework reduces workload at the planning phase and also benefits construction management during the later execution phase. In addition, BIM, which can be defined as a digital representation of a facility in which all facility information is represented and visualized in the model, is gaining momentum within the construction industry. BIM contains enriched information pertinent
to the facility, and facilitates the exchange and interoperability of information, making it capable of supporting many kinds of analysis, including cost and schedule analysis. Consequently, the development of BIM technology provides solid support for the integrated framework of cost estimation and schedule planning.

Although many BIM-based tools for cost estimation and schedule planning have been developed, they are each developed for only one aspect of construction management, such as scheduling or cost estimation. Moreover, most of these tools are limited in their application to determining quantities of product elements (e.g., number of doors, windows, and walls) from the BIM model as inputs. Detailed construction information (e.g., construction method, crew information), which would be necessary in order to generate detailed estimation and schedule at the construction operation level, is not taken into consideration. This paper thus proposes a BIM-based integrated framework for detailed cost estimation and schedule planning of construction projects. In the proposed framework, a BIM product model in Autodesk Revit is integrated with a construction process model retrieved from RS Means so to combine construction method knowledge with 3D product model, thereby facilitating cost estimation and schedule planning.

The paper begins with a review of state-of-the-art research with respect to detailed cost estimation, project scheduling, and BIM technology. Second, detailed explanations pertaining to the proposed framework are presented, and a simple building project is used as a test-bed to verify the integrated framework. Finally, findings from the implementation of the proposed framework are summarized, particularly as they pertain to potential future research.

2 Literature Review

2.1 Overview

Cost estimation and project scheduling have been of interest to scholars for decades, and numerous efforts within these domains have been conducted in recent decades. Conventionally, previous studies in these areas have concentrated on purposes such as schedule optimization, cost estimation models, and time-cost trade off, and have aimed to produce more accurate and optimized project schedule and estimated cost. However, with the recent emergence of BIM, BIM-based applications in the construction industry have been another focus for many researchers. Several BIM-based applications and tools for cost estimation and scheduling planning have been developed to improve the efficiency of construction practitioners. Given the objective of this research, this literature survey focuses on BIM-based quantity takeoff, BIM-based cost estimation, BIM-based project scheduling, and cost and schedule integration.

2.2 BIM-based Quantity Takeoff

Quantity takeoff is the foundation of other tasks in construction management, such as cost estimation and schedule planning, and its accuracy can directly affect downstream analyses and decisions. Traditionally, quantity takeoff is a manual process during which quantities of design elements are measured based on the design drawings or the 3D model, and this manual quantification is highly error-prone [12]. Thus, various automated approaches to extracting quantity information from 2D drawings or 3D models have been explored in the past, such as generating quantities using AutoCAD drawings [14]. Among these, BIM offers the best automatic approach by which to generate accurate quantity takeoff directly from 3D product models [13]. Also, BIM-based quantity takeoff is currently the most widely used BIM-based application in the construction industry. Most BIM tools are able to support the quantity takeoff feature, including the “Schedules” function of Autodesk Revit. Nevertheless, BIM-based quantity takeoff may not provide all the necessary data about the product model in the case in which the BIM model is not designed with sufficient detail. To facilitate automatic quantity takeoff to a sufficient level of detail, the BIM model must be “redesigned”, which requires even more effort than performing manual takeoff. As such, some studies have sought to explore an automatic approach to designing the BIM model in performing a quantity takeoff. Monteiro et al. [12] developed an add-on for ArchiCAD, which can automatically generate the formwork model based on the structural model of the building. Kim et al. [6] explored an automated modeling method by which to model a building’s interior. Once the detailed information can be represented in the BIM, the thorough quantity takeoff can be generated by the routine in the BIM tools. All these efforts pertaining to automatic modeling could, on the other hand, improve the efficiency of quantity takeoff.

2.3 BIM-based Cost Estimation

As previously described, quantity takeoff is part of the cost estimation process. Most research regarding BIM-based quantity takeoff has actually been conducted to serve the purpose of cost estimating. Besides quantity takeoff, however, another challenge entailing a considerable amount of manual work in estimating is to find and apply appropriate cost data to the takeoff.
Without exploring an automatic solution to this challenge, Ma et al. [11] introduced a semi-automatic way to conduct cost estimation for tendering of building projects based on the use of a design model through an open Industry Foundation Classes (IFC) standard. In fact, the use of ontology technology could help reduce or eliminate manual involvement and automate the process of searching for the most appropriate cost items. By formalizing ontology to represent the cost-driving features of building product models, practitioners can generate cost estimation more expeditiously [14, 15]. In this context, Lee et al. [8] have illustrated the ontology and BIM-based approach for building cost estimation, but with the limitation of only focusing on tiling work.

2.4 BIM-based Schedule Planning

In terms of scheduling, research efforts in the last decade have evolved from traditional 3D CAD model-supported construction planning [16] to BIM models with enriched information-based scheduling [9]. However, most of these have only been able to produce construction schedules at the product component level, not at the construction operation level. Today, with the discrete event simulation-based schedule approach being increasingly utilized to support the generation of construction schedule at the on-site operation level, researchers are exploring potential means by which to integrate this approach with BIM. One example is an interface system developed by Wang et al. [17] to generate the on-site operation level schedule. However, their research was limited to reinforced concrete (RC) construction projects and did not have flexibility in considering different construction methods.

2.5 Cost and Schedule Integration

To date, various models and systems through which to integrate cost and schedule have been developed, including an integrated cost and schedule model for repetitive construction processes [2], a schedule and cost management system for steel structural construction [7], and an integrated database framework for cost, schedule, and performance data [1]. Nevertheless, these efforts have sought mainly to manage and control the construction process during the project execution stage, whereas detailed cost estimation and construction schedule planning, as two main concerns for construction project management [5], demand a large amount of effort during the planning phase. In this regard, an integrated system with the support of BIM for detailed cost estimation and schedule planning would improve project planning efficiency during the planning phase and benefit all the stakeholders involved.

3 Integrated Framework

The proposed integrated framework for detailed cost estimation and schedule planning is presented in Figure 1. It mainly consists of three components, a construction-oriented 3D product model, an integrated work breakdown structure, and a construction process information database. The construction-oriented 3D
product model, developed in Autodesk Revit, is a building product information database from which quantity takeoff is extracted for cost estimation and schedule planning, whereas the construction process information database from RSMeans stores the construction process/work package information, such as the crew rate and productivity information of each work package. Both quantity takeoff and work package information are organized based on an integrated work breakdown structure (WBS) which, as the core of this framework, combines cost breakdown structure and WBS for construction scheduling in consideration of specific construction methods (e.g., requirements of temporary structures and crews), thereby facilitating the generation of detailed cost estimation and construction schedule. In the following sections, detailed explanations are given.

3.1 Integration of Cost and Schedule

Integration models for cost and schedule can be categorized into three groups, including a WBS-based model, faceted classification model, and work-packaging model [1]. In the present research, a new WBS-based model drawing on the advantages of both the WBS-based model proposed by Eldin [3] and a work-packaging model, is proposed to achieve the integration of detailed cost estimation and construction scheduling. More specifically, the new WBS-based model constitutes a single, integrated work breakdown structure linking and combining cost breakdown structure (CBS) items with WBS items. Additionally, CBS items in the proposed WBS should be designed corresponding to cost records customized for specific construction methods in the construction process information database, such as RSMeans. Moreover, the integrated WBS is established following a consistent format based on all construction activities involved in constructing the final product; resources (cost) or time or both are incurred in the activities.

3.1.1 Integrated work breakdown structure

As indicated in the example provided in Figure 1, the integrated WBS is organized into three levels. The first level is designed for types of building component, with the second level for specific individual components. By setting the second level, the proposed WBS addresses one main limitation of traditional CBS with no support of BIM – that it does not involve specific (e.g., spatial) information of individual building components. For instance, “02 Columns” in Figure 1 is further separated into “02.01 Column #1” and “02.02 Column #2”, etc., in order to consider work packages for individual components.

With detailed construction data available, the third level is divided into activities. As shown in the figure, “02.02 Columns #2” is divided into “Concrete form”, “Reinforcing bar”, “Concrete material”, “Concrete placing”, and “Concrete curing” at the third level. Among these items, “Concrete form”, “Reinforcing bar”, and “Concrete placing” will incur construction cost and time, and be taken into consideration for cost estimation and schedule planning. However, “Concrete material” is the only item considered for the purpose of estimating the material cost of concrete, and “Concrete curing” only incurs time during construction. As a result, the integration of cost and schedule is achieved at this level by combining cost and schedule activities.

3.2 Integration of Product and Process Models

In fact, the three-level hierarchical WBS not only achieves the integration of cost and schedule, but also integrates the product model with the construction process model. The reason lies in the fact that the product model is decomposed into individual building components which are listed at the second level in the integrated WBS, whereas the third level WBS consists of detailed construction processes constructing the corresponding individual component. Thus, by linking the third level to the second level of WBS, we essentially complete the match from the construction process model to the product model.

Additionally, the quantity takeoff from Revit and the construction process information from RSMeans should be generated or selected to match the lower-level (third level) WBS. In fact, data from RSMeans is construction-oriented and is formulated based on detailed construction process information (e.g., construction method). Aside from this, RSMeans utilizes a MasterFormat-based 12-digital-index system, known as the most adopted industry standard for organizing construction data, to organize cost items. Thus, RSMeans data can be easily mapped to detailed WBS manually. However, the product model in Revit is an assembly of building components and does not embrace any construction information used to generate the takeoff for construction activities. Therefore, we have to integrate Revit and RSMeans systematically to address this issue.

Fortunately, Autodesk Revit provides several powerful mechanisms by which to integrate data into Revit families (product model components), including “project parameters”. In general, project parameters are customized properties that can be added to elements by end-users and are able to hold the user-defined information. Consequently, the integration of Revit and
RSMeans can be realized by defining construction process information as project parameters for Revit elements. Moreover, to facilitate the implementation of the integration, one Revit add-on, as shown in Figure 2, is developed. The add-on automatically identifies and lists all types of building elements in the project model, as well as construction information items in the Microsoft (MS) Excel file (exported from RSMeans Online). Practitioners can then add the selected construction items as type parameters into the building elements belonging to the checked element type by clicking “Attach Parameters”, and can also define names for the parameters through the middle Textbox in the interface. Figure 2 shows one example of mapping construction information to building element for a concrete masonry unit (CMU) wall. The resulting properties of the CMU wall are presented in Figure 3. As indicated in the figure, five construction items from RS Means are stored under the “Construction” category in Type Parameter of CMU Wall.

With regard to quantity takeoff for each work package under each building element, the schedule feature of Revit can be utilized to perform the task. However, it should be noted that Schedule in Revit not only can produce the physical volume or area of a building element as quantity takeoff, but also is capable of allowing users to embed used-defined equations for the derived quantity, which makes Revit able to produce construction activity-oriented quantity takeoff. For instance, the quantity for grouting in CMU walls is defined as one quarter of the area of CMU walls in RSMeans, provided that only one quarter of CMU cores are fully grouted. In this case, the quantity of grouting for CMU walls can be produced by embedding “Area*0.25” in Revit Schedule. Figure 4 presents one example of quantity takeoff for CMU walls. Subsequently, the takeoff can be easily combined with unit cost and productivity data in order to perform detailed cost estimation and schedule planning.

4 Case Study

In this section, the procedures to implement the integrated framework are illustrated by using a simple building project as a case study. The project is a one-storey reinforced unit masonry building designed for use as a parking garage, with the dimensions of 60 ft. (width) x 90 ft. (length) x 12 ft. (height). The exterior load bearing walls are made of 12-in.-thick, 5000 psi CMU blocks, vertically reinforced by #6 bars in every second block, with reinforcing cores fully grouted. To enhance the structural integrity, all rebar extends 3 ft. above the top of the wall and is bent into the roof slab. The continuous wall footing is 24 in. (width) x 18 in. (depth) with 2 #6 bars, continuous, and #4 bars, 18 in. on-center, across the footing. The footings are formed on the sides, using fine-graded earth as the soffit. All concrete in the structure is 5000 psi, and is placed by crane and bucket. The slab-on-grade (SOG) is 6 in.-thick, with 6 x 6 6/6 welded wire fabric reinforcement. The subgrade consists of 6 in. of gravel, 3 in. of sand, with a 0.010 in.-thick polyethylene vapor barrier. Three interior reinforced concrete columns are placed down the center line of the building. All columns are 16 in. x 16 in. and are constructed using 1-use plywood forms and 4 #8 vertical reinforcing bars. Column footings are

Figure 2. Interface by which to map construction processes to product elements

Figure 3. User-defined properties of CMU wall

Figure 4. One example of quantity takeoff
6 ft. square and are constructed using 1-use plywood forms, 15 in. deep, with 6 #6 bars each way in the bottom of the footing. Footings are formed on all sides, with fine-graded earth as the soffit. The roof slab is an 8 in.-deep flat slab constructed using 2-use plywood forms, with #8 rebar 24 in. on-center each way. Rebar is placed in the bottom of the roof slab. #4 bars, 12 in. on-center each way, are located in the top of the slab in an 8 ft. square area over each column. 1-use wood curb forms are installed for the construction of the slab-on-grade and the roof slab.

4.1 Development of 3D BIM Model

According to the design description, the BIM model is developed in Autodesk Revit in detail. All the product information is modeled in Revit, including various types of rebar, in order to facilitate the generation of the comprehensive quantity takeoff. In order to design rebar in detail and build the 3D model for rebar, one Revit extension provided by Autodesk, “Reinforcement”, is employed. The extension is developed based on an object parametric modeling approach. It requires only a few parameters from users in order to generate the 3D model of rebar, and improves the efficiency of modeling. Figure 5 shows the 3D view of the building project.

4.2 Construction Information Database

Construction information, including cost or productivity for each work package, is extracted from RSMeans Online. During this process, construction practitioners must be involved in selecting appropriate items corresponding to the construction methods specified in the design description and to match the RSMeans items with the integrated WBS. Taking the formwork for concrete wall footings as an example, RSMeans provides a variety of items for formwork which are differentiated by the material of the formwork, number of uses, and type of building component, etc. However, only one item can be considered the most appropriate item for the designed wall footing, based on the conditions of “continuous wall footing”, “plywood” and “1-use”. In this regard, practitioners eventually should select item with a line number “031113450020”. Similarly, other items for the project can be selected in the same manner. Once all cost items are identified from RSMeans, they are exported into MS Excel for the subsequent estimation and scheduling. One aspect to be pointed out at this juncture is that the bare cost, rather than overhead and profit (O&P) cost, is used in the estimation; the indirect cost, meanwhile, including overhead, is estimated by taking a percentage of direct cost. The cost listed in Table 1 is the bare cost for labor, material, equipment.

4.3 WBS and Quantity Takeoff

The WBS for the building project is presented in Table 1. During the establishment of WBS, it is assumed that: (1) excavation and earth moving work will be performed by subcontractors, and not included in the WBS; (2) demobilization is also ignored in the WBS. It should also be noted that in the interest of brevity only two wall and two wall footings, and one column and one column footing are presented in Table 1, since the same quantities and cost and productivity information will apply for the remaining wall footings, columns, and column footings.

Then, following the steps illustrated in the previous section, the RSMeans construction data is assigned to each type of building component in Revit. The quantity takeoff is generated and exported into MS Excel through Revit’s Schedule feature. Table 1 tabulates the generated quantity takeoff. As described above, Schedule in Revit allows users to define equations in order to produce derived quantities, and this feature is especially instrumental in performing quantity takeoff for temporary facilities which are not represented in the 3D product model. In the case study, quantities for formwork and scaffolding are surveyed as such. For instance, the formwork quantity for each continuous wall footing is expressed as Equation (1); and scaffolding quantities are estimated separately by Equation (2) for steel tubular and by Equation (3) for planks. All the equations are inputted into Revit when the quantity schedules are created.

\[
Q_i^F = \frac{V_i}{W_i} \times 2
\]

\[
Q_i^{ST} = \Lambda_i
\]

\[
Q_i^{SP} = \frac{L_i^W}{L_i^F} \times 2
\]

Where \(i\) is the identification number of the building element; \(Q_i^F\) is the formwork quantity for wall footing; \(V\) denotes the volume of the wall footing in Revit; \(W_i\)
<table>
<thead>
<tr>
<th>1st WBS</th>
<th>2nd WBS</th>
<th>3rd WBS</th>
<th>Quantity</th>
<th>Line Number</th>
<th>Material</th>
<th>Labor</th>
<th>Equipment</th>
<th>Daily output</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>0101</td>
<td>010101</td>
<td>567</td>
<td>031221160100</td>
<td>0.00</td>
<td>0.68</td>
<td>0.73</td>
<td>1040</td>
<td>0.55d</td>
</tr>
<tr>
<td>SW</td>
<td>SW#1</td>
<td>010102</td>
<td>138.7C</td>
<td>031232171200</td>
<td>50.24</td>
<td>9.45</td>
<td>0.82</td>
<td>200</td>
<td>0.69d</td>
</tr>
<tr>
<td>02</td>
<td>0201</td>
<td>020101</td>
<td>270</td>
<td>031113450020</td>
<td>8.10</td>
<td>3.60</td>
<td>0.00</td>
<td>375</td>
<td>0.72d</td>
</tr>
<tr>
<td>WF</td>
<td>WF1</td>
<td>020102</td>
<td>0.17Ton</td>
<td>032110600500</td>
<td>1448</td>
<td>573.04</td>
<td>0.00</td>
<td>2.1</td>
<td>0.08d</td>
</tr>
<tr>
<td>#3</td>
<td>#3</td>
<td>020103</td>
<td>10.05 CY</td>
<td>033105702200</td>
<td>0.00</td>
<td>27.77</td>
<td>11.79</td>
<td>110</td>
<td>0.09d</td>
</tr>
<tr>
<td>02</td>
<td>0201</td>
<td>020104</td>
<td>10.03 CY</td>
<td>033105540040</td>
<td>219.89</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>02</td>
<td>0201</td>
<td>020105</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>02</td>
<td>0201</td>
<td>020201</td>
<td>174SFCA</td>
<td>031113450020</td>
<td>8.10</td>
<td>3.60</td>
<td>0.00</td>
<td>375</td>
<td>0.46d</td>
</tr>
<tr>
<td>WF</td>
<td>#2</td>
<td>020202</td>
<td>0.11 Ton</td>
<td>032110600500</td>
<td>1448</td>
<td>573.04</td>
<td>0.00</td>
<td>2.1</td>
<td>0.05d</td>
</tr>
<tr>
<td>02</td>
<td>0202</td>
<td>020203</td>
<td>6.47 CY</td>
<td>033105702200</td>
<td>0.00</td>
<td>27.77</td>
<td>11.79</td>
<td>110</td>
<td>0.06d</td>
</tr>
<tr>
<td>#4</td>
<td>#4</td>
<td>020204</td>
<td>6.47 CY</td>
<td>033105540040</td>
<td>219.89</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>02</td>
<td>0202</td>
<td>020205</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>03</td>
<td>0301</td>
<td>030101</td>
<td>90 SFCA</td>
<td>031113450000</td>
<td>2.61</td>
<td>4.42</td>
<td>0.00</td>
<td>305</td>
<td>0.3d</td>
</tr>
<tr>
<td>SOG</td>
<td>SOG</td>
<td>040102</td>
<td>51 CSF</td>
<td>032205500500</td>
<td>32.58</td>
<td>20.74</td>
<td>0.00</td>
<td>29</td>
<td>0.18d</td>
</tr>
<tr>
<td>04</td>
<td>0401</td>
<td>040101</td>
<td>146FCA</td>
<td>031113650000</td>
<td>3.32</td>
<td>6.29</td>
<td>0.00</td>
<td>215</td>
<td>0.68d</td>
</tr>
<tr>
<td>05</td>
<td>0501</td>
<td>050101</td>
<td>192SFCA</td>
<td>031113250000</td>
<td>3.49</td>
<td>7.28</td>
<td>0.00</td>
<td>185</td>
<td>1.04d</td>
</tr>
<tr>
<td>C</td>
<td>C#1</td>
<td>050102</td>
<td>0.16 Ton</td>
<td>032110600500</td>
<td>1448</td>
<td>573.04</td>
<td>0.00</td>
<td>2.1</td>
<td>0.08d</td>
</tr>
<tr>
<td>05</td>
<td>0501</td>
<td>050103</td>
<td>5 CY</td>
<td>033105702500</td>
<td>0.00</td>
<td>67.52</td>
<td>29.08</td>
<td>45</td>
<td>0.11d</td>
</tr>
<tr>
<td>#2</td>
<td>#2</td>
<td>050104</td>
<td>5 CY</td>
<td>033105350400</td>
<td>219.89</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>05</td>
<td>0501</td>
<td>050105</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>06</td>
<td>0601</td>
<td>060101</td>
<td>95 CY</td>
<td>033105704300</td>
<td>0.00</td>
<td>17.59</td>
<td>0.65</td>
<td>55</td>
<td>1.73d</td>
</tr>
<tr>
<td>R</td>
<td>R#1</td>
<td>060102</td>
<td>95 CY</td>
<td>033105350400</td>
<td>219.89</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>06</td>
<td>0601</td>
<td>060103</td>
<td>95 CY</td>
<td>033105704300</td>
<td>0.00</td>
<td>17.59</td>
<td>0.65</td>
<td>55</td>
<td>1.73d</td>
</tr>
<tr>
<td>07</td>
<td>0701</td>
<td>070101</td>
<td>192SFCA</td>
<td>031113250000</td>
<td>3.49</td>
<td>7.28</td>
<td>0.00</td>
<td>185</td>
<td>1.04d</td>
</tr>
<tr>
<td>MW</td>
<td>MW#1</td>
<td>070102</td>
<td>195SFCA</td>
<td>031113356500</td>
<td>1.80</td>
<td>7.48</td>
<td>0.00</td>
<td>180</td>
<td>1.08d</td>
</tr>
<tr>
<td>07</td>
<td>0701</td>
<td>070103</td>
<td>7.44 Ton</td>
<td>032110600400</td>
<td>1448</td>
<td>414.70</td>
<td>0.00</td>
<td>2.2</td>
<td>3.38d</td>
</tr>
<tr>
<td>#3</td>
<td>#3</td>
<td>070104</td>
<td>126 CY</td>
<td>033105701550</td>
<td>0.00</td>
<td>27.77</td>
<td>11.79</td>
<td>110</td>
<td>1.15d</td>
</tr>
<tr>
<td>07</td>
<td>0701</td>
<td>070105</td>
<td>126 CY</td>
<td>033105350400</td>
<td>219.89</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>07</td>
<td>0701</td>
<td>070106</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080101</td>
<td>1092 SF</td>
<td>042210280650</td>
<td>8.85</td>
<td>6.00</td>
<td>0.00</td>
<td>300</td>
<td>3.60d</td>
</tr>
<tr>
<td>V#1</td>
<td>V#1</td>
<td>080102</td>
<td>249SF</td>
<td>040516300350</td>
<td>3.32</td>
<td>1.95</td>
<td>0.22</td>
<td>640</td>
<td>0.42d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080103</td>
<td>767.2 lb</td>
<td>040519260060</td>
<td>0.90</td>
<td>0.53</td>
<td>0.00</td>
<td>650</td>
<td>1.18d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080104</td>
<td>10.95 SF</td>
<td>015423700990</td>
<td>0.00</td>
<td>131.63</td>
<td>0.00</td>
<td>8</td>
<td>1.37d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080105</td>
<td>11.55 Ea.</td>
<td>015423705700</td>
<td>0.00</td>
<td>14.58</td>
<td>0.00</td>
<td>72</td>
<td>0.10d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080106</td>
<td>7.05 CSF</td>
<td>015423700990</td>
<td>0.00</td>
<td>131.63</td>
<td>0.00</td>
<td>8</td>
<td>0.88d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080107</td>
<td>7.45 EA</td>
<td>015423705700</td>
<td>0.00</td>
<td>14.58</td>
<td>0.00</td>
<td>72</td>
<td>0.10d</td>
</tr>
<tr>
<td>08V</td>
<td>0801</td>
<td>080108</td>
<td>50.14</td>
<td>072610101200</td>
<td>9.66</td>
<td>9.27</td>
<td>0.00</td>
<td>37</td>
<td>1.36d</td>
</tr>
</tbody>
</table>

Note: SW: site work; WF: wall footing; CF: column footing; SOG: slab on grade; C: column; R: roof; MW: masonry wall; V: vapor. Under “Material”, “Labor”, and “Equipment” is bare unit cost ($/unit); d in Duration is day.
Table 2. Cost estimation of the building project

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount ($)</th>
<th>Total($)</th>
<th>Rate</th>
<th>Cost Basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>70,297</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td>137,652</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td>2,861</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>210,810</td>
<td>210,810</td>
<td>7%</td>
<td>Current Total</td>
</tr>
<tr>
<td>Engineering Fees</td>
<td>14,757</td>
<td></td>
<td>5%</td>
<td>Material</td>
</tr>
<tr>
<td>GST</td>
<td>6,883</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permits</td>
<td>7,000</td>
<td></td>
<td>27%</td>
<td>Labor</td>
</tr>
<tr>
<td>Payroll Burden</td>
<td>18,980</td>
<td>258,429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td>517</td>
<td>305,556</td>
<td>18%</td>
<td>Current Total</td>
</tr>
<tr>
<td>Corporate Overheads and Profit</td>
<td>46,610</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surety Bonds</td>
<td>10,399</td>
<td></td>
<td></td>
<td>Bond</td>
</tr>
<tr>
<td><strong>Base Estimate</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>$315,955.21</strong></td>
</tr>
</tbody>
</table>

represents the width of the wall footing; \( Q_{ST} \) denotes quantities for steel tubular scaffolding; \( A \) represents the vertical area of the wall; \( Q_{SP} \) is the quantity of scaffolding planks; \( L_w \) is the length of wall; and \( L_P \) is the length of scaffolding planks.

4.4 Cost and Schedule Determination

Direct costs in the estimation are grouped into three categories - namely, labor, material, equipment, and are calculated separately, following Equation (4).

\[
\text{Cost}_i = \sum_{j=1}^{n} R_{ij} \times Q_{ij}
\] (4)

Where \( i \) denotes the category of cost, (labor, material, or equipment); \( j \) represents the index of cost item; \( n \) is the number of cost items; \( R \) is the bare rate from RSMeans; and \( Q \) denotes the quantity from the BIM model. Other costs, such as indirect cost, are calculated by taking a percentage of the basis. The assumptions pertaining to various types of indirect costs considered in the project and the cost results are presented in Table 2.

In terms of schedule, the duration for each activity is obtained by dividing its quantity by its daily output as shown in Table 1, and the construction sequencing logic of the schedule is determined manually based on authors’ experience. For instance, “020202 rebar for wall footing” must be finished before “020203 concrete placing for wall footing” can be started. Finally, the project duration is determined as 53 calendar days (or 35 working days.)

5 Summary

In the presented research, a BIM-based integrated framework for detailed cost estimation and scheduling planning has been proposed, and has been explained and verified through a case study of building project. As the core of the proposed framework, the integrated WBS achieves cost and schedule integration as well as the integration between the product model and the construction process model. However, it is developed manually based on construction knowledge of a building project. Future research efforts are needed with respect to the automatic generation of WBS. Some other challenges regarding full automation of BIM-based detailed cost estimation and schedule planning are summarized as follows:

1. Temporary facilities such as formwork and scaffolding are missing from the 3D model. Consequently, quantities for formwork and scaffolding cannot be directly extracted from the BIM model without manual involvements.
2. Modeling of the temporary facilities in a BIM environment is difficult to achieve without the required construction knowledge. For instance, in the case study, CMU walls can be constructed concurrently by establishing multiple sets of scaffolds around the building; or they can be built sequentially by installing and moving only one set of scaffolds. In this case, the BIM program needs to have “intelligence” embedded in order to automatically build a 3D model of the temporary facility.
3. Incompatibility between the BIM-based quantity takeoff and downstream analyses, such as estimation, with respect to the definition of quantities also impedes the integration of the product and process models. For example, RSMeans defines the grout quantity for CMU as the area of the grouted wall, rather than the physical volume, area or perimeter of grouted core provided by Revit.

Acknowledgement:
Dr Gunnar Lucko, Catholic U. of America is acknowledged for specifying the “garage” project, which was used in this research and by students in the CIV E 406 (“Construction Estimating, Planning and Control”) Fall 2013 class at University of Alberta.

References


A Discrete Firefly Algorithm for the Scaffolding Modular Construction in Mega Projects

J.J. Liu\textsuperscript{a,b}, L. Hou\textsuperscript{b} and X.Y. Wang\textsuperscript{b}

\textsuperscript{a}School of Science, China University of Petroleum, Beijing 102249, China
\textsuperscript{b}Australasia Joint Research Centre for BIM, Curtin University, Australia

E-mail: liujj@cup.edu.cn, lei.hou@curtin.edu.au, xiangyu.wang@curtin.edu.au

Abstract -

Productivity is critical to large and complex capital projects in sectors like infrastructure and resources. It normally relates to a significant scheduling issue when pursuing both time and cost objectives. A poor planning of project schedule might incur both excessive labour cost and project delay. Looking at scaffolding activity, a foremost and prerequisite construction activity for the subsequence LNG equipment erection, this study identifies the present research gaps: very limited research emphasis has been placed on the impact of design of time-cost optimization in scaffolding, as well as the feasible solutions. Therefore, the practical guidance in scaffolding construction planning is treated as a critical research focal point in this study. Modular construction of scaffolding has been extensively implemented in LNG projects due to the risks of working on offshore platform and the enormous working area. The emphasis of this study therefore is placed on addressing time-cost trade-off problems (TCTPs) of scaffolding modular construction. In particular, the questions of how to resolve the relatively multi-objective optimization issues using the Firefly algorithm and how to model the time and cost objective functions under specified constrains are answered. It is concluded that this paper, for the first time, looks into the complexity of scaffolding scheduling and gives solutions using such a robust optimization algorithm. In parallel, a simulation case study has been carrying out to test this algorithm based solution, to see how it can help scaffolding planners develop practical project schedules in either cost effective or time saving manners, or other trade-offs in between.

Keywords -

LNG; Time and cost trade-offs (TCTPs); Firefly Optimization Algorithm; Scaffolding

1 Introduction

Productivity is critical to increase Australias GDP. Analysis shows that lifting labour and capital productivity can result in AUD $90 billion additional income per year by 2017 [1]. This is especially so for large and complex capital projects in sectors like infrastructure and resources. The data shows a 0.7% annual decline in productivity between 2005 and 2011 compared with a 2.4% increase from 1993 to 1999 [1]. This underscores the urgency of getting productivity right and it is a priority area that can reap large rewards in future income growth [2]. Nowadays, scaffolding becomes an irremovable concern to industries across oil and gas, building, and infrastructure, considering the relatively low productivity and high labor shortage and cost. Poor design, planning and scheduling of scaffolding often lead to issues such as idling, rework, unnecessarily long travelling time between activities, which substantially reduce productivity. In mega projects that of great complexity such as offshore LNG (Liquefied Natural Gas) platform, high-rise building and infrastructure, scaffolding concerns a wide range of activities with different natures. Thus, construction process of scaffolding has a significant impact on the subsequent quality, safety and profitability of construction projects. However, the research works devoted to scaffolding issues may seldom or never be performed before and therefore are short of empirical data on guiding enacting offshore scaffolding schedule, despite its crucial importance [3].

In response to this, this paper focuses on scaffolding scheduling issues in an offshore LNG construction project, where the scaffold work is very large scale and involves huge amounts of factors such as materials, work crew, equipment, etc. The cost and efficiency from designing scaffolding schedule to its actual erection could be regarded as the focal point when conducting productivity research. Scaffolding normally applies modular construction in LNG projects. On the one hand, since laying LNG equipment takes up a lot of size and height, scaffolding therefore also has huge amount of work. The entire structure of scaffold, after modularization, can be divided into several small modules for parallel construction in different workshops, which saves time and workload. On the other hand, working on offshore platform is very risky as falling from the platform might incur drowning accident. Given the features of scaffolding modular con-
strucution, the objective of optimizing scaffolding sched-
ule is to generate scaffolding design, erection and disman-
tling scheme, and their associated schedule and resource
requirements and estimated cost by holistically consider
multiple inputs (materials, work crew, equipment, etc.)
under the constraints of design and progress of the build-
ing work, safety, workspace, budget, specification and
codes, and so on. The emphasis should be placed on pro-
ducing a series of outputs that have been optimised to in-
form decision-making.

2 Literature Review

Offshore LNG scaffolding engineering design prob-
lems generally involve several objectives. These objec-
tives, related to the technique and economic performance
of the engineering system, are potentially conflicting in
nature. Multidisciplinary design optimization has grown
to the point of gaining near universal recognition in its
ability to lead to better designs [4]. Time and cost are
two critical objectives of modular scaffolding construc-
tion, which are intricately concomitant to each other. To
each workshop, using more workforces may speed up the
progress of construction. But the total cost added up from
each workshop would be definitely higher. Due to a limit-
ed research work on selecting options with corresponding
time and cost to complete highly productive scaffolding
activities, scaffolding scheduling is always accompanied
by uncertainty. The emphasis is thus placed on working
out the solution of allocating multiple appropriate re-
sources for each modular construction to obtain the ob-
jective of overall time and overall cost. The timeCost
trade-off problems TCTPs are multi objective optimiza-
tion problem [5]. The TCTPs address the project scheme
options under the constraints of both project duration and
project cost, so as to select the best trade-offs to com-
plete an activity. Given that randomness and fuzziness
may co-exist in project scheduling problem, Ke and Liu
[6] investigate various types of project scheduling prob-
lems using fuzzy activity, which can also be applied as
the supplement of diverse algorithms to solve the soft-
ware project scheduling problems[7]. To help project
planners develop practical project schedules without im-
pacting project quality, Kim et al. [8] take into account
the potential quality loss cost in TCTPs, and propose a
mixed integer linear programming model that considers
the excessive crushing activities. In terms of the resour-
ces-constrained (both renewable and non-renewable resources
are as constraints) multi-mode scheduling solutions, Ant
Colony Optimization (ACO) was studied by Li and Zhang
[9], who manifest that, against other metaheuristic meth-
ods, this method is particularly beneficial for industry
practitioners in real construction projects. Scheduling
problems are normally concerned with assembly in flow-
shop, such as queries scheduling and makespan, which is
one of most difficult NP-complete problems. Allahver-
Search (TS) approach with other types of algorithms such
as particle swarm optimization (PSO) and EDD, and ex-
perimentally prove its quality (more effective) and per-
formance (less error-prone). Simulated-Annealing (SA)
can be applied in addressing the similar issue. Accord-
ing to Varadharajan and Rajendran [12], SA yields the
most optimum and less computational solution in the net
non-dominated multi-objective genetic local search than
Elitist Non-dominated Genetic Algorithm (ENGA) [13]
and (Gradual Priority Weighting) GPW approaches[14].
The population-based approaches incorporating differen-
t strategies for generating and improving a population of
schedule have been very booming in achieving the similar
objective, for example, the widely used Genetic Algorith-
m (GA) or Hybrid Genetic Algorithm (HGA) [15], [16],
Ant Colony Optimization (ACO) [17], Particle Swarm
Optimization Algorithm (PSOA)[18], Differential Evolu-
tion Algorithm (DEA) [19], Artificial Bee Colony Algo-
risthm (ABCA)[20], etc. It is experimentally proved these
population-based algorithms are averagably more competi-
tive and efficient in searching local optimum in continu-
ous and discrete multi-objective scheduling problems.

Firefly Algorithm (FA) is a novel algorithm which was
firstly proposed in year 2010. Since then, numerous re-
search works and literatures (detailed in [23]) have man-
ifested its advantages over algorithms mentioned above
in dealing with a wide range of issues. The study does
not only establish the mathematical model of scaffolding
modular construction in mega projects, but also propose
a discrete self-adapted FA and demonstrate its viability in
producing time-cost solutions.

3 Problem Description and Mathematical
Model

Modular construction of scaffolding only considers
each module is produced by an independent workshop,
and the ultimate goal is the completion of each module so
that the following final assembly can start. In the math-
ematical model, it assumes that each worker maintains the
same productivity (working volume/time). Within each
scaffolding module, there are a number of working step-
s, where each working step requires a number of work-
ers. Our objective function consists of two parts. First,
there should incur the least cost after all the modules are
completed. Second, there should be the minimum time
consumption for the most time-consuming module. The
assembly sequence within each module abides by prece-
dence relationship (a group of task components are sub-
ject to order requirement). Resources are confined by
the minimum and maximum number of workers available. Our mathematical model for the scaffolding modular construction problem is presented. First, notations used in our model are described and then the mathematical model is explained.

The notations are described here.

\[ I : \text{maximum number of modules} \]
\[ J_i : \text{maximum number of assembly steps to produce scaffold module } i \]
\[ P : \text{set of precedence relations between procedures to product a scaffold module} \]
\[ i : \text{module index, } i = 0, 1, \ldots, I + 1, (i = 0 \text{ or } I + 1 \text{ are dummy module}) \]
\[ j : \text{assembly step index in each module, } j = 0, 1, \ldots, J_i + 1, (j = 0 \text{ or } J_i + 1 \text{ are dummy steps}) \]
\[ d_{ij} : \text{duration of performing } j\text{-th assembly step in module } i \]
\[ c_{ij} : \text{cost of performing } j\text{-th assembly step in module } i \]
\[ t^F_{ij} : \text{finish times of last assembly } J \text{ in module } i \]
\[ t^S_{ij} : \text{start time of performing } j\text{-th assembly step in module } i \]
\[ t^F_{ij} : \text{finish time of performing } j\text{-th assembly step in module } i \]
\[ t^D_{ij} : \text{deadline of module } i \]
\[ x_{ij} : \text{resource variables, it can be continuous or discrete} \]
\[ L : \text{lower bound of resources} \]
\[ U : \text{upper bound of resources} \]
\[ C_{\text{max}} : \text{maximum total cost} \]
\[ T : \text{time delay when finishing all the modules} \]

The mathematical model of our problem (P) is presented as below:

\[
\begin{align*}
\min C_{\text{max}} &= \sum_{i=1}^{I} \sum_{j=1}^{J_i} c_{ij} \\
\min T &= \sum_{i=1}^{I} \max \left(0, t^F_{i,J_i} - t^D_{i} \right) \\
t^S_{(i-1)J} + d_{(i-1)J} &\leq t^S_{ij}, \quad \text{for all } i \tag{3} \\
t^S_{ij(J-1)} + d_{i(J-1)} &\leq t^S_{ij}, \quad \text{for all } i \text{ and } j \tag{4} \\
d_{ij} &= f^S_{ij}(x_{ij}) \tag{5} \\
c_{ij} &= f^F_{ij}(x_{ij}) \tag{6} \\
L &\leq \sum_{i=1}^{I} \sum_{j=1}^{J_i} x_{ij} \leq U \tag{7}
\end{align*}
\]

where \( i = 0, 1, \ldots, I + 1; j = 0, 1, \ldots, J_i + 1 \).

In above model, the objective function (1) shows the total modules cost. The second objective of our problem, as shown in objective function (2), is the total time to finishing all the modules, and should be minimized. The precedence relationships between activities and between workshops are described by (3) and (4), respectively. Equations (5) and (6) relate, respectively, the duration and the cost of performing \( j\text{-th assembly step in module } i \) to the resource variables. Formula (7) decides the scope of the resources. Based above mathematical model, a discrete Firefly algorithm will be applied to handle the scaffolding modular construction problem in next section.

4 Firefly Algorithms (FA) for Addressing TCTPs

In this section, a discrete FA algorithm will be present and then applied to TCTPs.

4.1 Discrete Firefly Algorithm

Firefly algorithm is a novel nature-inspired algorithm inspired by social behavior of fireflies. Fireflies are one of the most special, captivating and fascinating creature in the nature. By idealizing some of the flashing characteristics of fireflies, firefly-inspired algorithm was presented by Yang [21].

Firefly-inspired algorithms use the following three idealized rules:

(i) All fireflies are unisex which means that they are attracted to other fireflies regardless of their sex;

(ii) The degree of the attractiveness of a firefly is proportion to its brightness, thus for any two flashing fireflies, the less brighter one will move towards the brighter one and the more brightness means the less distance between two fireflies. If there is no brighter one than a particular firefly, it will move randomly;

(iii) The brightness of a firefly is determined by the value of the objective function. For a maximization problem, the brightness can be proportional to the value of the objective function.

In the discrete firefly algorithm, there are four important issues:

Attractiveness: In the firefly algorithm, the main form of attractiveness function \( \beta(r) \) can be any monotonically decreasing functions such as the following generalized form:

\[
\beta(r) = \beta_0 e^{-\gamma(r)}}\), \( n \geq 1 \tag{8}
\]

where \( r \) is the distance between two fireflies, \( \beta_0 \) is the attractiveness at \( r = 0 \) and \( \gamma \) is a fixed light absorption coefficient.

Distance: The distance between any two fireflies \( i \) and
\[ r_{ij} = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2} \quad (9) \]

where \( x_{i,k} \) is the \( k \)-th component of the \( i \)-th firefly.

**Movement:** The movement of a firefly, \( i \) is attracted to another more attractive (brighter) firefly \( j \), is determined by

\[ \bar{x}_i = x_i + \beta(r) \times (x_i - x_j) + \alpha(r - \frac{1}{2}) \quad (10) \]

where the second term is due to the attraction while the third term is randomization with \( r \) is a random number generator uniformly distributed in \([0, 1]\). In this paper, a outline update mode is applied in original FA, that means the light intensity is updated and the new position of firefly is evaluated after all the firefly finish moving.

The steps of the FA are given below [21]:

**Step 1.** Generate initial population of \( n \) fireflies \( x_i, (i = 1, 2, \ldots, n) \) randomly each of which represents a candidate solution to the optimization problem with objective functions \( f(x) \) and decision variables \( x_i = x_{i1}, x_{i2}, \ldots, x_{in} \).

**Step 2.** Compute light intensity using Eq. (8) for each firefly \( \beta = \beta_1, \beta_2, \ldots, \beta_n \). The distance between fireflies is computed from Eq. (9).

**Step 3.** Move each firefly \( i \) toward other brighter fireflies using Eq. (10). If there is other brighter firefly move it randomly.

**Step 4.** Evaluate new solutions and update light intensity.

**Step 5.** Rank the fireflies and find the current best solution.

**Step 6.** Repeat steps 2-5 until termination criterion is satisfied.

Based on the effectiveness of the firefly algorithm in optimizing continues problems, it is predictable that this algorithm would be impressive to solve discrete optimization problems which creates the motivation for proposing a discrete firefly algorithm. Sayadi et al. [22] developed a discrete version of FA (DFA) which can efficiently solve NP-hard scheduling problems, while a detailed analysis has demonstrated the efficiency of FA over a wide range of test problems, including multi-objective load dispatch problems. Furthermore, DFA can also solve scheduling and travelling salesman problem in a promising way [23].

**Discretization:** When the firefly \( i \) moves toward firefly \( j \), the position of firefly \( i \) is changed from a binary number to a real number. Therefore, we must replace this real number by a binary number. The following sigmoid function restricts \( f(x_{ik}) \) to be in the interval \([0, 1]\):

\[ f(x_{ik}) = \frac{1}{1 + e^{-x_{ik}}} \quad (11) \]

or

\[ f(x_{ik}) = \frac{e^{2x_{ik}} - 1}{e^{2x_{ik}} + 1} \quad (12) \]

where \( f(x_{ik}) \) denotes the probability of \( x_{ik} \) equalling 1 (See Figure 1). Then we can map interval \((0, 1)\) into the integral interval which is the feasible area of the problem. The results of numerical experiments show that Eq.(12) is better than Eq.(11) in most cases. So the Eq.(12) is used in this paper.

The steps of the DFA can be summarized as the pseudo code shown in **Algorithm 1**.

### 4.2 An adaptive Discrete FA for addressing TCTPs

For a better performance of DFA, an adaptive search strategy is proposed which consists of the parameters varies with the current iteration number \( t \).

\[ \alpha = \alpha_0 - \frac{1}{1 + e^{-(t - t_{\text{max}})}} \quad (13) \]

where \( t_{\text{max}} \) is the number of the maximum generation, \( \alpha_0 \) is the maximum value of \( \alpha \). Figure 2 demonstrates the variation of \( \alpha \) with respect to \( t \). This adaptive variation can play a balance adjustment of convergence and diversity.

In our TCTPs of the scaffolding modular construction problem there is one kind of variable \( x_{ij} \), which denotes the number of works performing \( j \)-th assembly step in modular \( i \). Both minimized objective functions and four constraints are taken into account in the mathematical model. The linear weight method is used to convert two objective functions into single objective by the following equation:

\[ F(x_{ij}) = \omega_1 C_{\text{max}} + \omega_2 T \quad (14) \]
where $\omega_1$ and $\omega_2$ are weights decided by user according to his expect on each objective function. Generally, $\omega_1 + \omega_2 = 1$. The position for the $k$-th firefly in the $t$-th generation can be denoted as $x_{kij}^t$. For convenience, $x_{kij}^t$ is expressed in a vector $(x_{k11}^t, \ldots, x_{kij}^t, \ldots, x_{k1n}^t, \ldots, x_{ijn}^t)$. As to dealing the constraints, we adapt the technique of the penalty function to convert the constrained optimization problem into unconstrained optimization.

$$\min P(x_{ij}) = F(x_{ij}) + G(x_{ij})$$  \hspace{1cm} (15)

Algorithm 1 Procedure of DFA

1: Generate initial population of fireflies $x_i, (i = 1, 2, \ldots, n)$;
2: Suppose that $f(x_i)$ is the objective function of $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$;
3: Light intensity $LI_i$ at $x_i$ is determined by $f(x_i)$;
4: Set light absorption coefficient $\gamma$, randomization parameter $\alpha$ and maximum generations $t_{max}$.
5: while $t < t_{max}$ do
6: for $i = 1 : n$ do
7: for $j = 1 : n$ do
8: if $LI_j > LI_i$ then
9: Move firefly $i$ towards $j$ in $d$-dimension by equation (13), (8) and (10)
10: end if
11: end for
12: end for
13: Attractiveness varies by equation (8)
14: Discrete the position of $i$-th firefly by equations (11) or (12)
15: Evaluate the new position of $i$-th firefly and update light intensity $LI_i$.
16: Rank the fireflies and find the current best;
17: end while
18: Show result and visualization;

where

$$G(x_{ij}) = \mu \sum_{i=1}^{j} \sum_{j=1}^{f} \max \{0, t_{i(j-1)}^S + d_{i(j-1)} - t_{ij}^S\}$$

$$+ \lambda \sum_{i=1}^{j} \max \{0, t_{i(j-1)}^S + d_{i(j-1)} - t_{ij}^S\}$$

and $\mu, \lambda \rightarrow \infty$.

Hence the discrete Firefly algorithm can perform for the scaffolding modular construction problem.

4.3 Case Study

A real 3-module LNG construction case study derived from an offshore LNG platform is selected to fit into the proposed model. Here, assembly steps had several options of time and resources. The module $i$, assembly step $j$, precedence relationships $P_{ij}$, cost of each step $c_{ij}$, duration of each step $d_{ij}$, and deadline $t_{ij}^D$ of each module are detailed in Table 1.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$j$</th>
<th>Description</th>
<th>$c_{ij}$</th>
<th>$d_{ij}$</th>
<th>$t_{ij}^D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>cantilever</td>
<td>-</td>
<td>$3x_{11} + 2$</td>
<td>$24 - 5x_{11}$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>vertical</td>
<td>01</td>
<td>$3x_{12} + 1$</td>
<td>$20 - 3x_{12}$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>reinforcing</td>
<td>02</td>
<td>$2x_{13} + 3$</td>
<td>$15 - x_{13}$</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>reinforcing</td>
<td>13</td>
<td>$3x_{21} + 1$</td>
<td>$15 - 2x_{21}$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>tube</td>
<td>13</td>
<td>$x_{22} + 2$</td>
<td>$10 - x_{22}$</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>traversing</td>
<td>21-22</td>
<td>$2x_{32}$</td>
<td>$15 - 3x_{32}$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>cross brace</td>
<td>01-31</td>
<td>$2x_{31} + 1$</td>
<td>$12 - x_{31}$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>planking</td>
<td>01-31</td>
<td>$2x_{32} + 5$</td>
<td>$18 - 2x_{32}$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>guard system</td>
<td>01-31</td>
<td>$x_{33} + 4$</td>
<td>$14 - x_{33}$</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>safety net</td>
<td>01-31</td>
<td>$x_{34} + 3$</td>
<td>$15 - x_{34}$</td>
</tr>
</tbody>
</table>

Utilizing data or expresses into optimization model (15), we applied the presented DFA to find the solution of problem (P). Where the parameters are preset as follow: population size is 40, Max Generation is 100, $L = 0$ and $U = 50$. Discretization is performed by using of equation (12), and then the real variables $x_{ij}$ are mapped into integral interval $[1, U]$. After a run of the proposed algorithm, a quasi-optimal duration, cost, and workers allocation of the problem is presented in Table 2.

Figure 3 shows the performance of the algorithm for the case and Figure 4 shows the pareto front of the case.
Table 2. Optimal assembly duration and cost.

<table>
<thead>
<tr>
<th>i</th>
<th>j</th>
<th>Description</th>
<th>Duration(day)</th>
<th>Cost($10^3$)</th>
<th>Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>cantilever</td>
<td>19</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>vertical reinforcing tube</td>
<td>18</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>post</td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>horizontal reinforcing tube</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>traversing lever</td>
<td>7</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Joist</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>cross brace</td>
<td>11</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>planking</td>
<td>9</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>guard system</td>
<td>6</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>safety net</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td>53</td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. The convergence curve of results

5 Conclusions

The paper studied the project scheduling problem with modular construction activity. For obtaining the minimal total modules cost and minimal total time to finishing all the modules, a multi-objective constrained optimization model was established and a modified intelligent algorithm (DFA) to find the optimal solution of this model was proposed. FA is an excellent intelligent algorithm which has been successfully applied in several real discrete optimisation problems [23]. The focal point of the paper is to investigate the scaffolding modular construction issues, which belong to a series of discrete TCTPs. Through a substantial project-based case study, the results manifest the credibility of the optimisation by producing a better solution of workforce allocation for the sake of time and cost balance.

6 Acknowledgements

Part of this research was supported under Australian Research Council Linkage Project scheme (project number: LP130100451).

References


[8] Kim, J., Kang, C. and Hwang, I. A practical approach to project scheduling: considering the potential quality loss cost in the timeCost tradeoff prob-


Enhanced Online LS-SVM Using EMD Algorithm for Prices Prediction of Building Materials

Ying-Hao Yu\textsuperscript{a}, Hsiao-Che Chien\textsuperscript{a}, Pi-Hui Ting\textsuperscript{b}, Jung-Yi Jiang\textsuperscript{a} and Pei-Yin Chen\textsuperscript{a, *}

\textsuperscript{a} Department of Computer Science and Information Engineering, National Cheng Kung University, Taiwan
\textsuperscript{b} Department of Business Administration, Chang Jung Christian University, Taiwan
E-mail: yinhaun@gmail.com, max6060777@hotmail.com, tphui@mail.cjcu.edu.tw, jungyi@ismp.csie.ncku.edu.tw, pychen@csie.ncku.edu.tw

Abstract -
Cost estimation is economically critical before starting off a construction project. One of the essential assignments for materials’ prices prediction is to control the cost of inventory. Even though the prediction system based on support vector machine (SVM) recently has been emerged as a favourable choice, the prediction accuracy of SVM is usually deteriorated with nonstationary price data. Thus the way to explore workable price prediction still remains a challenge to be resolved for materials’ cost control. In this paper, an enhanced online least squares support vector machine (LS-SVM) is proposed to predict the trend of building materials prices. Our design is to incorporate with empirical mode decomposition (EMD) to deconstruct nonlinear and nonstationary data for the set of intrinsic mode functions (IMFs), which are represented in sinusoid-like waveforms. Superior prediction, therefore, can be attained by predicting IMFs with online LS-SVMs. According to our simulation results, proposed EMD designs notably improve prediction accuracy from online LS-SVM and are workable for the cost estimation of building materials.

Keywords -
Support Vector Machine; Nonstationary Data; Empirical Mode Decomposition; Intrinsic Mode Functions; Online LS-SVM

1 Introduction
Effective cost estimation is paramount to stakeholders for evaluating viability of a construction project. Categorization of this assignment is about to control resources comprehensively of materials, labour, and equipment [1]. Of these resources, severe fluctuation of material cost from domestic and international economic influences is usually imputed for the principal cause of failure [2]. This stresses the indispensability of materials’ prices prediction during pre-project planning.

In past years, the expert systems based on support vector machine (SVM) have become a mainstream in statistical machine learning and prediction. The initial motivation of SVM was to classify patterns by linearly or nonlinearly separating classes using a hyperplane [3]. This idea soon had been improved for linear or nonlinear function’s estimation and is also known as the support vector regression (SVR) [3], [4]. After the release of SVM and SVR, Suykens and Vandewalle proposed a refined version called least squares support vector machine (LS-SVM) [5]. In this work, least squares loss function and equality were designed to substitute for ε-insensitive loss function and inequality constraints in SVM and SVR [6]. Such reformulation significantly reduces the computing load for large data set and makes it more popular in many prediction systems [6].

Nevertheless, when support vector machine works with nonstationary data, insufficiency of this kind of system will manifest. The issue arises from SVM’s identical form, which utilizes a linear function \( f(x) \) involving kernel function to resolve nonlinear problems. This design sometimes is improbable in that a unique function cannot satisfy whole sequence of non-stationary time series [7]. Such limitation is inapplicable for financial analysis, e.g., exchange rate and price prediction [8].

Based on the merits of SVM, some enhancements have been developed for nonstationary systems. For instance, in the work of Chang et al. [7], single linear function of SVM is replaced by multiple functions for nonstationary time series. Zhang et al. suggest deconstructing nonstationary signal beforehand by using wavelet packet transform [9], and LS-SVM cooperating with differential evolution is designed by Chen et al [2]. For more efficiency, the algorithm of empirical mode decomposition (EMD) has been utilized to deconstruct nonstationary data for sinusoid-like signals IMFs, then SVMs can mostly predict (track) almost-stationary IMFs.
Another algorithm named local mean decomposition (LMD) was reported to take over EMD by better signals quality and prediction accuracy, but this algorithm is highly depending on optimal smoothening processes to signals [12].

In this paper, we propose a LSSVM-based system for prices trend prediction of building materials. Prediction accuracy of proposed system was tested with two building materials of copper and aluminium. Our system utilizes online algorithm to dynamically update training database with daily prices [13], [14]. For more accuracy, the algorithm of EMD is adopted to deconstruct nonstationary price data before online LS-SVM. Superior accuracy of online EMD LS-SVM can be further achieved by improving the prediction performance of IMF1.

This paper is arranged as the follows. In the Section 2, the algorithms of our prediction designs such as EMD, LS-SVM, and online algorithms will be detailed. Comparisons among prediction results with different LSSVM-based designs will be listed in the Section 3. A short discussion about the method to enhance online EMD LS-SVM is in the Section 4. Finally, the conclusion is drawn in the Section 5.

2 Prediction Designs

Proposed online EMD LS-SVM is shown in Figure 1. Each material’s price data are firstly deconstructed by EMD algorithm for IMFs and a residual item $R_n$. These signals are then processed respectively by a trained online LS-SVM. The trend of material’s price can then be composed by sum of online LS-SVMs’ outputs. Algorithms of EMD, LS-SVM, and online algorithms are expressed in the following sections.

![Figure 1. The configuration of proposed online EMD LS-SVM](image)

2.1 Empirical Mode Decomposition

Empirical mode decomposition (EMD) was originally designed to deconstruct nonstationary time series data to IMFs. Huang et al. designed this algorithm for examining IMFs’ instantaneous frequencies by Hilbert spectral analysis (HSA) in order to avoid complicated computing [15]. Generally, the locus of an IMF is sinusoid-like signal and similar to the harmonic of original signal. The difference is that IMF’s signal might have various amplitude and frequencies [15].

The algorithm to extract IMFs can be summarized as the following steps:

1. Extract local extrema points (maxima and minima) of tested signal $x(t)$.
2. Determine upper envelope $x_u(t)$ and lower envelope $x_l(t)$ by linking up local maxima and minima respectively.
3. Derive the first mean $m_1(t)$ between upper and lower envelopes by:
   \[ m_1(t) = \frac{(x_u(t) + x_l(t))}{2} \]  
4. Define the first error $h_1(t)$ as:
   \[ h_1(t) = x(t) - m_1(t) \]
   where $h_1(t)$ should conform to the properties of IMF as:
   a. The difference between extrema and zero-crossing points number of the whole data set must be $\leq 1$.
   b. The mean value of envelopes which are composed of the local maxima and minima should be zero at any point.
   If $h_1(t)$ satisfies the requirements of an IMF, the first IMF function can be confirmed as $C_1(t) = h_1(t)$ — otherwise extracting procedure goes back to the step 1 and replaces $x(t)$ with $h_1(t)$.
5. Determine the residual item $R_1(t)$ as:
   \[ R_1(t) = x(t) - C_1(t) \]
6. Repeat the steps from 1 to 5 for IMF2 by replacing...
x(t) with \( R_{\text{in}}(t) \), and extracting processes will stop when \( R_{\text{in}}(t) \) becomes a monotonic function.

As shown in Figure 1, the nonstationary signal can be deconstructed into IMFs and a residual function. Since EMD’s outputs are the subsets of \( x(t) \), original signal can be expressed as:

\[
x(t) = \sum_{i=1}^{c} C_i(t) + R_{\text{res}}(t)
\]

### 2.2 Least Squares Support Vector Machine

Considering a given training data set which is defined as \( D = \{(x_i, y_i), \ldots, (x_l, y_l)\} \), \( x \in \mathbb{R}^n \), \( y \in \mathbb{R} \). The LS-SVM algorithm defines a linear function \( f(x) \) as:

\[
f(x) = \langle \omega, \phi(x) \rangle + b
\]

where \( \langle \cdot, \cdot \rangle \) denotes the dot product, \( \omega \) is the weight vector, \( b \) is a bias, and \( \phi(x) \) represents a mapping function to map the input vectors into a high-dimensional feature space. The goal of prediction is to find a function \( f(x) \), which has limited error to the actual targets \( y_i \) from training database. Thus equation (5) becomes an optimal problem for:

\[
\min \frac{1}{2} \| \omega \|^2 + \frac{1}{2} \gamma \sum_{i=1}^{l} e_i^2;
\]

subject to \( y_i = \langle \omega, \phi(x_i) \rangle + b + e_i \), \( (i = 1, \ldots, l) \) (6)

where \( e_i \) denotes the variable of error for misclassifications, and \( \gamma \) is defined as the penalty parameter to minimize estimation error and maintain function’s smoothness [6], [11].

To resolve equation (6), Lagrangian function can be utilized to find out \( \omega \) and \( e \). It can be written as:

\[
L_{\text{LS-SVM}} = \frac{1}{2} \| \omega \|^2 + \frac{1}{2} \gamma \sum_{i=1}^{l} e_i^2 - \sum_{i=1}^{l} \alpha_i \left( \langle \omega, \phi(x_i) \rangle + b + e_i - y_i \right)
\]

where \( \alpha_i \) is Lagrange multiplier, which can be either in positive or negative value. The conditions for optimality of equation (7) are:

\[
\frac{\partial L}{\partial \alpha_i} = \omega - \sum_{i=1}^{l} \alpha_i \phi(x_i) = 0
\]

\[
\frac{\partial L}{\partial b} = \sum_{i=1}^{l} \alpha_i = 0
\]

\[
\frac{\partial L}{\partial e_i} = \gamma \cdot e_i - \alpha_i = 0, \ (i = 1, \ldots, l)
\]

Finally, the LS-SVM for function estimation can be re-written as:

\[
f(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + b
\]

where the \( K(x_i, x) \) is known as the kernel function in the form of Gaussian radial basis function (RBF) as:

\[
K(x_i, x) = \exp \left( -\frac{\| x_i - x \|^2}{2\sigma^2} \right)
\]

The \( K(x_i, x) \) must satisfy Mercer’s theorem, and \( \sigma \) is the width of RBF. After the implementation of kernel function, we can nonlinearly map training data onto an infinite-dimensional space to resolve nonlinear problems [6], [11].

### 2.3 Online Algorithm for LS-SVM

Prediction accuracy of LS-SVM is depending on the features of trained data. If the features of upcoming data (signals) are different to trained data, original prediction function will lose tracking to the trend of upcoming data. This leads users confronting with such problem have to re-train system in order to learn the features of latest data. To resolve this problem, an “online” mechanism has been implemented in system to dynamically update training database during prediction. This updating mechanism can be summarized as the following steps [16]:

---

304
3 Simulation Results

Predicted building materials prices of copper and aluminium are shown in Figure 2. The data of sampled materials prices were from the futures market of London. In our test, the EMD LS-SVM were first trained with 1012 daily prices for each material and then the prices on 233 more days from different year were tested by combining with online predicting algorithms. Collected data sets of both materials in Figure .2 are obviously nonlinear and non-stationary. The Figure 2 indicates the prediction results of proposed online EMD LS-SVM can track the actual material prices efficiently.

When compare with the different prediction schemes based on LS-SVM, the criterion of mean absolute percentage error (MAPE) is chosen for evaluation indicator as the follows [11]:

$$MAPE = \frac{1}{M} \left[ \sum_{i=1}^{M} \left| \frac{r_i - f_i}{r_i} \right| \times 100 \% \right]$$

where $r_i$ denotes the actual price of building material. The $f_i$ is the predicted price, and $M$ is the sampling number by days.

Table 1 lists prediction accuracy with different LSSVM-based designs surveyed by using MAPE. It can be seen that online algorithm inevitably improves the traditional LS-SVM by updating training database dynamically with incoming price data. This prediction result is even better than the EMD LS-SVM algorithm. Such a contradiction between the EMD LS-SVM and online LS-SVM arises from the sinusoid-like IMFs, which are still partially non-stationary with price data and deteriorates the tracking performance of trained LS-SVMs.

However, even though online LS-SVM has better tracking performance than LS-SVM and EMD LS-SVM for non-stationary and sinusoid-like waveforms, we can further improve online prediction accuracy if it cooperates with EMD. As shown in Table 1, the last prediction results indicates over 40% MAPE improvement from online LS-SVM can be achieved after incorporation of EMD algorithm.

In the end, the prediction accuracy in Figure 2 is not only sufficiently demonstrated for daily prices but also feasible for the weekly and monthly prices predictions by cooperating with online and EMD algorithms. Although we only demonstrated prediction results based on daily prices, with the same skill in the future, weekly and monthly prices can also be predicted after the accumulation of daily prices.
Figure 3. Comparisons between original and improved tracking for IMF1 signals
Table 1. Error comparisons among different LSSVM-based predictions

<table>
<thead>
<tr>
<th>MAPE (%)</th>
<th>Copper</th>
<th>Aluminium</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS-SVM</td>
<td>2.38543</td>
<td>6.36419</td>
</tr>
<tr>
<td>EMD LS-SVM</td>
<td>2.07504</td>
<td>4.6045</td>
</tr>
<tr>
<td>Online LS-SVM</td>
<td>0.87321</td>
<td>0.86573</td>
</tr>
<tr>
<td>EMD Online LS-SVM</td>
<td>0.46545</td>
<td>0.50844</td>
</tr>
</tbody>
</table>

4. Discussions

Notwithstanding EMD LSSVM-based prediction can achieve sufficient accuracy, comparisons between LMD- and EMD-LSSVM recently had been made by Dong et al. [17]. In their works, LMD-LSSVM outperformed EMD-LSSVM by lower error rate and training duration of LMD. Since EMD has been maturely developed, a question might be asked for refining EMD for better prediction performance. After surveying the signal quality of IMFs and tracking capability of online EMD LS-SVM, prediction with EMD can be further improved by manipulating on IMF1 carefully.

As shown in Figure 3(a) and 3(c), IMF1 in each test scenario is composed of high frequency transitions and is severely non-stationary comparing with the other IMFs. The reason is that IMF1 involves abrupt transitions of original data, so these signals in IMF1 caused failure of tracking by using online EMD LS-SVM. This phenomenon can be observed in Figures 3(a) and 3(c); both tracking signals coloured with dotted blue lines were nearly flatten out.

For resolving unsuccessful tracking with IMF1 signals, our answer is to mitigate the components (nonstationary parts) that cause the failure of tracking. Here we suggest improving online EMD LS-SVM’s tracking by first double sampling IMF1’s signal. This manipulation is achieved by interpolating a virtual price, which is from the average prices of every two days. Since the data envelopes of maxima and minima are symmetric to time axis, the average values of higher frequency transition in IMF1 will be very close to the time axis (null). In other words, by double sampling IMF1’s signal, some predictions for minima-to-average or average-to-maxima will be approximated to a stationary relationship. Consequently, online LS-SVM’s tracking capability with double sampling rate can then be improved. After tracking with double sampled signals, the tracked parts of virtual prices can be removed in order to restore IMF1 to original resolution but with better online LS-SVM tracking performance.

The improvements of tracking IMF1’s signal by manipulation of double sampling are shown in Figures 3(b) and 3(d). By comparing with MAPEs before double sampling, the error rates are all significantly decreased. The remainders in IMF1 which cannot be tracked by online EMD LS-SVM are mostly with larger amplitude and could be classified as the noise. Here proposed double sampling mechanism on IMF1 makes an online EMD LS-SVM error rate 0.379% for copper and 0.423% for aluminium. Both results are also over 50% improvement from online LS-SVMs’ MAPE, as shown in Table 1. It can be seen that proposed double sampling on IMF1 has the potentiality for a better EMD-based prediction in the future.

5. Conclusion

This paper has proposed a LSSVM-based system for prediction of building materials’ prices. Higher prediction accuracy is achieved by online mechanism for dynamically updating training database. Moreover, the deconstruction of nonlinear and non-stationary data by empirical mode decomposition (EMD) can enhance prediction accuracy by tracking sinusoid-like IMFs and residual signals using online LS-SVMs. Comparing with latest LMD algorithm, double sampling IMF1’s signal can provide competitive improvement of prediction accuracy. Based on our successful works on enhanced prediction of materials prices, the future work will focus on the improvement of tracking capability for IMFs and expand prediction ranges for weekly and monthly materials prices.

Acknowledgement

Supports from Ministry of Education Taiwan for the Top University Project to the National Cheng Kung University (NCKU) is gratefully acknowledged.

References


An Interactive Progress Monitoring System using Image Processing in Mobile Computing Environment

Hongjo Kim, Kinam Kim, Sungjae Park, Jihoon Kim, and *Hyoungkwan Kim

School of Civil and Environmental Engineering, Yonsei Univ., Seoul, Korea
E-mail: hongjo@yonsei.ac.kr, kinam@yonsei.ac.kr, sjpark4691@yonsei.ac.kr, ghoon@yonsei.ac.kr, hyoungkwan@yonsei.ac.kr (*Corresponding author: hyoungkwan@yonsei.ac.kr)

Abstract -
A timely progress monitoring is essential for the success of a large-scale construction project. It enables a site manager to properly prepare resources and make plans for the remaining part of the construction activities. Recently, mobile computing and image processing have been investigated as a means of automating progress monitoring. Mobile computing is advantageous in wireless data recording, retrieval, and transfer whereas image processing is able to analyze the site images to extract progress information. However, their potential applications in construction monitoring were rather separately studied; synergistic effect of both techniques has not yet been fully materialized. This paper presents an interactive mobile progress monitoring system to enhance the existing progress monitoring practices. The system utilizes the interactive feature of tablet computer to combine the strength of image processing and mobile computing. When a user selects an object on a construction site image in a mobile computing environment, the system provides a list of attributes for the object of interest. In this interactive environment, the user can easily match the object to the right attributes such as location, material type, and relation with other objects. This initial matching can then allow for automatic matching of other objects on the site to the right attributes of their own. The method can raise accuracy of image processing and significantly reduce the effort required to do the pre-processing for progress monitoring. A cable-stayed bridge construction project, a case study, is used to validate the proposed system.

Keywords -
Construction management, Progress monitoring, Image Processing, Mobile Computing, Cable-stayed bridge

1 Introduction

In a large scale construction project, multiple activities are simultaneously conducted; a well-coordinated effort is needed to timely prepare the necessary resources, such as labor, equipment, and materials, for uninterrupted smooth execution of the project. However, changeable and uncertain site conditions often necessitate change of the original design or plan. The change management requires efficient communication among the participants of the project. One of the important subjects for communication is the progress status. Delay in the transfer of the progress information prevents project managers from making optimal management decisions, resulting in increase of the project cost. Thus, implementation of an effective communication system to monitor and share progress information is crucial for project success.

Mobile computing technology is emerging as a viable solution to handle the construction site information. Radio frequency identification (RFID), ZigBee, and smartphone technologies are being studied for improved site monitoring processes. To name a few examples, Kim et al. (2011) used RFID and ZigBee technologies to implement the ubiquitous sensor network for construction material monitoring [1]. Chen and Kamara (2011) proposed an information management framework to use mobile devices for acquiring and transferring site information [2]. Kim et al. (2013a) proposed using smartphone technologies for site monitoring, task management, and real-time information sharing [3].

Image processing is being investigated as a way to automatically capture and analyze construction site information. Wu et al. (2010) developed the 3D CAD-based filtering method to increase the accuracy of object identification in construction site images [4]. Chi and Caldas (2012) proposed image processing to monitor the safety status of earth moving operations [5]. Kim et al. (2013b) automated the process of 4D CAD model updating using image processing [6]. Rebolj et al. (2008) conducted construction activity monitoring using 4D model and image processing and shared the resultant information with site personnel using mobile devices [7].

Synergistic effect can be expected by using both mobile computing and image processing technologies for improved construction management. However, there is a lack of effort in combining the two technologies, and the opportunity has not yet been sufficiently gripped. This paper presents an interactive progress monitoring system based on image processing and mobile computing. Capitalizing on the strength of the two technologies, the progress information can be retrieved on site with high accuracy and the information can easily be shared with the project participants. One of the main contributions of this research is the flexibility of image processing techniques for progress monitoring. In the previous efforts in this area, users of progress monitoring techniques had to customize their algorithms to the site and naturally involved a lot of laborious tasks. The proposed system can be utilized as a tool for minimizing such efforts. The system requires the minimum intervention of users to automatically identify the whole progresses. The system was validated using a case study involving construction of a cable-stayed bridge.

1.1 Case Study Description

The proposed progress monitoring system was applied to the upper deck installation activities of the cable-stayed bridge construction. The bridge, located in Jecheon, Choongcheongbuk-do, Korea, has a length of 442m and a width of 14.5m, with two pylons. A total of 28 segments of upper decks comprise the superstructure of the bridge. Each deck was installed in the order of edge girders, floor beams, stringers, concrete panels, and in-situ concreting. The process of component installation, from edge girder to concrete panel, was supported by a derrick crane, and the whole process, even including in-situ concreting, was repeated for each deck installation.

In order to capture the images for deck construction, a closed-circuit television (CCTV) camera was installed at the cross beam of a pylon. The camera had the resolution of 704×480 and was equipped with pan, tilt, and zoom (PTZ) capabilities. The construction scenes were captured at a rate of one picture per second and the saved images were transferred to the main server in the form of JPEG.

2 An Interactive Progress Monitoring System

Figure 1 shows that the interactive progress monitoring system is composed of two hardware units: mobile device and main server. Mobile device is to interact with users for information input and output, whereas the main server is to process images and offer the site information. Since image processing in mobile computing environment requires high-level computing power, the server was used for the image processing part. The mobile device was used as the interface through which users interact with the system. In this system, MATLAB, Hypertext Preprocessor (PHP), and Microsoft Windows Server were used. PHP language was favored because it could generate a mobile website that could be accessed by any mobile device, irrespective of the operating system. The mobile website developed for the system, thus, could be accessed by Android users and Apple iOS users.

![Figure 1. Interactive progress monitoring in mobile computing environment](image)

To increase the level of automation of the system, the MATLAB source files were converted to executable files. Once the executable files were developed, they were, as the name indicates, executable by the commands transmitted through the mobile website. For the generation of the executable files, the MATLAB compiler toolbox was utilized. The compiler toolbox...
was to build executable files from MATLAB source codes, with the help of additional program compiler; in this study, Microsoft Visual Studio 2010 was used. The executable files generated in this way could be run in a computer system where MATLAB Compiler Runtime (MCR) was installed. As a result, the image processing functions could be performed seamlessly based on the commands from the mobile website.

From the functional perspective, the proposed system consists of two parts: progress identification and progress information management. The progress identification is to extract the current status of the project using the image processing, while progress information management is to retrieve progress information concerning the component of interest and update the database based on new information.

The progress identification part plays the role of raising accuracy of object identification in the image processing, and thus reduces the effort required to do the preprocessing for progress monitoring. First, a user selects the construction scene of interest, and then the image of the scene such as the one shown in Fig. 2 is sent to the mobile device. Next, the user can choose the component of interest by using the box like the one shown in Fig. 3(a). The image inside of the box is cropped and later used for pattern matching with a 3D CAD mask. The user can also choose the identity of the component using a list such as the one shown in Fig. 3(b). Both information (image and identity) are now sent to the server. The server now starts the image processing shown in Fig. 1 by relying on the information (image and identity).

For the image processing, the cropped RGB color image is first converted to a grayscale image. The grayscale image is now, with the help of the identity information, associated with the 3D CAD mask. Here, the 3D CAD mask is the image obtained in the CAD environment assuming the same angle, magnification, and perspective origin of the CCTV as shown in Fig. 4. The mask can be used for effective identification of construction components [4]. The 3D CAD mask (Fig. 4) is compared with the cropped image using the normalized cross correlation, to check the existence of the component. Once the whole process is completed, using the relative locations among the components of the segment, the same process is repeated for the other components until all components are checked. In this way, the progress identification can be done in the automatic manner and the current progress information is displayed to the user.

The progress information management part allows the user to retrieve progress information from the database; any input acquired from the mobile device can be used to update the database. For this function, the user first needs to select the site of interest in order to receive the corresponding site image. The user now can choose a particular component to extract its related information including the supplier, date of installation, and installation contractor. The user can also update the database with new information, such as most recent site pictures and comments of site inspectors. As for now,
the progress information management part is not yet fully implemented as a part of the system.

Figure 4. 3D CAD mask filter of the selected component

3 Conclusions

This paper proposed an interactive system to effectively monitor progress of construction sites based on mobile computing and image processing. Mobile device was used as the interactive interface platform by which the user can extract progress information from and input project information into the system. Image processing was used to automatically calculate the current progress level of the project. Combined together, the two technologies increased the level of automation of construction progress monitoring and image processing accuracy. Additionally, omnipresence of site information through the mobile device allowed the users to have prompt decision making process.

Future studies are still required. As previously mentioned, the progress information management part needs to be fully implemented. In addition, the system needs to be improved so as to be able to extract progress information from the images captured directly from the mobile device. Such improvements are expected to make the system readily applicable in real construction projects for significant increase of construction productivity and quality.

References


Abstract -

The theme of this contribution fits astride two strands of Construction Management Studies: “health and safety risk management” on construction site and “constructability” design approach. Specifically we faces two problems usually considered separately: the lack of specificity, synergy, effectiveness between safety planning and construction process and the lack of concurrence between design and construction process. The proposed approach assigns to design the key role of optimizing both constructability and safety planning in a unique and simultaneous procedure. The purpose is to improve “working plan” design level by specific construction simulation and the related safety assessment in the perspective of the better constructability performance and final building results. The proposed method is a Construction Management Approach supported by design tool based on working directly on design drawings. Starting from the representation of construction details, their construction can be simulated. This is obtained analyzing construction details by means of progressive drawings that express and simulate the breakdown of construction phases, chronologically processed and logically related. In these specific drawings all the elements of the construction site are graphically placed. The procedure allows both to perform the specific and detailed safety assessment for the realization of construction detail components, and to verify the constructability of them. The safety and constructability analysis is so performed in a graphical way during the design process not waiting the construction phase. So we are still able to intervene, if necessary, with amendments under the designer control.

Keywords -
Construction Management; Project Management; Safety; Constructability; Buildability; Design.

1 Introduction

In order to optimize the construction process and to achieve the best results, two strands of Construction Management studies should be investigated: “health and safety risk management” on construction site and “constructability” design approach.

The issue of safety on construction sites was born with the construction activity: the use of dangerous materials, the use of equipment and machineries in precarious and provisional environments and, above all, the need to “work at height”, make the construction activity with a greater risk of accidents.

At the same time was born the need to direct and coordinate safety factors with all the other resources throughout the project instruments by using modern management techniques to achieve the objectives [1, 2].

The issue should be of interest for all researchers involved in Construction Management. It is wrong who still thinks it is a matter of tedious formalities, probably unnecessary.

In order to achieve the integration of “constructability” and “safety”, investigation on working knowledge of general management and familiarity with the special knowledge domain related to the project are indispensable (Figure 1), [3].

Figure 1. Basic Ingredients in Project Management (PMBOK Guide)
Supporting disciplines such as computer science, decision support systems and drawing techniques, may also play an important role [4].

This paper faces two different kinds of problem, usually considered separately:

- the lack of specificity, synergy, effectiveness between safety planning and construction process;
- the lack of concurrence between design phase and construction process (Figure 2).

In order to take under control constructability and safety during design phase and not only during construction process, the proposed approach assigns to design the key role of optimizing both constructability studies and safety planning in a unique and simultaneous procedure.

This approach fits within the relationship “construction phase” – “design phase” (Figure 3).

The purpose is to improve “working plan” design level by specific construction simulation and the related safety assessment in the perspective of the better constructability performance and final building results.

The simulation is possible during design phase, that is the moment of building and construction site foreshadowing, together with all their related aspects.

Such issue comes from many Construction Management studies that deal with, in particular, the Constructability and Buildability theoretical field [5,6,19].

Computation and scheduling of the intervention can change and the designer may lose the unified management of the process, losing this way the synthesis between intentions and proposal. The gap may be apparent and limited to the operational dynamics while preserving the essence of the unity between design intent and result of the construction [7].

Conversely, the project may be, in part or in full, of hard construction: the dynamics cannot be congruent with the intentions. The final result may be not consistent with the requirements and it will therefore be not acceptable in terms of quality [8].

Some important studies have been developed with the aim to manage constructability and safety by means of design [14, 15, 16], other studies involved BIM to facilitate a smarter and safer infrastructure and building construction [17].

The proposed method (simulating dimensional, technical and technological characteristics of the building and construction site) tries to overcome the questions above specified.

2 Suggestion for an effective design method

This contribution concerns the Constructability, which embraces the functions both of project management and of design, covering a wider scope than the Buildability. Constructability interacts with the project management techniques that utilize optimally knowledge and experiences on building effective, to improve the achievement of the project objectives. Constructability involves the entire design process [9, 18].

In these theories plays key role the project intended as a platform to maximize the interests of its most effective function, within the Construction Process in general and within the Building Construction in specific.

Talking about "optimization" is always dangerous, because the parameters should be defined and references needed to assess the actual improvement of this instrument.

The design influences construction cost, for example, because any decision made at the initial stage of a project life cycle has far greater influence than those made in later stages or during construction phase (Figure 4), [10].
Starting from the aforesaid studies, the method put the optimization of building process into the hands of construction simulation, later described, and design drawings.

In this perspective the safest route is to direct experimentation on the actual impact in terms of improving the "easy to build".

The basic question remains: "Which are the design elements that we can manage in order to improve their effectiveness in the construction site?".

In first instance, the answer is obviously twofold: the design contents and representation techniques. These two means are of course intimately related: the representation is nothing more than the explicit and objective view of the contents, which, in their textual expression, may be not clear.

In fact the representation is much more, as it allows you to contextualize the content, especially in terms of requirements and performance. The representation allows you to make content consistent, in a geometric and technological upgrading, with the specific artifact to be built and with the specific environment in which the construction will be inserted [11].

This consideration of "visualization and consistency", has provided the basic idea of the method we propose. The idea is that the graphical representation is not only "passive image" of the content, but it is the "active checks" in terms of specificity and consistency. So, why not trying to improve the content by performing the opposite: can an analysis of the representation define new contents? And more, why not accentuate the "dynamism" of the project whose representation usually goes for "static sections" of the final result?

Constructability and safety can be joined by design: this is the reason why it is interesting to work directly on design drawings. The proposed method can be defined as a Construction Management approach supported by a design tool directly based on design drawings [12].

3 The method

The core of this approach is to intervene on the working drawings to simulate the construction process, usually only implicit in the representation of the "result". As preliminary phase it is necessary to define the construction site global needing: boundary conditions, site peculiarities, external constrains. This is useful in order to be able to manage every single element that can participate or even interfere with the construction phase.

After this, the method can be performed: it consists in the application of the methodological procedure described in par. 3.1 by means of the design tool described in par. 3.2.

3.1 Methodological procedure

The classic representation of the "finished product" holds a critical background: completeness graphics imposes a "designed order" often associated with a construction sequence whose effectiveness has yet to be demonstrated. Even if the performance takes place in a conscious way, with attention to the technological stratification in terms of the logic of performance, it may happen that it has not "constructive effectiveness". The graphical representation may have no correspondence with the needs the construction site and the specific production process.

The methodological procedure involves a preliminary choice of construction details of the building. These technological details must be meaningful to their difficulty of execution or their repetition within the building structure.

Every selected detail expected in the final design is examined and its "constructability analysis" is processed, according to the logic diagram shown in Figure 5.

Starting from the representation of selected constructive details, their construction can be simulated. This is obtained analyzing them by means of progressive drawings that express and simulate the breakdown of construction phases, chronologically processed and logically related (Figure 6).
3.2 Operative framework

Methodological procedure is performed by means of an encoded design tool described in the follow.

In terms of graphic representation, in view of a likely professional application, an A3 sheet size (according to ISO 216) is used in pursuit of greater manageability in construction site.

The table is divided into three main areas (Figure 8), with a strict logical and sequential concatenation according to the scheme earlier shown in Figure 6, but with a free graphic expression.

1. **Graphic representation of the construction detail and the specific phase of work to accomplish.**

   It must be the predominant element in the table. In terms of communication, what you are referring to it has to be immediately understood. Compared to a traditional graphical representation, it has to provide a large range of information relating to construction needs. To this purpose, it may be useful to integrate the planimetric detail with diagrams and section to clearly contextualize it within the building, or possibly with images - in case you are in the presence of a required conversion of an existing building.

   It has to be represented:
   - **the position of operators**: workers can be summarized with symbols in the form of a dimensionless logo to simplify the engagement representative;
   - **the presence of machinery**: machines can be summarized, as workers, with symbols in the form of dimensionless logo;
   - **the presence of temporary structures**: it’s important to have scale draw of the temporary structures because of their physical interaction and interference with the building structure under construction.

The following color code is used (Figure 8):
- **Black** for the parts of the building already constructed;
- **Red** for the analyzed construction phase;
- **Blue** for temporary works necessary to the workers to reach the correct workplace;
- **Green** for the temporary works of structural support.

The dynamic variation of colors from a table to another expresses the evolution of the construction.
2. **Description of the analyzed working phase with the specific type of manpower, equipment and machinery needed.**
   
   This part, complementary to the first one, integrates and explicates all the construction information of the analyzed working phase (Figure 8).

3. **Risk assessment and identification of individual and collective measures to prevent and protect.**
   
   This section provides an analysis of the data explained by the graphical representation. We highlight the critical issues that come from the constructive simulation of the specific phase, both in terms of building site settlement and in terms of worker safety. This section also explicit critical issues due to the interaction "man-machine", "man-provisional structure", "man-structure" (Figure 8).

---

**Figure 7.** Tables drawn according to the Operative Framework and connected as in Operative Flow Chart in figure 6 – Example of roof redevelopment
3.3 First applications and results

During last years, we have applied the proposed design and assessment tool for the building process in a wide range of situation (Figure 9). The followed criteria in choosing the case studies were mainly regarding:

- **Typology of building intervention.**
  The tool can be adapted for each typology of intervention, starting from the maintenance till coming to the new construction, passing through the building redevelopment.

- **Structural system typology.**
  We have some invariant steps according the specific structural technology (i.e. concrete frame buildings, wooden structures, steel structures).

- **Typology of Graphical representation.**
  It is necessary to choose the more appropriate way to represent with effectiveness the building we deal with. We performed 2D and 3D, in order to make the representation the most communicative we could.

- **Type of workers.**
  We tried to modify the tool’s language adapting it to the final users, which are nothing else that workers. Multilingual versions of the tables were produced, as well as representations dedicated to self-builders.

By the linear combination of the above set of criteria, hundreds of tables were produced on the base of real case studies. Some of the produced tables directly influenced the construction process, as well as they suggested different technological solutions to the designer, with the aim to make the building consistent with the design idea.

Figure 9. Diagram showing the possible applications
4 Conclusion & Discussion

At first, an attempt was made to assimilate this process in some of the planning security tools provided by European and Italian law. The instrument that seems more adaptable is Safety and Coordination Plan (PSC, as defined in Italian law). In experimental applications performed in the last few years, we have used the definition of “Graphic PSC”. We realized however, that this definition was too restrictive: the proposed method it is not only a graphic representation but is based on a theoretical approach that leads to a constructability and safety performance based design.

The final product of our design procedure is an assessment tool for the building process.

The method allows both to perform the specific and detailed safety assessment for the realization of construction detail components, and to verify the constructability of them.

The safety and constructability analysis is firmly performed in an encoded graphical way during the design process, not waiting the construction phase. So we are still able to intervene, if necessary, with amendments under the designer control.

At the moment we are working to implement the tool in a software application. If we apply the proposed design to a wide database, by the help of a BIM software [13], we will be able to automate the procedure to perform the constructability and safety performance based design exactly during the design phase.

References


Abstract -
Tower crane layout design and planning within construction site is a common construction technical issue and regarded as a complex combinatorial problem. To transport heavy materials, such as rebar, formwork, scaffolding, equipment and steel, tower cranes are needed and should be well located to reduce construction cost and improve safety management. Currently, practitioners in the industry are over-reliance on individual experience and subjective judgment during decision-making process. The purpose of this paper wants to develop a well-defined approach, which integrating Building Information Modelling (BIM) and firefly algorithm to come up with an optimal tower crane layout for construction projects. Firstly, BIM technology is utilized to automatically generate the quantity of materials which need to be transported. Then firefly algorithms are used to determine the locations of tower cranes, supply points and demand points according to transportation requirement, time and cost. Thirdly, the optimal tower crane layout scheme will be visualized by 4-Dimension (4D) BIM to verify its constructability and safety based on computer simulation and individual experience. Finally, a practical case is selected to evaluate the developed approach. In addition, some lessons learned and issues are highlighted that help direct future research and implementation. The optimization results of the example are very promising and it demonstrates the application value of the approach.

Keywords -
Tower crane; Firefly algorithm; Building Information Modelling (BIM); Site layout

1 Introduction
Tower crane layout design and planning within construction site is a common construction technical issue and regarded as a complex combinatorial problem. To transport heavy materials, such as rebar, formwork, scaffolding, equipment and steel, tower cranes are needed and should be well located to reduce construction cost and improve safety management. Affected by many uncertainties (variables) and variations, tower crane layout planning is a typical multi-objective problem. To facilitate the decision-making process for this problem, Tem et al proposed a non-structural fuzzy decision support system, which integrated both experts’ judgment and computer decision modelling, to solve the appraisal of complicated construction problems. The system allowed assessments based on pairwise comparisons of alternatives using semantic operators that can provide a reliable assessment result even under the condition of insufficient precise information [1]. In order to reduce conflicts between groups of tower cranes, Irizarry and Karan Hence developed an integrated Geographic Information Systems- Building Information Modelling (GIS-BIM) model to effectively identify the feasible locations for defined tower cranes [2]. Tam and Tong used artificial neural networks to model the non-linear tower crane operations, and developed genetic algorithms to determine the locations of the tower crane, supply points and demand points by optimizing the transportation time and costs [3]. Ju and Choo proposed a parameterized super element formulation for modelling the multiple-pulley cables in a crane system based on the friction-free assumption between the cable and pulleys. These cable passages had significant effect on both static tensions and dynamic properties of tower cranes and the proposed super element model provided more accurate and realistic results in dynamic analyses of the crane system [4]. Sacks et al developed an approach for unique association of isolated equipment operations with planned construction activities based on comparison of the values of various characteristics, calculated for each equipment operation, against preset filters of characteristic values for all expected basic construction activities [5]. Duong et al concerned with
the control of an under-actuated three-dimensional tower crane system using a recurrent neural network (RNN), which is evolved by an evolutionary algorithm [6]. Ning et al used a continuous dynamic searching scheme to guide the max-min ant system so as to solve the dynamic tower crane layout problem under the two congruent objective functions of minimizing safety concerns and reducing construction cost [7]. In order to eliminate collisions between mobile crane and onsite facilities, virtual cranes were provided with motion-planning algorithms that enable them to find collision-free and time-efficient paths for each piece that needs to be erected [8].

Efficient material transportation plays an important role in reducing costs and time. Tam et al had developed traditional linear regression models and nonlinear neural network models for predicting hoisting times of a tower crane [9]. To transport heavy materials, tower cranes are needed and should be well located to reduce operating costs and improve overall efficiency. Huang et al had developed Quadratic assignment problem to simulate the material transportation procedure [10]. Leung et al had developed a quantitative model for predicting the hoisting times of tower cranes for public housing construction using artificial neural network and multiple regression analysis [11]. Input shaping is an effective method for reducing motion-induced vibration. Blackburn et al. had investigated the effect of nonlinear crane dynamics on the performance of input shaping. They proposed nonlinear equations of motion and novel command-shaping algorithms for reducing vibration during the nonlinear slewing motions of tower cranes [12]. Aspiring to adopt a non-statistical quantitative approach to safety assessment due to the operation of tower cranes, Shapira and Simcha implemented a multi-attribute decision-making tool to elicit knowledge from experts and formalize it into a set of weighted safety factors [13].

Visual monitoring and alarming is an effective way to avoid accidents and to increase the efficiency of project management in the construction industry [14]. BIM is an emerging technology in construction industry and have been implemented in site layout visualization and optimization. Al-Hussein et al presented a practical methodology for integrating 3D visualization with special problem simulation for tower crane operation [14]. Lee et al proposed a robotic tower-crane system with a laser device, an encoder, and an accelerometer, and tested the feasibility of the system under indoor, outdoor, and swinging conditions [15]. Huang et al developed Construction Virtual Prototyping (CVP) system to assess the executability of a construction planning including site layout, temporary work design, as well as resource planning [16]. Hwang developed a collision-prevention approach, which comprised of real-time data collection platform, a visualization module, and a decision module, to real-time monitor equipment operations and assess the possibility of collision [17]. Yang et al demonstrated the use of a surveillance camera for assessing tower crane activities during the course of a workday. In particular, it seeks to demonstrate that the crane jib trajectory, together with known information regarding the site plans, provides sufficient information to infer the activity states of the crane [18]. Kang et al had developed a system for providing detailed planning and visualization in a virtual construction environment as well as for assisting crane operators in real-time during erection. [19]. Tower crane operators often operate a tower crane with blind spots. To solve this problem, video camera systems and anti-collision systems are often deployed. Lee et al introduced a newly developed tower crane navigation system that provides three-dimensional information about the building and surroundings and the position of the lifted object in real time using various sensors and a building information model [20]. However, due to the low transmission rate of a wireless network, sometimes it is not feasible to use a camera in a remote monitoring and alarming system for a tower crane [21]. Li and Liu developed a data-driven remote monitoring and alarming system for tower cranes, which integrated field data and 3D simulation. In addition, an experience-based and practical alarming system was embedded into the virtual monitoring system [21].

The purpose of this paper wants to develop a well-defined approach, which integrating Building Information Modelling (BIM) and firefly algorithm to come up with an optimal tower crane layout for construction projects. Firstly, BIM technology is utilized to automatically generate the quantity of materials which need to be transported. Then firefly algorithms are used to determine the locations of tower cranes, supply points and demand points according to transportation requirement, time and cost. Thirdly, the optimal tower crane layout scheme will be visualized by 4-Dimension (4D) BIM to verify its constructability and safety based on computer simulation and individual experience. Finally, a practical case is selected to evaluate the developed approach.

2 The Framework of Integrating BIM and Firefly Algorithm to Optimize Tower Crane Layout

This section describes a framework of integrating BIM and firefly algorithm to optimize tower crane layout (as shown in Figure 1) which includes four steps: creating BIM model, determining minimal number of tower cranes, determining optimal location of each tower
crane, and BIM-based visualization and validation. Each is described in detail below.

Figure 1. The framework of integrating BIM and firefly algorithm to optimize tower crane layout

2.1 Step 1: Creating BIM Model

In this step, there are two main tasks need to be implemented. The first one is calculating the required number of tower cranes based on lifting material quantity, construction schedule and tower crane types. For example, the total lifting quantity of a building is 28000 t, construction duration is 300 days, tower crane load is 1.5 t per each lifting times, and each tower crane can lift 32 times per day, we can conclude the required number of tower cranes is 2 (the calculation method is 28000/(1.5*32*300)=1.94). The second one is assessing the requirement number of tower cranes based on feasible areas for locating tower cranes. When we get the results of the two tasks, the large value is the minimal number of tower cranes.

2.3 Determining Optimal Location of Each Tower Crane

In this step, there are three tasks need to be performed.

2.3.1 Problem description

For establishing a mathematical model for the location of cranes group, some assumptions are listed below [22]:
(1) Geometric layout of all supply (S) and demand (D) points, together with the type and number of cranes, are predetermined.
(2) For each S-D pair, demand levels for transportation are known, e.g., total number of lifts, number of lifts for each batch, maximum load, unloading delays, and so on.
(3) The duration of construction is broadly similar over the working areas.
(4) The material transported between an S-D pair is handled by one crane only.

2.3.2 Mathematical model

The completed model consists of two sub-models: minimal conflicts between cranes, and a single-tower-crane optimization model for optimal location in terms of minimal hook transportation time.

To measure possibility of conflict, a parameter (NC), called the conflict index is introduced. The transportation of a crane corresponds to a triangle with apexes representing the supply point, demand point, and crane location. The number of intersections between two triangles reflects the severity of conflicts, i.e., the more intersections the more likely are conflicts. Additionally, the intensity of material flows also affects possibility of conflicts. Hence, conflicts between cranes i and k can be represented as

\[ NC_{ik} = \sum_{i=1}^{L} \sum_{j=1}^{I} n_{ij,kl}(Q_{ij} + Q_{kl}) \]

where \( n_{ij,kl} \) defines the number of intersections of the two triangles, respectively, consisting of crane i and task \( f_j \), crane k and task l. \( Q_{ij} \), the number of lift of j-th tasks in i-th task groups. \( Q_{kl} \), the number of lift of l-th tasks in k-th task groups.

Furthermore, the possibility of conflicts between two crane task pairs should be proportionate to \( n_{ij,kl}(Q_{ij} + Q_{kl}) \). Therefore, for all cranes and all tasks, the conflict index (reflecting general possibility of conflicts) can now be calculated as

\[ \min f_c(x) = \min NC_{ik} \]
On the other hand, if \( \{Dx_j, Dy_j, Dz_j\} \) and \( \{Sx_j, Sy_j, Sz_j\} \) refer, respectively, to the location of \( S \) and \( D \) of a task, and \( (x_i, y_i) \) denotes location of foundation of a crane. The travel time of \( i \)-th hook for task \( j \) \( T_{ij} \) can be expressed as

\[
T_{ij} = T(D_j', S_i) + T(S_i, D_j) + U(D_j)
\]

where \( i = 1, 2 \ldots l; j = 1, 2, \ldots, J \), and

\( T(D_j', S_i) \), hook travel time without loads from \( D \) of task \( j' \) (produced by last request) to \( S \) of present request \( j \);

\( N_k^j \), the repeated number of task \( j \) in batch \( k \),

\( T(S_i, D_j) \), hook travel time with loads from \( S \) to \( D \);

\( T(D_j, S_i) \), hook travel time without loads from \( D \) to \( S \);

\( L(S_i) \), hook delay time for loading at \( S \);

\( U(D_j) \), hook delay time for unloading at \( D_j \).

Hence, the minimal average transportation time for all crane hooks is

\[
\min f_2(x) = \min \text{ATT}(x, y) = \frac{1}{K} \min \sum_{i=1}^{l} \sum_{j=1}^{J} T_{ij}
\]

(2)

2.3.3 Firefly Algorithm for the Problem

Firefly algorithm is a novel nature-inspired algorithm inspired by social behaviour of fireflies. Fireflies are one of the most special, captivating and fascinating creature in the nature. By idealizing some of the flashing characteristics of fireflies, firefly-inspired algorithm was presented by Yang [32].

Firefly-inspired algorithms use the following three idealized rules:

1. All fireflies are unisex which means that they are attracted to other fireflies regardless of their sex;
2. The degree of the attractiveness of a firefly is proportion to its brightness, thus for any two flashing fireflies, the less brighter one will move towards the brighter one and the more brightness means the less distance between two fireflies. If there is no brighter one than a particular firefly, it will move randomly;
3. The brightness of a firefly is determined by the value of the objective function. For a maximization problem, the brightness can be proportional to the value of the objective function.

In the firefly algorithm, there are four important issues: Attractiveness: In the firefly algorithm, the main form of attractiveness function \( \beta(r) \) can be any monotonically decreasing functions such as the following generalized form:

\[
\beta(r) = \beta_0 e^{-\gamma r^n}, \quad n \geq 1
\]

(3)

where \( r \) is the distance between two fireflies, \( \beta_0 \) is the
attractiveness at \( r = 0 \) and \( \gamma \) is a fixed light absorption coefficient.

**Distance:** The distance between any two fireflies \( i \) and \( j \) at \( x_i \) and \( x_j \) is the Cartesian distance as follows:

\[
r_{ij} = \| x_i - x_j \| = \sqrt{\sum_{k=1}^{d} (x_{ik} - x_{jk})^2}
\]

where \( x_{ik} \) is the \( k \)-th component of the \( i \)-th firefly.

**Movement:** The movement of a firefly, \( i \), is attracted to another more attractive (brighter) firefly \( j \), is determined by

\[
\bar{x}_i = x_i + \beta(t) \times (x_i - x_j) + \alpha \left( r - \frac{1}{2} \right)
\]

where the second term is due to the attraction while the third term is randomization with \( \sigma < 1 \) being the randomization parameter and \( r \) is a random number generator uniformly distributed in \([0, 1] \). In this paper, an outline update mode is applied in original FA, that means the light intensity is updated and the new position of firefly is evaluated after all the firefly finish moving.

The steps of the FA can be summarized as the pseudo code shown below.

---

**Stop if**

- \( t < t_{\text{max}} \) \( \% \% \) \( t_{\text{max}} \) is maximum generation
  - \( i = 1 : n \) \( \% \% \) all \( n \) fireflies
    - \( j = 1 : i \)
      - If \( (L_{ij} > L_{ij}) \)
        - Move firefly \( i \) towards \( j \) in \( d \)-dimension by equation (5)
  - Attractiveness varies by equation (3)
- Evaluate the new position of \( i \)-th firefly and update light intensity \( L_{ij} \), \% outline update
- Rank the fireflies and find the current best

**End**

---

2.4 **BIM-based Visualization and Validation**

The optimal layout scheme is calculated based on ideal construction environment, some factors which cannot be quantified are not considered in the mathematical model. In order to assure the constructability of the optimal tower crane layout scheme, BIM is used to conduct assessment based on construction expertise and experience of project participants. In addition, 3D visualized construction scheme of tower cranes will be generated so as to guide field workers install and dismantle tower cranes accurately.

3 **Case Study**

Case A is selected to validate the proposed method. This project includes 6 high-rise business buildings, 21 floors on the ground and 3 floors underground. The gross floor area is 420,000 m² and occupied area is 40,000 m².

3.1 **Determining the Minimal Number of Tower Cranes**

In this section, firstly the required number of tower cranes is calculated based on information from BIM model. The following information demonstrates the detailed calculate process.

1. Calculating rebar lifting times
   - The rebar quantity of rebar in a standard level is about 100kg/m²x2000 m²=200 tons. Based on the tower crane load, which is 1.5 tons per lifting, the rebar lifting times in a standard level is: 200/1.5=133.

2. Calculating formwork lifting times
   - There are about 40 frame columns in standard level, tower crane can life one frame column’s formwork at a time because the formwork weight is in the range of tower crane load, so the lifting times of frame column’s formwork is 40.

   - The length of shear wall in standard level is 100 meter and the length of single formwork will be 3.3 meter on account of the higher story height. So the lifting times of the shear wall formwork in standard level is 100m/3.3m=30. Since there are more corner walls in this project, the lifting times is adjusted to 30x1.5=45.

   - The beam and slab area in standard level is about 2,100 m². Based on the tower crane load, which is 90 m³ per lifting in consideration of the formwork stack height. Hence, the lifting times of beam and slab formwork in a standard level is: 2100/90=24.

   - The estimation amount of beam and slab formwork support system is 100 tons. The average lifting weight of tower crane is 1.5 tons at a time, so the lifting times of beam and slab formwork support system is: 100/1.5=67.

   - Synthesizes all types of formwork erection and dismantle, the total lifting times of formwork and support system in standard level is: (40+45+24+67)*2=352.

3. Calculating other materials lifting times
   - Lifting times of scaffolding steel pipes: 15
   - Lifting times of door and window formwork: 15

---
Lifting times of embedded water and electricity pipes: 10

(4) Calculating the total lifting times
The total lifting times in a standard level is:
\[135 + 352 + 15 + 15 + 10 = 527\]

(5) Calculating the required number of tower cranes
Based on the construction experience and tower crane specification, each tower crane can lift 4 times per hour and 10 work hours per day. In addition, project duration is 9 days per a standard level. Hence, the required number of tower cranes is: ceiling \[\frac{527}{4\times9\times10}\] = 2. Because there are 6 similar buildings in this project, at least 12 tower cranes are needed in this project.

Secondly, according to feasible areas for locating tower cranes according to geometric layout of supply and demand points, the require number of tower cranes is 12. When comparing the two results, the large value (i.e.12) is the minimal number of tower cranes.

3.2 Determining the Optimal Location of Each Tower Cranes

Figure 4 shows the optimal tower crane layout scheme, each location of tower crane is calculated based on the mathematical model developed in section 2.3.

Figure 4. The optimal tower crane layout scheme

3.3 BIM-based Visualization and Validation

The optimal tower crane layout scheme is calculated based on a static and hypothetical scenario. In this project, the installation of each tower crane is divided into five times. BIM is used to calculate the installation height of each tower crane at a time in order to reduce collision. Tower cranes installation heights in stage one are shown in figure 5 and table 1. Tower cranes installation heights in stage two are shown in figure 6 and table 2. Tower cranes installation heights in stage three are shown in figure 7 and table 3. Tower cranes installation heights in stage four are shown in figure 8 and table 4. Tower cranes installation heights in stage five are shown in figure 9 and table 5.
This paper has developed a well-defined approach, which integrates BIM and firefly algorithm to come up with an optimal tower crane layout for construction projects. The results show that: (1) less time is needed to create a tower crane layout scheme compared with traditional methods, especially in multiple tower cranes layout; (2) the optimal tower crane layout scheme generated by the proposed method is better than the original scheme in less total material transportation cost and collisions; and (3) field workers can understand and perform the tower crane layout scheme easily and accurately due to its visualization and interaction.

References


A Model for Construction Contractor Selection Using Competitive Intelligence (CI)

Mahdi Safa a, Arash Shahi b, Carl T. Haas c, Majeed Safa d, Keith Hipel e Sandra MacGillivray f, Dawn Fiander-McCann g

a,b,c Department of Civil and Environmental Engineering, University of Waterloo, Canada

d Department of Agricultural Management and Property Studies, Lincoln University, New Zealand

e Department of System Design Engineering, University of Waterloo, Canada

f,g Valency Inc., Canada

E-mail: msafa@uwaterloo.ca, arash.shahi@uwaterloo.ca, chaas@uwaterloo.ca, majeed.safa@lincoln.ac.nz, kwhipel@uwaterloo.ca, sandra@valencyinc.com, dawn@valencyinc.com

Abstract: While comprehensive and ongoing competitive intelligence (CI) is employed in a variety of industries to provide valuable input for broad strategic decisions, the construction industry lags behind in adopting this technique. This paper presents a CI model for use in the construction contractor selection process, which is a critical element of construction project management and one that inherently entails risk and risk management. The use of CI for contractor selection is an important development in light of the realization on the part of many companies that the diffuse nature of the information and lack of robust analysis create numerous inconveniences during the decision-making process. Based on the application of the CI method for a competitive environment, the proposed model has the potential to improve the process for assessing and selecting contractors. This paper describes the proposed model, including background information, structural details, guidelines for its use and implementation, and key data analysis findings.

Keywords: Construction management, Competitive Intelligence (CI), Megaproject, Contractor selection, Risk

1 Introduction

Current trends suggest that construction project management may be well “behind the curve” in effectively applying the competitive intelligence (CI) approach. Construction project management is a challenging and complex process involving coordination of many tasks and multiple parties (such as consultants and contractors) with different priorities and objectives. Effective decision making approaches for construction projects require the deployment of various strategies, tactics, and tools. The selection of a contractor is one of the critical and strategic decisions that need to be made on a continuous basis during the life of a project. Hence, the main purpose of introducing a CI model in this context is to support decision makers with the contractor selection process. Since competitive intelligence (CI) has become more important to a firm’s knowledge development and decision making efforts [1, 2], decision makers as the CI professionals must play an active role in the selection of contractors.

Several CI models and programs are in use in other industries, but the theories, methods, and results related to these industries cannot always be applied directly to the construction industry with the same level of success. The scope of this research focuses specifically on the industrial sector projects in the construction industry, but the concepts and models developed are applicable to other construction industry sectors. In general, the industrial construction sector involves projects such as power plants, refineries, and process facilities. It is characterized by a high concentration of participants (contractors, owners, etc.) and by a high level of engineering and project management sophistication. The research presented in this study concentrates on industrial construction and, in particular, on megaprojects.

Decisions made during the bidding process are concentrated at the managerial level, the point at which public officials and designated decision makers have the power to accept or reject a contractor for a specific project or its subprojects.

The contract is one of the most important parts of the bidding process. To accomplish well-executed projects, we must have knowledge about the contract management and contractor selection process and the
best ways to manage contracts more effectively. The contractor selection as well as many other multi-criteria decisions impacting the overall project should be made during the front end planning (FEP) of the projects. The bulk of the project costs, the major risks, and contractor selection strategy are defined during the FEP stage [3].

Therefore, using appropriate CI tools to facilitate informed decisions early in the process is critical in making sure that measurable improvements can be realized with respect to the contractor selection. It should be noted that the ethical and legal risks of the selection process can be minimized by establishing clarity regarding the legal ways to gather information and to interact with source of information [4]. This approach builds a trust relationship and creates transparency with the contractors’ communities [4].

There is no standard and universal definition for CI as experts and scholars with different background and experience have different views on CI. Blenkhorn and Fleisher (2013) defined CI as “the ethical and legal process of discovering, analysing, and delivering intelligence from publicly available, non-proprietary, and proprietary information sources for the purpose of becoming more competitive in the marketplace”[5]. In other words, “CI is a necessary, ethical discipline for decision making based on understanding the competitive environment”[6]. This information is about potential contractor’ abilities and desires to assist project teams in making the correct strategic decisions. It should be noted that, it is critical that project teams have access to evolving knowledge and instruction in the field for decision makers to remain proficient in competitive intelligence [7].

2 Construction Contract Management

Challenges

In large construction projects, optimally dividing the work among contractors is challenging. This process is typically executed during the pre-construction phase by project leadership team, who rely on their experience and judgment, making it difficult to demonstrate that the results are optimal or to defend the decision making rationale after the fact. Addressing these problems requires an auditable and robust method that still incorporates the expertise of the project leadership team.

The scope of knowledge required for conducting the selection phase is influenced by the type of project and the job situation [8]. The development of the CI model therefore includes consideration of the size, location, and schedule of the project as well as any cash flow constraints, the owner’s philosophy, and the overall project strategies.

Significant changes, learning, adaptation, and growth are inevitable when a CI is implemented. As mentioned earlier, the evaluation of the contractor is a vital component of this system and is related to risk and risk management [9-11]. During the bidding and negotiating process, selecting the best contractors with respect to the relevant CI practice is highly critical for the overall success of the project. To select the best contractors and to prepare the most realistic and accurate bid proposals, the experts must be aware of all financial, technical, organizational culture and general information about the projects.

During the contract management process, a significant risk is ineffective partnership strength, which occurs when the relationships between managers (decision makers) and contractors are too personal. Unfortunately, a large part of decisions are affected by such connections. While this kind of personal selection practice has associated benefits, such as stability, mutual trust, and reduced transaction and search costs, it also entails a number of challenges. For the development of the system during this phase, a critical recommendation is the avoidance of any unsystematic CI tendencies and reliance. Instead, emphasis will be on a systematic method for designating a contractor. However, the managers ultimately make the final choices, while the CI model results approach the ideal optimal decisions.

The model developed in this research was created based on interviews with the construction project managers who are in charge of numerous mega projects in North America. The selection of the best contractor for each work package influences not only the success of the construction project but also the quality of the results produced by any model based on those decisions as outputs. In particular, during the bidding process, the optimal selection of the contractors is vital because it results in an accurate and realistic bid proposal. During this phase decisions are focused at the managerial level, the point at which public officials and designated decision makers have the power to accept or reject a contractor for a specific construction project based on management-level considerations.

An evaluation project team (CI professionals) is frequently involved in the contract management process. Five typical criteria are usually considered, four of which are quantitative, and one of which has a qualitative value: (1) cost, (2) time, (3) field service and engineering rates, (4) experience, and (5) the financial stability of the contractor. However, a comprehensive CI model is derived from a complex and creative process. The model should incorporate a wide range of activities, elements and attributes, including:
1. Using news aggregators, databases and journal subscriptions (Press analysis)
2. Pricing research
3. Interviews with potential contractors
4. Interviews with experts
5. Historical database analysis
6. Governmental and publicly held records
7. Avoiding surprises and identifying threats and opportunities
8. Understanding where a company is vulnerable and may have decreased reaction time

3 A Guide for Designing a CI Model for the Contractor Selection Process

CI is aimed at promoting the success of a construction firm by improving the use of information in a company and also the use of external sources. The performance of a construction firm depends much on the performance of its contractors and subcontractors during the different phases of a project. John E. Prescott provided a framework for designing a CI model that has been exercised for this study [12]. The design of a CI program for the proposed model in this study involves five key decision domains. The following table provides a summary of these decision domains.

<table>
<thead>
<tr>
<th>#</th>
<th>Decision Domain</th>
<th>Key Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CI Projects</td>
<td>Project-based approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Focus on decisions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prioritize intelligence needs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CI team</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Try a demonstration project</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pitfalls</td>
</tr>
<tr>
<td>2</td>
<td>CI Products</td>
<td>Timely, Accurate, Relevant, etc.</td>
</tr>
<tr>
<td>3</td>
<td>CI Achievements</td>
<td>Early warning of opportunities and threats</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strategic decision making support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tactical decision making support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Competitor monitoring and assessment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strategic planning support</td>
</tr>
<tr>
<td>4</td>
<td>CI Ethics</td>
<td>Develop a code of ethics</td>
</tr>
</tbody>
</table>

Decision domain 1: Projects are the foundation of a CI model. Construction as a whole is usually dominated by “one-off” projects [13-15] so, each project is unique. The CI model helps the managerial team in gathering, analysing, and applying information about projects, potential contractors, regulators, partners, and clients for the short- and long-term planning needs of a firm [16]. The results of the intelligence audit through selecting the best practices could pose the benefits of CI.

Decision domain 2: the outputs of the CI model must have some value to the management team. Creativity is a valuable guide for CI professionals who are developing CI products.

Decision domain 3: the CI model needs to have an achievement such as identifying opportunities and threats, support for strategic and tactical decision making, and competitive monitoring and assessment.

Decision domain 4: Ethics is one of the most important topics of the CI field. The majority of ethical problems have concentrated on the methods used in the collection of information. In this case, the contract selection process should be an ethical instrument to provide a formalization of commitments, enforceable by the company’s principle. Construction firms’ concerns over reputational risk have led to steady growth of due diligence in researching and selecting contractors in ethical ways. The contract selection process therefore involves an increasing focus on ethical issues.

Decision domain 5: A CI department can be located anywhere in the organization. A standardized work process increases the efficiency of the model. The use of a pre-configured workflow also increases the rigor that is applied to the work process of design reviews and equips the engineering team to accurately track the resolution of comments regarding nonconforming items. The review indicated that someone in the organization should be recognized as, the CI manager (champion). This person is the focal point for the CI effort. There are potentially three additional roles for individuals assisting in the CI effort. Each of the roles requires different skills, and in some cases, training. The first role conducts the human intelligence (the capacity of the contractor to connect with the owner) network coordination. A second role involves the collection of...
secondary information through information technology. The third role is the analyst. Analysts convert information into intelligence. The analyst needs to develop skills in a variety of areas including forecasting, profiling, financial analysis, and statistics.

The following section explains how a number of analytical tools could be applied for the analysis approach employed in the CI model. Specifically, the SWOT, Porter’s five forces, and the PEST methods are introduced using a major contractor as a case-study. In this paper, authors have selected Company “A” (contractor) as an example of a globally recognized and leading Canadian engineering and construction contractor. Only public data and information have been used for this analysis, and they are presented in the next section prior to the introduction of the CI models.

4 Current State of the Canadian Construction Industry

The construction industry is diverse and significantly large in Canada, currently a $171 billion industry and divided into residential sector and non-residential sector [17]. It is strong and constantly evolving, consumes 40% of the country’s energy and 50% of Canada’s primary resources. It employs 1.24 million people, which is about 7% of Canada’s total workforce [18]. Between the years 2002 and 2011, the annual GDP of the construction industry was growing at a compound annual rate of 3.2% [19].

However, industry growth is projected to slow down slightly during the next decade mainly due to slower activity in the residential sector [20]. There will be an increased activity in non-residential sector, since the Canadian federal government has approved numerous large infrastructure projects worth approximately 200 billion dollars [21]. Another trend in this sector is investment in “greening” of the construction industry.

It is estimated that about 100,000 workers will need to be added to the current number by 2020, which would be about a 10% rise from current level [22]. The construction industry in Canada is comprised of about 250,000 firms, which are usually very small. About 90% of companies in the residential category have less than 5 employees, and only 1% has more than 50 employees. In the non-residential sector, about 70% of firms have 5 or less employees [23]. Some of the largest firms in this industry by revenue are listed in the Table 2.

Table 2 Largest firms in construction industry by revenue [24]

<table>
<thead>
<tr>
<th>Company</th>
<th>2012 Revenue in billion</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$8.1</td>
</tr>
<tr>
<td>B</td>
<td>$6.9</td>
</tr>
<tr>
<td>C</td>
<td>$3.1</td>
</tr>
<tr>
<td>D</td>
<td>$2.9</td>
</tr>
<tr>
<td>E</td>
<td>$2.2</td>
</tr>
<tr>
<td>F</td>
<td>$1.5</td>
</tr>
<tr>
<td>G</td>
<td>$1.4</td>
</tr>
</tbody>
</table>

These large and other long existing firms have an advantage over new potential entrants, since they have established reputation, experience and stable relationships with subcontractors, suppliers and financial institutions. New entrants are at a big disadvantage, since contracts are usually obtained based on long time experience, reputation and overall capabilities of the firms. At the global level, Canada is expected to move from the seventh to fifth largest construction market in the world by year 2020 [25].

At this moment, the global construction sector is worth approximately US $7.5 trillion [26]. The construction industry closely follows the overall economy growth numbers and there is a 4.5% increase in construction industry predicted for this year. While Europe seems to be stagnant with little current or short-term increases, the emerging regions in developing countries around the world will likely be increasing their activities in construction industry in the next decade [27, 28].

5 Company A

Company A is the largest engineering and construction company in Canada and one of the top five design firms in the world. As one of the Canadian construction industry leaders, A has expertise in the following areas: agriculture and rural development, environment and water management, infrastructure, mass transit, mining
and metallurgy, oil and gas, operations, maintenance, pharmaceuticals, biotechnology, power, procurement and telecommunications. Currently, the Canadian market generates 62% of its revenue. The remaining 38% of revenue comes from Latin America, Europe, Africa, Middle East, United States, Asia-Pacific with individual contributions ranging from 1% to 10%. Revenue by industry segment is distributed more uniformly comparing to revenue by geographical area. It ranges from 26% for Infrastructure and Environment all the way to 5% for other industries [29].

A has slipped 19 notches on the Top International Ranking from 2012 to 2013 [30]. A’s performance has suffered some major blows. The company CEO attributes this drop to fixed-price contracts being detrimentally impacted by a mining sector slowdown. Another view of this same scenario comes from industry analysts who claim some of the company’s activities have become controversial. Rumours of A’s suspect bidding practices and political involvement may well have played a part in 2013’s net income drop to 10-50 million, down from the 355 million initially forecast [31].

5.1 Future of Company A

For the short term, the company hired a financial advisor to advise it on selling equity stake in its power transmission business. The company stated that all options were open for the unit, including a private sale or a strategic partnership. For year 2014, financial results will still remain negatively affected by the troubled contracts and rumours of questionable business dealings. Numerous projects will be facing cost reforecasts - even the company’s latest contracts in the hospital and road sectors - which have not been previously the subjects of revisions. A expressed its hopes that the reforecasts will only be one-time events that will further not affect future profitability. In addition, the company plans to take a separate big-bath charge of $75-million to reorganize its European operations, which are in disarray in a number of countries, notably in France. Moreover, by the year 2015, the majority of the fixed-price contracts will be expired and the negative effect over the financial results will be eliminated [32].

Beyond 2015, A will target growth in key engineering and construction markets with a focus on mining and resources like oil and gas and it will attempt to change its revenue mix to increase the contribution made by service contracts. A’s plan is to enhance its project financing capabilities to strengthen its competitiveness in major projects and infrastructure concessions [32].

5.2 Current Strategy

Current focus of the Company A is on infrastructure and clean power, and specifically on high growth and high margin sectors, building on its strengths. It is a market leader in Canada for renewable energy and it strives to keep it that way. In terms of infrastructure, Company A is focused on the growing transportation sector in North America and its strategy is to leverage experience they have with building airports, bridges, roads and highways, since such projects will remain a good source of sustainable growth. Being a diverse global company, Company A has a specific future strategy for every region. For example, it sees Africa as a strategic long-term growth area and Canada as a key region to grow market share.

In the future, Company A will be focused on three distinct industry areas: oil and gas, mining and environment and water. For the oil and gas area, it is looking to compete in this high margin business at a global level. It is hoping to build on its existing platform, expertise and know-how in this area. With respect to the mining industry, Company A is hoping to grow and solidify its tier 1 position and establish long-term profitability. Although it is aware that it scales well in this industry, it is looking to make further progress.

Finally, Company A is also contemplating participating in the environment and water market also on a global scale. It claims to have multiple opportunities in this area available, and have synergies it can build on with the oil and gas sector. Furthermore Company A is looking to better integrate and share resources, and promote greater interaction between business units, especially relaying on the newly appointed COO [33].

5.3 Analysis of Company A

The applications of the strategic analysing tools are introduced in this section. All these tools use the historical information and data for inducing future assumptions.

5.3.1 SWOT Analysis:

Company A SWOT analysis is a simple strategic tool that facilitates in understanding the strengths, weaknesses, opportunities and threats.

**Strengths:** Company A is the market leader in the Canadian construction industry and a significant international player. Considerable size and extensive experience over a wide range of industries makes it capable of approaching very large projects. It has a strong international business experience and world-class expertise. One subsidiary company Energy Inc. makes it the only company on Canadian market capable to design one type of nuclear reactors.
Weaknesses: In recent years, the company faced reputational issues due to questionable business practices. There is alleged evidence of unethical practices that have been used to win bids for major projects and the former CEO has been charged with fraud. Some locked fixed-price contracts negotiated in the previous years turned out to be costlier than expected. The mining industry, sector where the company is deeply involved, did not grow as expected in the recent years. Combined, these elements may have resulted in a financial decline over the last three years.

Opportunities: The mining industry is expected to recover in the near future with potential for considerable growth. The worldwide increased demand for power capacity can be another source of revenue. Continuously aging world population will increase demand on health care facilities, which is an area of interest and expertise of the Company A.

Threats: As a result of damaged reputation and accusations of illegal business practices, the company and all its subsidiaries are restricted by the World Bank to participate in bids for projects sponsored by the organization. This will restrict the company from winning extremely large international projects. The Canadian nuclear power market does not appear to be active in the near future, so the company will not benefit a lot from its market dominance. Another potential threat could be that the mining sector’s recovery may be slower than anticipated.

5.3.2 Porter’s Five Forces Analysis

Porter’s five forces of competitive position analysis can be used for assessing and evaluating the competitive strength and position of a potential contractor.

Industry Competitors: There are about 250,000 firms in Canada, the majority of which are very small. 90% of these companies have less than five employees and only 1% have 50 or more. There are a small number of large and reputable construction companies that could compete for the largest projects in Canada. Among them are Aecon Group Inc., Graham, PCL Construction, Ellis Don Construction and Bird Construction. This has a positive impact and Company A could expect to potentially obtain many new large and profitable projects. However, the current scandals will have a negative impact on Company A, and other large companies may be given preference.

Threat of New Entrants: Although there are emerging new companies in construction industry, there is a low risk from potential entrants, since a lot of resources and time are necessary to build a strong and reputable construction company. The entry barriers are high as well as the overall risks for new entrants, which is to advantage for Company A.

Suppliers Bargaining Power: Construction materials can be obtained from numerous suppliers and any supplier can be substituted by another at no significant cost, effort or time suggesting that there is a high competition between suppliers to construction companies’ advantage, as prices are driven down. Sub-contractors are also plentiful and the switching costs are low, which has a positive effect for Company A.

Clients Bargaining Power: For the less special and less complex project, the customers of construction companies have many firms available for completing the projects, thus, the customers/buyers have a stronger bargaining power. In general, customers select contractors by the lowest bid. However, the complexities for the selection of firms is less for special and more complex projects, since not all construction firms have the required capabilities. In such cases, the bargaining power is lower.

Substitutes: There are a limited number of large and well-established construction companies that could be selected for the largest projects. However, once the assigned projects are completed, any other large company can be chosen for the new project. A’s most important advantage over others is that it is licensed to design the CANDU reactors. The overall impact on A is positive in that the company has an established reputation and can leverage its capabilities, experience and well-developed project management and it is recognized as a leader in the industry. One major negative impact on the future of Company A has the World Bank’s ban from participating in large projects sponsored by them.

5.3.3 PEST Analysis

PEST analysis is another strategic tool for understanding the political, economic, socio-cultural and technological environment that can be used for evaluating market direction (growth or decline). A has to stay on top of political changes, such as government subsidies for certain projects or initiatives common in this industry and newly introduced environmental legislations. It must stay involved in markets of countries rebuilding from war or natural disasters. In terms of economic factors, Company A has to monitor infrastructure spending by various bodies of government, how the disposable income of families is changing (for involvement in residential sector), and how business spending and government spending are trending.

For the social aspect, the green initiative is a very important emerging trend in the recent years. Low population growth and a retiring workforce may lead to potential employee shortages. The technological factor
is also important in this industry and Company A should ensure it stays involved in developments of new technologies, new sources of energy and other innovations. PEST Analysis is summarized in Table 3.

### Table 3 PEST analysis of A

<table>
<thead>
<tr>
<th>POLITICAL</th>
<th>ECONOMICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>monitor government subsidies</td>
<td>monitor infrastructure spending</td>
</tr>
<tr>
<td>monitor environmental legislation</td>
<td>monitor disposable income of families</td>
</tr>
<tr>
<td>stay involved in rebuilding countries</td>
<td>stay on top of business spending</td>
</tr>
<tr>
<td></td>
<td>stay on top of government spending</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOCIAL</th>
<th>TECHNOLOGICAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>follow green initiatives</td>
<td>keep up with new innovations</td>
</tr>
<tr>
<td>watch population growth</td>
<td>keep up with new energy sources</td>
</tr>
<tr>
<td>consider retiring workforce</td>
<td></td>
</tr>
</tbody>
</table>

### 6 Conclusion

Competitive advantage in a globalized construction market can be gained or lost based on how well a construction firm is able to apply CI principles. The adoption of a rational process for the evaluation and selection of contractors has a direct and positive influence on such a competitive advantage. The proposed CI model should thus be considered a valuable approach for guiding the enhancement of an efficient bidding decision process: its use can decrease the time required for a decision and can generate additional options that may not otherwise be identified. Like other strategic analysis models, the CI model requires the use of strategic analytical tools (analytical methods). In the study presented in this paper, the SWOT, Porter’s five forces, and the PEST methods were applied for analysing A as a sample contractor. The developed model also creates audit trails that can provide explanations of the reasons for decisions. The opinions of industry experts indicate that the model has the potential to become a valuable tool for construction contract management.

### References


[20] Source: Canadian Industry Statistics, On-line:


Safety Management for Existing Buildings in Tunnel Construction

Limao Zhang\textsuperscript{a,b}, Xianguo Wu\textsuperscript{a}, Miroslaw J. Skibniewski\textsuperscript{b,c}, Hongyu Chen\textsuperscript{d}

a. School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074, China
b. Department of Civil & Environmental Engineering, University of Maryland, College Park, MD, 20742-3021, USA
c. Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Poland
d. Wuhan University of Technology, Wuhan, Hubei 430070, China

E-mail: limao_zhang@hotmail.com, wxg0220@126.com, mirek@umd.edu, hy_chen@hotmail.com

Abstract -
This paper develops a systematic approach with detailed step-by-step procedures for safety management of existing buildings adjacent to tunneling excavation. The potential safety risk of a specific nearby building is assessed within four different risk levels, with the spatial neighbor relation (hazard parameter) and the building health condition (vulnerability component) taken into account. Corresponding protective measures for buildings at different risk levels are provided according to risk assessment results. A fine balance between the system safety and cost constrains is reached, where the evaluated risk level plays a decisive role in the adoption of numerous simulation analysis tools. A case concerning the protection of a five-story framed building adjacent to a twin-tunnel in China is utilized to verify the applicability of the proposed approach. The impact of the single and twin tunnel excavation on the soil displacement and building foundation deformation is further analyzed in details. Results demonstrate the feasibility of the proposed approach, as well as its application potential. The proposed safety management approach is also worth popularizing in other similar projects, and can be used to increase the likelihood of a successful project in a complex project environment.

Keywords -
Safety management; adjacent buildings; tunnel construction; case study; numerical analysis

1 Introduction

Due to continuous growth in urbanization worldwide, a large number of new metro tunnels are being constructed or planned for high-speed railways within congested urban areas, especially in developing countries, like China. The tunneling excavation works in the soft ground inevitably lead to ground movements, which may cause adjacent surface buildings to deform, rotate, distort, and possibly sustain unrecoverable damages, especially those founded on shallow foundations [1]. The exploitation of urban underground space presents several geotechnical engineering problems, one of which is the effect of underground tunnel construction on surface and subsurface structures [2]. Damage to buildings adjacent to tunneling excavation can be a major design consideration in tunnel construction because of the challenge regarding the measurement and performance of underground structures [3]. Accordingly, the impact of the tunnel excavation on adjacent buildings is of major interest for tunneling construction in urban areas, due to the high interaction between tunneling and existing structures [4]. Therefore, in order to assure the safety and serviceability of nearby buildings during tunnel construction, it is necessary to explore the safety risk mechanism of the excavation-induced damage to nearby buildings, and propose corresponding preventive measures for adjacent buildings ahead of time [5].

Tunnel-soil-building interaction is considered a highly complicated process, and it is very difficult to rigorously analyze the tunnel-soil-building interaction problem [6]. With the ability to take all relevant factors into account, such as ground heterogeneity, non-linear behavior of soils, soil-structure interaction and construction methods, the finite element method (FEM) proves to be an effective and realistic tool for guaranteeing the safety of tunnel construction [7]. In general, this FEM-based numerous analyses approach provides an effective solution for analyzing the potential construction safety since the complex tunnel-soil-building interaction can be simulated in this approach. However, it can be time consuming and very expensive, since the simulation of the tunneling excavation process can be very slow [8, 9], especially when a large number of existing buildings have to be analyzed.

As a matter of fact, current FEM-based analyses are mainly applied in some specific structures which have important significance, but rarely adapted in general nearby buildings. With the development and utilization of urban underground space, the number of existing buildings adjacent to the construction of metro tunnels is showing an increased growth. For time and cost considerations, it is difficult or nearly impossible to carry out numerous analyses for each adjacent building. To date, most of previous researches have been on the prediction of ground settlement and the tunnel-induced movements on nearby foundation systems. Very few
researchers carried out the overall safety risk analysis and management for nearby buildings in tunnel construction with the cost and project risk taken into account. How to strike a balance between system safety and cost constrains becomes a challenging problem, which falls in the scope of this research interest. A universally accepted standard regarding the safety risk analysis and management for adjacent buildings has not been reached in tunnel construction fields so far. In the meantime, most of the studies have focused on single tunnels, and less works have been devoted to twin tunnels without taking into account the effect of tunnel-building interaction [10, 11]. Compared to single tunnels, there are more factors which contribute to the interactions between twin tunnels and surface buildings [12]. In this research, a systematic and comprehensive safety management approach with detailed step-by-step procedures is developed for the protection of existing buildings adjacent to the metro tunnel under construction. The potential safety status of a definite nearby building is assessed within four different risk levels. Corresponding prophylactic measures for nearby buildings at different risk levels are further provided according to risk assessment results.

2 Project overview

Wuhan Yangtze River Tunnel (WYRT), known as the first road tunnel under the China’s longest river Yangtze River, is an important route connecting two large cities of Wuhan, namely Wuchang and Hankou. It is a double-spool tunnel with a diameter of almost 12 m, a total length of almost 5,049.2 m and a total investment of 335 million dollars. The location of the WYRT construction is shown in Fig. 1. In the south and north sides of the Yangtze River, WYRT is designed to pass under through five pre-existing urban trunk roads. Affected by the extremely complicated geological conditions, including the uneven soft stratum and supper-shallow buried depth of the tunnel, several world-class challenges are encountered in the whole tunneling process [13]. Inevitably, the tunneling excavation can generate significant disturbances to surrounding environments, which may have negative effects and cause potential damages to the surface buildings, especially in densely built area.

Due to scarce land resources in the metropolitan area of Wuhan, there are numerous buildings overlying the tunnel in this central urban area. To be specific, brick masonry and reinforced concrete buildings dominate in this area. Most of these buildings are 2-7 stories high, and are typically supported on shallow foundations with a buried depth of 1-4 m. Many of nearby buildings are built in the late 1980s or early 1990s, and cracks can be observed almost in every building. Currently, limited published data and construction experiences related to such kinds of large-span and double-spool tunnel projects constructed under densely occupied buildings in soft soil ground are available [1]. It is therefore necessary to investigate the impact of tunneling excavation on the adjacent buildings, and then carry out the overall safety management of these adjacent buildings in tunneling environments.

3 Safety management approach

Step 1: Influence area determination

In urban areas, it is essential to protect pre-existing structures and underground utilities from being damaged due to ground movement caused by the construction of a metro tunnel. Therefore, it is particularly important to know the influence degree and scope as a result of a metro tunnel construction [14]. To analyze the impact of tunneling excavation on adjacent buildings, it is necessary to determine the influence area where the surface buildings can be affected potentially at the first step. Currently, the empirical method is widely used to predict ground movements as tunneling proceeds in construction practice. Martos firstly proposed that the shape of the settlement trough could be well represented by a Gaussian or normal distribution curve. Later, Peck analyzed settlement data from a large number of tunnels and mines by fitting Gaussian curves, and offered suggestions about the equation of settlement.
The approximate ground settlement curve can be used to provide an easy understanding of the influence area induced by tunneling excavation. Surface buildings in the estimated influence area can then be identified for further investigations. According to the geological conditions of the WYRT construction, the parameters information ($\phi=24^\circ$, $z=13$ m, $R=5.69$ m) are entered as inputs, and ground settlement curves are subsequently conducted in different scenarios of stratum loss ratios ($V_l$), as seen in Fig. 2. Obviously, the ground settlement increase sharply as $V_l$ increases, indicating that the stratum loss should be strictly controlled during the tunneling process. Meanwhile, the single side of the ground settlement trough is about 20-30 m wide. Conservatively, surface buildings that are 30 m offset from the projections of the tunnel centerline can have potential to be destroyed or damaged.

**Step 2: Building health investigation**

Most old buildings are aging and do not have complete load-bearing capability designed, and some kinds of structural damages are likely to occur in existing buildings in the process of long-term operation [15, 16]. The health condition of an existing building itself provides a basis as to how much the additional deformation or loads it is able to bear. This aging factor is rarely considered in previous FEM-based analysis due to the complicity and essential characteristics of the aging facilities, which will further affect the accuracy of the final calculation results to some extent. Therefore, in order to analyze the safety and serviceability of one existing building, it is necessary to monitor, assess, and predict structural integrity and durability of the building structures and their various components.

By numerous monitoring, tests and experimental studies and practices, some kinds of specifications or standards regarding the evaluation of the building health condition have been proposed. These specifications and standards provide an easy solution for the evaluation works, especially when a large group of buildings need to be evaluated. The former Soviet Union government issued a standard regarding the graduation of the protection of brick buildings adjacent to deep excavation, in which the total deformation ($\Delta L$) was used to assess the degree of the potential damage. The Ministry of Construction in China issued “Standard for appraiser of reliability of civil buildings (GB 50292-1999)” in 1999, in which an indicator $\frac{R}{\gamma_0 S}$ was used to appraise the load-bearing capacity of concrete structural components. Herein, $\gamma_0$ refers to the importance coefficient of structural components, $R$ refers to the structural resistances, and $S$ refers to the mechanical effect. The load-bearing capacity of concrete structural components can subsequently be classified into five levels. However, this standard did not consider the technical conditions of the building foundation and its superstructures based upon $\frac{R}{\gamma_0 S}$. A new guide document “Standard for structure safety appraiser of buildings (DB11 T639-2009)” was issued in 2009 in China. In this standard, the health conditions of the main building components, including the superstructure, substructure and building envelope in details, can be taken into account, and the overall structure safety of the existing building is assessed within four different levels, namely “A (Good), B (Normal), C (Poor), D (Endangered)”, as seen in Table 1. Due to its comprehensiveness and operability, this standard is easily accepted by engineers in construction fields.

<table>
<thead>
<tr>
<th>Levels</th>
<th>Description</th>
<th>Building Health Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Good</td>
<td>The building structure is safe and reliable without any serious defects or dangerous building components. The building can be used safely under the normal applying load. The building structure is safe without any dangerous building components. The building can be used safely under the normal applying load. The capacity of partial structural components cannot satisfy the requirement of ultimate state under normal serviceability. Some structural components are unsafety, leading to partially endangered buildings. The capacity of the load-bearing components cannot satisfy the requirement of ultimate state under normal serviceability. Major structural components are unsafety, leading to totally endangered buildings.</td>
</tr>
<tr>
<td>B</td>
<td>Normal</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Endangered</td>
<td></td>
</tr>
</tbody>
</table>

**Step 3: Safety risk assessment**

Risk assessment is the determination of quantitative or qualitative value of risk related to a concrete situation and a recognized hazard [17]. It is common practice to find in most safety sections on
current codes around the world that Risk = Hazard × Vulnerability × Value of the consequences [18]. In this risk framework, with regard to the safety risk assessment of an existing building nearby a metro tunnel, the spatial neighbor relation between the metro structure and adjacent buildings can represent the hazard component to some extent. As seen in Figs. 3 and 4, the location x (horizontal distance between this location and the tunnel centerline) plays a significant role in the tunnel-induced ground settlement. In most cases, the magnitude of the tunnel excavation effect seems to be slowed down as the building foundation is becoming far away from the metro structure. The spatial neighbor relation can be mainly measured in two directions, including the horizontal distance and the vertical distance. Accordingly, both the horizontal and vertical distances are used to define the hazard component. In this research, the hazard parameter is then divided into five neighbor levels, namely “1 (Very far), 2 (Far), 3 (Close), 4 (Very close), 5 (Extremely close)”, as seen in Table 2.

Table 2. Gradation of spatial neighbor relation between the tunnel structure and nearby buildings.

<table>
<thead>
<tr>
<th>Levels</th>
<th>Description</th>
<th>Spatial Neighbor Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very far</td>
<td>The location of the building is beyond the scope of tunneling excavation effect, typically more than 30 m away.</td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td>The horizontal distance falls in a range of 10-30 m. 1) The horizontal distance falls in a range of 3-10 m; and 2) The depth of the existing building foundation bottom is deeper than the tunnel buried depth.</td>
</tr>
<tr>
<td>3</td>
<td>Close</td>
<td>The depth of the existing building foundation bottom is shallower than the tunnel buried depth. 1) The horizontal distance falls in a range of 3-10 m; and 2) The horizontal distance between the metro tunnel structure and the adjacent building is less than 3 m.</td>
</tr>
<tr>
<td>4</td>
<td>Very close</td>
<td>The depth of the existing building foundation bottom is deeper than the tunnel buried depth.</td>
</tr>
<tr>
<td>5</td>
<td>Extremely close</td>
<td>The horizontal distance between the metro tunnel structure and the adjacent building is less than 3 m.</td>
</tr>
</tbody>
</table>

As aforementioned in Step 2, the health condition of an existing building can generally reflect its resistance capacity to additional deformation or loads. As a result, the qualitative description of the vulnerability parameter can be specified in the building health condition. The value component is not considered separately in this research, since the value of both the metro tunnel and nearby buildings is too high to be evaluated due to their contributions to urban development. Furthermore, each nearby building is assumed to be protected with policy factors taken into consideration. Thus, a simplified risk assessment framework regarding the impact of tunneling excavation on adjacent buildings can be achieved. The potential safety risk of a specific existing building in tunneling environments can be divided into five different levels, namely "I (Safe), II (Low risk), III (Medium risk), IV (High risk)”, as seen in Table 2. The higher the level, the higher the risk for the adjacent building. For instance, assuming a nearby building with a neighbor relation of Level 5 (Extremely close) and a health condition of Level B (Good), the safety risk level can then be rated Level III (Medium risk).

Step 4: Safety management strategies

Risk is defined as a state of potential damage which can be avoided or put under control by suitable and careful measures adopted for safety control [19]. With regard to adjacent buildings, safety control and management aims to avoid risks and keep them from being damaged. Numerical analyses have proved to be the most accurate and realistic approach for tunnel safety analyses, especially in complex environments [20, 21]. However, numerical simulation is laborious and time-consuming, and will turn out to be practically uneconomical when adopted blindly [22]. Decision making provides a means for systematically dealing with complex problems to arrive at a decision [23]. According to the safety risk assessment results, the degree of the potential safety risk of a specific nearby building can be obtained, and then relevant safety control measures can be proposed in advance for risk response. As aforementioned, due to the time and investment limitations, there is no need to carry out the FEM-based numerous analyses for a pre-existing
building at a low risk level. In this research, the degree of the potential safety risk acts as a decisive role in the depth of the safety analysis and the determination of corresponding control measures. Table 3 illustrates prevention and control measures for existing buildings at different risk levels. As seen in Table 3, the FEM-based safety analysis is employed in the situation where the existing building lies in a risk level of III or IV. In the meantime, the actual implementation effect can be further analyzed when corresponding control measures are adopted, which can also provide feedbacks and suggestions for adjustments or optimizations in previous steps.

4 Case study

4.1 Background

Wuhan is the largest city in Central China with a population of 10.02 million (2011 data). WYRT is constructed to relieve the pressure of urban traffic jams across the Yangtze River. WYRT is a double-spool tunnel with a total length of 5,049.2 meters, while the Left Line is 2,550 m (LK2+720 ~ LK5+270) and the Right Line is 2,499.2 m (RK2+778 ~ RK5+277.2). Two slurry shield machines with a cutter diameter of 11.38 m are used to push the tunnel from Wuchang district to Hankou district. In accordance with the construction schedule, one slurry shield machine is utilized to excavate the Left Line from Wuchang to Hankou, while the other is utilized to excavate the Right Line after 2 months’ lag in the same direction. As aforementioned, crowded buildings are encountered in the influence area induced by tunnel excavation, among which a five-story frame teaching building (FFTB) is chosen as a case to verify the applicability of the proposed safety management approach.

FFTB, built in 1984, is a reinforced concrete framed structure. It provides 50 rooms for almost 3,000 students for training purpose at Wuhan University of Technology. FFTB is the first adjacent building that the shield machine passes through. The tunnel crosses right under the foundation of FFTB at LK4+943, with a height of 6.47 m from the tunnel roof to the building foundation base. Fig. 6 presents the horizontal drawing of FFTB that is adjacent to the twin-tunnel WYRT. Fig. 4 illustrates the cross-section drawing of this case. Due to limited published data and construction experiences regarding safety management of nearby buildings adjacent to twin-tunnels, some research works on the prediction of the building behaviors and distortions have to be performed before the tunnel construction.

4.2 Risk assessment result

With regard to FFTB near the construction of WYRT, the developed safety risk assessment framework (see Fig. 3) is first used to evaluate the overall risk level, which can then provide a basis for determining corresponding control measures for risk reduction. As seen in Figs. 6 and 7, WYRT passes right through the foundation of FFTB, and the depth of the building foundation is shallower than the tunnel buried depth. Thus, the spatial neighbor relation between the tunnel structure and the nearby building is rated a level of 5 (Extremely close) according to Table 2. Meanwhile, the health condition of FFTB falls to a level of B (Normal) after a thorough investigation. In this situation, the potential safety risk of FFTB can be assessed at a level of III (Medium risk) based upon the safety risk assessment framework as seen in Fig. 3. According to safety management strategies (see Table 3), in order to assure the safety and serviceability of FFTB, it is therefore necessary to carry out numerous simulation analyses, and then propose corresponding control measures based upon analysis results.

4.3 Numerous simulation analyses

To simulate the impact of tunnel excavation on the adjacent building, a full numerical model (see Fig. 5) is developed using a 3D coordinate system. It is defined that the X-axial denotes the distance from the tunnel centerline in the lateral direction, the Y-axial is the coordinate in the longitudinal direction, and the Z-axial is the depth below the surface. Fig. 5 (a) represents the finite element model of FFTB, where beams, plates, columns and foundations are the main load-bearing components. Spatial four nodes element C3D4 is used to simulate these components. Fig. 5 (b) represents the
finite element model of the tunnel structure, where segment, grouting concrete and shield shell are principal load-bearing components. Spatial eight nodes element C3D8R is used to simulate the segment and grouting concrete, while the shell element is used to simulate the shield shell. Fig. 5 (c) represents the interactive mechanics effect, where the soil plays a critical role in tunnel-building interaction. Spatial eight nodes element C3D8R is used to simulate the soil. In the whole simulation model, there are 43,714 nodes and 47,563 elements in total.

**Table 3. Prevention and control measures for existing buildings at different risk levels.**

<table>
<thead>
<tr>
<th>Levels</th>
<th>Description</th>
<th>Prevention and Control Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Safe</td>
<td>1) No need to take special pre-reinforce measures; and 2) Carry out regular monitoring of surface subsidence and building foundation deformation during the construction.</td>
</tr>
<tr>
<td>II</td>
<td>Low risk</td>
<td>1) Take necessary pre-reinforce measures for surrounding soil before the tunneling excavation; 2) Strengthen the monitoring of surface subsidence and building foundation deformation, and implement tendency analysis of the excavation-induced pile deformation every week; and 3) Strictly control technical parameters in tunneling excavation process in accordance with monitoring data.</td>
</tr>
<tr>
<td>III</td>
<td>Medium risk</td>
<td>1) Take special pre-reinforce measures for surrounding soil in excavation area; 2) Carry out field tests and numerical simulation analyses for safety risk analysis before tunneling excavation, and optimize construction schemes and technical parameters according to analysis results; and 3) Strengthen the monitoring of surface subsidence and building foundation deformation, and strictly control technical parameters in tunneling excavation process in accordance with monitoring data.</td>
</tr>
<tr>
<td>IV</td>
<td>High risk</td>
<td>1) Take special pre-reinforce measures for surrounding soil in excavation area, as well as the soft soil area around the building foundation; 2) Carry out field tests and numerical simulation analyses for safety risk analysis before tunneling excavation, and optimize constructions schemes and technical parameters according to analysis results; and 3) Strengthen the monitoring of surface subsidence and building foundation deformation, and invite domain experts to conduct the professorial tendency analysis of the excavation-induced bridges damage.</td>
</tr>
</tbody>
</table>

Fig. 5. Finite element models of the building, tunnel and soils: (a) building; (b) twin-tunnel; (c) full model.
In terms of numerous simulation analyses, methods to analyze the complex tunnel-soil-building interaction can be classified broadly into two categories so far. To study the tunnel-induced damage to the existing building, the impact on the building foundation deformation is studied. As analyzed above, the foundations $B_4$, $A_4$, and $A_5$ are likely to perform the maximum surface settlement. Thus, these three foundations are chosen to investigate the impact of tunnel excavation on the displacement of building foundation. Fig. 6 illustrates the displacement diagram of building foundations in X, Y and Z directions. These results are further analyzed as follows:

1) Fig. 6 (a) presents the evolution of the lateral displacement (X direction) of the building foundation during the tunnel excavation. The lateral displacement increases away from the twin-tunnel centerline during the single tunnel excavation. However, the lateral displacement increases close to the twin-tunnel centerline during the twin tunnel excavation. The foundation $A_4$ displays a maximum lateral displacement of 3.7 mm during the single tunnel excavation, while the foundation $A_5$ displays a maximum lateral displacement of 3.1 mm during the twin tunnel excavation. The maximum lateral displacement appears in the process of the tunnel excavation, rather than the end states. In general, the lateral displacement shifts in two opposite directions during the whole excavation, which can cause a massive challenge for the concrete toughness of FFTB. Specifically, FFTB will experiences two wild swings during the entire excavation process in X direction, and then longitudinal cracks are likely to appear in weak links of structural components of FFTB. Accordingly, the lateral displacement monitoring should be paid much more attention.

2) Fig. 6 (b) presents the evolution of the longitudinal displacement (Y direction) of the building foundation during the tunnel excavation. In the single tunnel excavation, the longitudinal displacement of each foundation increases when the tunnel face becomes close to the foundation section, and then decreases when the tunnel face moves away from the foundation. The tendency of the longitudinal displacement is very similar among all foundations ($B_4$, $A_4$ and $A_5$). The same situation exists in the twin tunnel excavation. Finally, the maximum total longitudinal displacement appears in the foundation $B_4$ with a displacement of 7 mm.

3) Fig. 6 (c) presents the evolution of the building foundation settlement during the tunnel excavation. In the single tunnel excavation, the vertical settlement increases continuously, and reaches about 55–60% of its final value when the tunnel face crosses the foundation section. The maximum settlement is observed at the foundation $A_4$ with a total settlement of about 24 mm. Meanwhile, the single tunnel can cause a differential settlement of about 8.2 mm between the foundations $A_4$ and $B_4$. It is similar for the twin tunnel excavation. Finally, the settlement of each foundation reaches about 34 mm of its final value.

![Fig.6. Displacement diagram of the building foundation: (a) X direction; (b) Y direction; (c) Z direction.](image-url)
4.4 Safety control measures

According to numerical simulation analyses, the impact of tunnel excavation on the foundation deformation of FFTB can be controlled by taking reasonable construction measures. The excavation process of the tunnel is an important factor affecting the structural stability of the nearby building. The following control measures should be adopted during the tunnel construction.

(1) Foundation reinforcement. To reduce the effect of soil deformation and improve the bearing capacity of the existing building foundation in soft soil, the foundation reinforcement comes to be the first choice. Generally, these schemes, such as the extended foundation dimension, static bolt piles and root piles are three schemes widely used in foundation reinforcement. Through a comparison of these schemes in the view of safety, reliability and low cost, the root pile scheme turns out to be a more competitive option due to a faster convenience with high efficiency, especial for low-story buildings near shallow-buried tunnels in soft soil. The grouting material is injected with high pressure about one month before the shield machine crosses through the foundation. The diameter of the root pile is designed to be 250 mm, and the observed concrete copings are used to link root piles to existing foundations of FFTB.

(2) Parameter optimization for the slurry shield machine. Tunnel-induced ground disturbance mainly results from frictional effect during the tunnel excavation, and it is therefore necessary to strengthen the operational safety of the shield machine, especially in circular and vertical curve sections. In the first 30 m before the tunnel face, the shield machine should drives slowly at a speed of 4-6 mm/min in both single and twin tunnel excavation, in order to adapt the environment around. While crossing under the existing foundation, the machine should drives fast at a speed of 10-16 mm/min. Besides that, in order to decrease the ground disturbance induced by rectification of shield machine, the horizontal deviation should be controlled within ±50 mm. In regard to a shallow-buried tunnel, the elevation deviation should be controlled within ±20 mm, but below the designed tunnel axis.

(3) Quality control for backfill grouting. The interspace of the shield tail is another principal reason causing the displacement of surrounding soil and building foundation. In order to reduce the subsidence developing to the surface, backfill grouting proves to be an effective technique in various tunnel projects. Grouting quantity, pressure and speed are the three main factors affecting the grouting quality. At first, the practical grouting quantity should be 50%-80% more than the theoretic interspaces of shield tail. Grouting pressure should be a little higher than the soil pressure, and is generally confined in a range of 0.2-0.4 MPa. Finally, grouting speed should be adjusted in accordance with the driving speed, and then the homogenous penetration can be reached during the entire excavation process.

5 Conclusions

In recent years, safety management of existing buildings near metro tunnels has attracted broad attention due to the rapid development of underground transport systems. Tunnel-soil-building interaction is considered a highly complicated process, and various factors involve in safety violations in tunnel construction practices. A novel safety management approach with detailed step-by-step procedures is developed for the assurance of adjacent buildings, including 1) construction influence area determination; 2) building health investigation; 3) safety risk assessment; and 4) safety management strategies. The potential safety of a nearby building is assessed within four different risk levels, namely “1 (Very far), 2 (Far), 3 (Close), 4 (Very close), 5 (Extremely close)”. The assessed risk level plays a decisive role in the depth of safety analysis, aiming to strike a reasonable balance between the project system safety and cost constrains. It is suggested that the numerous FEM-based analysis should be employed in the situation where the existing building is rated a risk level of III or IV. A case concerning the protection of a five-story framed building adjacent to a twin-tunnel that is WYRT in China is presented. The impact of the single and twin tunnel excavation on both the soil displacement and the building foundation deformation is further analyzed according to numerous simulation results. Implementation effects verify the applicability of the proposed approach. The proposed safety management approach presented in this research can be translated to other tunneling projects within similar urban settings, and can be used to increase the likelihood of a successful project in a complex project environment.

Tunneling excavation can exert a profound effect on the safety of surface and subsurface structures, such as adjacent bridges, underground buried pipelines and others. Further research works will concentrate on investigating the risk prevention mechanism to excavation-induced environmental damage, as well as to environmental protection strategies.

Acknowledgements

The National Science and Technology Support Plan (51378235), Wuhan City Construction Committee Support Projects (201217, 201334) and China
Scholarship Council (CSC) are acknowledged for their financial support of this research.

References


Abstract -

The construction industry is multi-disciplinary and collaborative in nature. Project managers are expected to understand the relations, roles and responsibilities in this collaborative working environment. Construction project managers need to be equipped with skills to process and understand the principles of interdisciplinary working. In order to keep pace with industry requirements, it is necessary for universities to teach subjects in tertiary education courses that educate and motivate construction students towards interdisciplinary working. This paper is based on a research project aimed at understanding opportunities and challenges for introducing subjects that require students from different disciplines to work together on an integrated project. In order to teach interdisciplinary working principles to project management students, a new post-graduate subject, Integrated Project Delivery, was introduced in the Master of Project Management at the University of Technology Sydney (UTS) as part of a construction sub-major. The subject was designed and teaching materials prepared based on inputs from architecture, engineering and construction management academics. In the teaching of the subject, interdisciplinary student teams were formed based on educational background and professional experience. An (pedagogical) action research approach was adopted to study the challenges and benefits of new ways of learning in line with the UTS learning strategies being adopted by academics. The main finding of this research concluded that project-based learning is very valuable to both students and industry alike as it promotes working on a live project enthusiastically and gaining industry experience in new ways of working adopted by industry. Working in multi-disciplinary teams requires students to respect the other team participants from different backgrounds, inculcating values of team spirit and discouraging adversarial behaviours. Furthermore this research provides insights about student expectations from a post-graduate subject and their experience of a multi-disciplinary working environment. This paper discusses both the successes and the issues faced during the teaching of this subject and suggest future directions for research and effective approaches to implementing interdisciplinary working in an educational environment.

Keywords -

Project-based learning; Multi-disciplinary teams in education; Flipped and blended learning; Collaboration; Interactive learning spaces; action research.

1 Introduction

The introduction of advanced technologies in the last decade has stimulated evolutionary changes in the construction industry, with impacts on engineering, architecture, construction management and related construction professions. Emerging trends in demographics, economics and globalization have increased competition in current business models and thus collaboration has become mandatory for survival [1]. There is a necessity for more practical multi-disciplinary collaboration amongst industry professionals to deliver advanced construction projects while meeting client demands. An interdisciplinary working environment and integrated approach towards project delivery is considered to optimize efficiency and deliver better outcomes. In current education models most students spend the majority of their time within their individual disciplines working on projects that do not assist in building interdisciplinary teamwork or interdisciplinary communication skills [2].

Rapid changes in the construction industry arising from the prompt adoption of the latest and most advanced technologies is making it necessary for professionals to regularly update their knowledge and skills throughout their career [2]. The implementation of effective interdisciplinary collaboration in day-to-day professional work environment of the Architecture, Engineering and Construction (AEC) industry has not kept pace with advances in technology. To keep pace
with the introduction of technology and industry requirements, construction project managers should pay more attention to the human side of professionalism and develop skills to be more flexible and adaptable to inevitable changes in their career and roles. Various studies have revealed a lag in the adoption of emerging trends in construction education as compared to the rapid advancements occurring in AEC industry [3, 4]. In the US, Europe and Australia, AEC students continue to be educated in separate faculties/schools, with little or no integration or collaboration between the disciplines. Within the higher education setting, advanced multi-disciplinary collaboration frequently encounters resistance as the traditional silos of architecture, engineering and construction prove difficult to bridge. As in industry, lack of cooperation between the professions also exists in academia, and questions often arise as to who should be responsible for, or take ownership of, cross-disciplinary subjects – especially when such subjects span across faculties. Furthermore, education in construction management is under immense pressure to evolve in step with the demands of a rapidly evolving industry and to develop skills, strategies and attitudes that will prepare graduates for their working lives and inculcate adequate professionalism.

This research paper looks at how the subject content, teaching methods, and interdisciplinary teams in a postgraduate subject can be aligned to promote students’ understanding of collaborative working and to develop the graduate attributes demanded by current construction industry practices. The research also evaluates industry expectations from construction management graduates and analyses student expectations from a subject offering interdisciplinary collaboration and project-based learning.

2 Background

In the case of attempts to introduce interdisciplinary learning in education, Fruchter [5] documents a collaboration between six universities across Europe, Japan, and the United States. An experimental framework was implemented, along with a tool kit to assist the interdisciplinary team members to collaborate successfully and to evaluate the outcomes. There are also several other instances of universities who have implemented and reported on individual cross-faculty collaborative design courses, such as the University of Sydney [6], Carnegie Mellon University [7], and Georgia Institute of Technology [8]. Other Universities in Australia, such as the University of South Australia and the University of Newcastle, have started to develop multi-disciplinary AEC courses at undergraduate level, but none appear yet to have attempted fully multi-disciplinary AEC courses at the postgraduate Masters level, despite institutions in the US, Europe and the UK having done so. The documented case studies often focus on the advances in information technology and its adoption amongst students, rather than how these strategies enhance educational practices in construction management.

The implementation of ‘blended’ learning procedures is comparatively recent in Australian education models. Blended learning combines face-to-face learning with distance (out of class) learning with a goal of harnessing the positive characteristics of both teaching methodologies[9]. Constructivist theorists believe consequential learning can be accomplished through the implementation of realistic, simulated problems to be solved [10]. On the other hand, variation theorists emphasize the importance of students’ experience of variation to address critical aspects of the subject matter [11]. Oliver and Trigwell [12] support the idea of combining blended learning with interdisciplinary team work to solve a complex problem, arguing that this allows students to experience the patterns of variation necessary to engender different perspectives and different ways of seeing things. Despite the various benefits of these education methods, there is still a lack of studies around interdisciplinary education in post-graduate courses which employ project-based learning in a ‘flipped’ and ‘blended’ learning environment, such as is being encouraged and trialled at universities like UTS.

The new postgraduate subject, Integrated Project Delivery, introduced into the UTS Master of Project Management degree attempts to fill a crucial gap in the education of construction management students. It combines a flipped and blended learning format with the advantages of new digitally advanced learning and teaching spaces that support and enhance students’ learning experience. Extensive consultation with representatives of the relevant professional bodies and AEC industry was undertaken to ensure that the new subject meets current industry demands, equipping students with the attributes required by industry and thus assisting to increase the employability of UTS graduates.

3 Methodology

This research has examined the needs of the construction industry and the development of educational strategies to deliver skilled graduates equipped with the requisite professional experience. It has been concerned with developing practical knowing for advancing the human dimension of professionalism among students. Action research methods were identified as most
suitable for this study; with the plan-act-observe-reflect cycle used to collect data through reflective journals maintained by the academics involved. Thus the research brings together action and reflection, theory and practice in the pursuit of practical solutions to meet industry demands and more generally the flourishing of individual persons and their educational and professional communities [13]. Figure 1 represents the action research cycle adopted for this research and its relation to various stages in this project.

Figure 1: Project Action Research Cycle.
Stages included 1) Plan: Planning subject content matter and deliverables, methods of subject delivery; 2) Action: Subject delivery and workshops; 3) Observe and Evaluate: Observing effects of new ways of learning and collaboration during workshops and; 4) Reflect: Collecting student feedback and reflections from Industry Focus group interviews.

Further, four phases of this research can be described as: Phase I (Plan): This phase included conducting a literature review and referring it for purposes of subject delivery. This phase also included enrolling students from different disciplines into the subject. Phase II (Action): This phase dealt with subject delivery and conducting focus group interviews. About 26 students participated in this subject and a reference group of 12 participants was selected for focus group interviews. Phase III (Observe): Data was collected throughout the semester at he workshops held for the subject. Observation regarding student reaction and behaviour for new way of learning was recorded. Student reflections about how the subject was taught and feedback through the university’s feedback system. Industry expectation from such a subject was gathered through a focus group for future direction. Data was analysed with help of Nvivo 10, and findings were recorded. Phase IV (Reflect): Based on findings, and discussion among staff ideas for future teaching were gathered. These ideas have been incorporated for further teaching of this subject.

The new postgraduate subject, Integrated Project Delivery, in the UTS Master of Project Management (MPM) degree, incorporates a project-based learning approach. Within this subject, the action research cycle includes the use of drawings developed for a building project to solve simulated problems based on real-life scenarios derived from industry. The subject was taught using a 2×2 block workshop mode over the course of a semester supported by online reading materials and online forums. The subject tried to combine both flipped and blended learning about interdisciplinary design and construction management and was designed to provide students with the experience, tools, and methods needed to improve their understanding of construction management processes.

Students participating in the MPM course belonged to variety of backgrounds such as engineering, architecture and construction management and formed multi-disciplinary teams, working together to generate designs, contractual documents and tenders, while experimenting with different work practices to take maximum advantage of project based learning. The students collaborated from remote locations via web-based approaches, while working on a live project. As such, they were encouraged to utilize technologically advanced methods of communication, correspondence and coordination that combined face-to-face interactions with the virtual. The subject also made use of the sophisticated future learning spaces at UTS, which are equipped with interactive white boards, pods and other facilities to assist students to collaborate during team meetings. Students used the software installed on the Pod computers and interactive white boards in the collaborative learning space [14] to manipulate and mark-up drawings and models. The flat tables and moveable whiteboard screen zones were ideal for facilitating group interaction. Students were asked to view online tutorials and lessons in their own time, and to come to the classes ready to work in groups and to be mentored by external guest tutors at regular intervals.

At the conclusion of the subject, a student feedback survey was conducted to gather data that could be analyzed to understand various factors that contributed to the learning experience in this subject.
A reference group was formed comprising representatives from relevant industry bodies and professional associations and senior managers from industry. A needs analysis was carried out with the reference group to determine industry expectations of UTS graduates working on complex construction projects. Data collected from the reference group and during two teaching workshops during the semester was analyzed to gauge the success of the subject and to inform decisions regarding further subject development. Therefore data was collected from a wide range of sources, including focus group discussions with industry professionals; the use of the Self & Peer Assessment Resource Kit (SPARK) in the student workshops with protocols for peer review in blended learning environments proposed by McKenzie, Pelliccione [15]; and student feedback surveys (SFS), which included additional questions relevant to the evaluation of the teaching and learning strategies adopted. Nvivo 10 was used for coding and analysis purposes. The analysis provided some surprising results, articulated in the following sections.

3.1 Subject Delivery Process

Since the subject was part of a Postgraduate course with several students working in industry, it was considered appropriate that the subject content be delivered in workshop mode separated by a break to encourage students to experience working as virtual teams. Two workshops were conducted for the delivery of subject content. Each workshop lasted for two days; one was conducted at the beginning of semester, with the second held towards the end of semester.

Workshop 1: The first workshop was utilized to introduce the subject and its new learning and teaching methods. Student teams were formed based on professional and educational background.

A site visit to a selected live project was arranged and students had an opportunity to interact with real project clients as well as members from the project’s architectural team and project management consultants. This helped them to gain experience on a live project, discovering the site constraints, environmental conditions, and project requirements and applying these reflectively to a problem-driven assignment.

Additional to the site visit, lectures were delivered by academics and industry professionals to familiarize students with various topics including: collaborative working in interdisciplinary teams, current practices in project management, advanced information technology tools for project management and communication, latest trends in the construction industry in Australia, working with government authorities, collaboration practices on large-scale projects, tendering processes and contractual agreements.

Assignments: After the first workshop students were asked to propose design changes and prepare tender documents based on their new designs as part of an assignment submission. Students were encouraged to use web-based communication tools for remote meetings. A virtual space was also created in UTS Online, a web-based tool to provide online learning to students. This virtual space provided for the maintenance of records, submission of weekly updates on frequency of communication, project progress and correspondence schedules. Self-assessment and group evaluation were conducted at two points during the period intervening between the two workshops to enhance productivity and outputs. The first SPARKPlus evaluation was set before the first assignment, such that results from evaluation and assignments were announced nearly at the same time. This allowed students to reflect on their teamwork and its effect on assignments. This in turn helped them to analyze the requirements of collaborative working.

Workshop 2: The second workshop was arranged towards the end of semester. Day 1 of the second workshop allowed students to interact with the contractors who had submitted tenders. At this time academics and industry professionals provided students with experience in selecting contractors, based not only on their quotations but also on an understanding of their work background and experience in similar projects. Day 2 involved presentations by the students of their work and understanding of the subject.

Following the second workshop, students were given a final assignment consisting of writing a brief personal reflection on the subject matter and learning experience, subject deliverables, methods of subject delivery and recommendations for future.

4 Findings and Discussions

Findings from this research are presented by following the sequence of (educational) action research, as shown earlier in Figure 1.

Stage1: Abstract Conceptualization. A review of the literature suggests that it is necessary to integrate current interdisciplinary working practices adopted by industry into the educational curriculum to prepare students to be work ready. Achieving the right mix of students in a cross-disciplinary subject is a cumbersome task that requires careful coordination with student administration (enrolment) authorities. This became particularly difficult when students with negligible professional construction experience were enrolled in the subject resulting in unbalanced teams. Taking into consideration the context of the case study project (the construction project management of a car park and
associated infrastructure) the issue of unbalanced teams posed a major issue and in one case led to the merger of two student teams.

The problem of unbalanced teams manifested itself in the assignment submission, which was based on construction tenders and contracts. Another concern highlighted at this time was the split workshop mode. Students who were used to doing a continuous block and had other subjects running between the workshops found it difficult to maintain interest and engagement in the subject. Contingency plans to negate these issues will need to be further developed.

Stage 2: Active Experimentation. Classroom workshops which included implementation of planned actions. Data was collected during the workshops to understand the students’ reactions to various new concepts adopted by this subject. Figure 2 shows team responses to a few key points observed during the workshops.

![Figure 2: Data collection results from workshop. The percentage on bar chart represents students support. As evident from the chart, the teams were not very comfortable with the small team size, which varied from 3-5 members from cross-disciplinary backgrounds. Most students suggested that a larger team size with two team members from each discipline would have assisted work distribution and more efficient outputs. Nearly 90% of the teams favored working in collaboration with members from different backgrounds. They stated that such integration assisted in looking at the same problem from different perspectives. Despite the experience gained from project-based learning, working on a live project was found to be stressful by a number of students. It demanded more accuracy from assignment submissions, which caused anxiety for those students who were not familiar with interpreting construction drawings. Another significant finding was that students would have preferred the site visit to be scheduled outside of the time allocated to the first workshop. Students would have preferred the workshop hours to be entirely devoted to clarifying questions and building rapport with other team members. Pre-reading materials were found very useful by teams to understand interdisciplinary working strategies and current progress in AEC industry towards integration. However, students realized that in order to work efficiently for assignments, their teams had to be proactive and adopt some of the proper methods for smooth working in collaboration, as listed below:

1. Developing a detailed understanding of each team member’s capabilities and documenting these.
2. Defining roles and responsibilities based on abilities and preparing a conflict resolution framework.
3. Establishing a communication framework including meeting schedules and correspondence routines.
4. Having a detailed discussion on the use of technologies for communication and preparing team agendas based on good project management principles.
5. Defining team protocols regarding document submissions.
6. Documenting all correspondence by individual team members as a means of recording the team’s progress on the project between workshops.

Stage 3: Concrete (Project) experience. Blend theory with practice. Following the workshops, focus group interviews were held with industry and academic experts. Their inputs on industry expectations from students were collected based on the following four general questions: 1) What is the current state of integrating works done by different teams in delivering projects in Australian AEC industry?; 2) What are some of the issues for integrating the work performed by different teams using current practices in industry, and how can these be addressed?; 3) What are the benefits of educating students about Integrated Project Delivery (IPD) practices to industry?; 4) What activities of collaborative working would you recommend be included in a subject that is designed to teach IPD? Feedback received from the industry focus group regarding the current state of integration is summarized in Table 1.

![Table 1. Industry Focus Group Feedback](#)

<table>
<thead>
<tr>
<th>Resisting Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of government policies.</td>
</tr>
<tr>
<td>Lack of contractual agreements.</td>
</tr>
<tr>
<td>Larger project teams.</td>
</tr>
<tr>
<td>Legality issues.</td>
</tr>
<tr>
<td>Tremendous governmental paperwork.</td>
</tr>
<tr>
<td>Lack of decision-making frameworks.</td>
</tr>
<tr>
<td>Differences in priorities and conflicts.</td>
</tr>
<tr>
<td>Traditional ways of working.</td>
</tr>
<tr>
<td>Lack of clarity on roles for large-scale projects.</td>
</tr>
<tr>
<td>Governance issues for complex processes.</td>
</tr>
</tbody>
</table>
Perception of project deliverables varies. Lack of well-defined process coordination. Lack of commitment / motivation by clients. Lack of project managers trained to work in a collaborative working environment.

Assisting Integration

Pre-defined project requirements. Commitment to innovation. Early engagement of stakeholders. Project start-up workshops. Selecting team members based on abilities. Willingness of team members. Suitable contractual agreements. Well distributed work load amongst members. Detailed Bid and specifications helps in Public-Private Partnership (PPP) projects. Common goal for project stakeholders. Back to front approach in construction process is beneficial for successful integration.

It is also important to consider how including this subject in education can benefit industry. This discussion with the focus group is summarized in Table 2.

Table 2: Benefits of including interdisciplinary collaboration in education.

<table>
<thead>
<tr>
<th>Benefits to Industry</th>
<th>Benefits to Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporate work practices and procedures. Industry makes graduates aware of current practices. Students are taught applied knowledge. Prepare students vocationally for industry. Relevance to industry and project context is realistic. Practicality of subject material. Teaching flexibility to select professions is important. Understanding differences in important. Students are aware of roles and responsibilities. Graduates are aware of project stages deadlines and times. Students have procedural information about all disciplines.</td>
<td>Students get applied experience by learning on a live project. Companies do not offer cadetships or internships and thus project-based learning helps them to achieve practical knowledge. Enhances standards of detailed knowledge in students’ background. Insight into the particularities of their own profession is provided. Students understand differences and develop abilities to address problem from different perspectives.</td>
</tr>
</tbody>
</table>

Stage 4: Reflective Observation. Think critically and analyze objectively. SFS and personal reflections are part of reflective observation. The questions in the survey can be classified into categories which are: i) Subject matter content and delivery; ii) Flipped Learning; iii) Blended learning.

Table 3 represents a classification of the responses collected in SFS.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>D</td>
<td>Disagree</td>
</tr>
<tr>
<td>N</td>
<td>Neutral</td>
</tr>
<tr>
<td>A</td>
<td>Agree</td>
</tr>
<tr>
<td>SA</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>

Figure 3 represents generalized student feedback on the subject based on satisfaction with the educational experience and subject deliverables. It is clear that the way the subject was delivered was not entirely positively received by the students. Based on these results, it is necessary to investigate what features adopted in delivering this subject were most disliked by students.

Figure 3: Survey results on subject satisfaction

Further analysis of the data collected via SFS surprisingly reveals that students were satisfied with the overall learning experience of this subject. There was however disagreement about way the subject was delivered.

Figure 4: Survey results on students learning experience
Data previously collected has demonstrated that project based learning and working in interdisciplinary teams are well accepted and appreciated concepts amongst students. Therefore, further investigation was done to differentiate between flipped and blended learning. From Figure 5, it is evident that the blended learning environment was not a problem in terms of subject resources and student experience.

Further analysis of the SFS data led to the conclusion that the flipped learning process was not popular amongst students in this subject. Figure 6 represents SFS results for flipped learning implementation in the subject’s studio/tutorial sessions. There was strong disagreement regarding the way subject matter was delivered during tutorials. Students thought that it did not assist in understanding the subject requirements. The results demonstrate general subject satisfaction but demand an enhanced subject delivery process.

While the students appreciated the web-based learning approach and were happy with the subject resources it is obvious that the flipped learning process adopted needs reconsideration before the subject is offered again.

As a part of the reflective cycle it was also necessary to record reflective inputs from industry professionals. Table 3 summarizes industry focus group suggestions and recommendations on such subjects in postgraduate university courses.

Table 3: Industry reflection on subject

<table>
<thead>
<tr>
<th>Industry recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project selected should be achievable with an interesting project brief.</td>
</tr>
<tr>
<td>New project with complex design and aesthetic value and an attractive project.</td>
</tr>
<tr>
<td>Select a public space and define a hypothetical project which is similar to urban planning project.</td>
</tr>
<tr>
<td>Realistic project and follow up the project management process with the implementation of some integration issues.</td>
</tr>
<tr>
<td>Small project to run from beginning through to handover of completed building.</td>
</tr>
<tr>
<td>A problem which is applicable to all disciplines.</td>
</tr>
<tr>
<td>Structuring integrated studies in already existing subjects.</td>
</tr>
<tr>
<td>Subject content can be part of communication, leadership, lean construction, professional practice subjects.</td>
</tr>
<tr>
<td>There should be no specific course on BIM but students should be taught BIM nevertheless.</td>
</tr>
<tr>
<td>Companies should offer cadetships or internships.</td>
</tr>
<tr>
<td>Students to be taught applied knowledge.</td>
</tr>
<tr>
<td>Traditional architecture subjects have integration topics and subjects which should be continued.</td>
</tr>
<tr>
<td>Practicality of subject material is important.</td>
</tr>
<tr>
<td>Teaching flexibility to select profession is important.</td>
</tr>
<tr>
<td>Selective enrollment process for students.</td>
</tr>
<tr>
<td>Reflective process for roles played.</td>
</tr>
<tr>
<td>Project based learning and minimum theoretical teaching</td>
</tr>
<tr>
<td>Include staff from all disciplines content delivery.</td>
</tr>
</tbody>
</table>
5 Conclusions

While the idea of teaching multi-disciplinary teams using a live project is generally favoured by both the industry and students, improvements are required in several aspects of how the subject is taught. These being:

1. The project selected should be slightly more complex to provide a variety of challenges that require teamwork.
2. The brief should be expanded beyond the actual project to enable students to apply their various skills.
3. The university administrative set up (or enrolment process) needs to be modified to facilitate the right blend of students in the class.
4. Some more lectures need to be included during the workshops for download of technical information.
5. Flipped learning requires students to complete more preparatory activities prior to attending the workshops.
6. Helping students to learn soft skills to engage well in multi-disciplinary teams is more important than teaching them the use of technology.

One of the major benefits and innovations of this subject is that students understand the significance of process design and protocols for working in multi-disciplinary teams.

6 Acknowledgements

This research project was funded by Learning and Teaching Grant awarded to the research team by the University of Technology Sydney.

References

Safety Practices in the Lebanese Construction Market: Contractors’ Perspective

R. Awwad*, M. Jabbourb and O. El Souki

*aCivil Engineering Department, Lebanese American University, Lebanon
bCivil Engineering Department, Lebanese American University, Lebanon
cCivil Engineering Department, Lebanese American University, Lebanon
E-mail: rita.awwad@lau.edu.lb, melanie.jabbour@lau.edu, omar.elsouki@lau.edu

Abstract -

The construction industry in the world is considered to be one of the most dangerous industries, claiming each year the lives of many workers. In particular, the boom in the construction industry in the Middle East over the past decades was accompanied with significant increase in death tolls and injuries among labor forces. Lebanon is a small developing country in the region who survived a harsh civil war that ended in 1990, and since then, its economy has become more dependent on the construction industry. The country has been undergoing a reconstruction and rehabilitation phase that led to a prosperous growth in construction, however, not accompanied by any improvement of the construction laws to include safety regulations and standards. The purpose of this research is to assess the existence and extent of application of safety practices by Lebanese contractors, management commitment to safety, implementation challenges and opportunities for enhancing safe practices. Based on the aforementioned, this study highlights the benefits and deficiencies of the current safety practice in the Lebanese construction market and concludes by providing recommendations to positively influence industry professionals and foster their commitment to safety. The aim of this paper is achieved through conducting one to one interviews with Lebanese contractors. Findings showed a lack of awareness among small and medium size contractors about the necessity of developing and applying safety training programs on the jobsite. However, large contractors showed a much stronger commitment to safety through planned safety management programs and regular inspections but within a limited budget.

Keywords

Construction management, injuries, hazards, construction safety, construction worker, safety and health management systems, Lebanese construction market.

1 Introduction

The construction industry has been the leading cause of the most dangerous injuries among all other industries [1]. According to the U.S. Bureau of Labor Statistics in the U.S. department of Labor, the construction industry resulted in the most fatal injuries in 2012 [2]. The nature of work occurring on construction sites is hazardous since it involves the use of heavy equipment, erection of temporary structures extending to considerable heights, the use of unhealthy materials, as well as having to work sometimes in extreme natural conditions for lengthy time periods. Construction injuries not only have direct costs of treatment and compensation but they also have several indirect impacts on the workers such as low morale and loss of productivity. This has direct impacts on the project progress including completion delays, resource replacement costs, and higher insurance premiums. In order to remedy for this situation, safety measures and programs have been introduced and have even become required by law in some countries. In fact, a cost-benefit analysis of accident prevention conducted by Ikpe and et al. showed that the benefits of accident prevention outweigh the costs of accidents by a ratio of 3:1 [3].

Most companies in developed countries design their safety programs following OSHA standards however the main challenge remains to ensure that management and supervisory personnel are committed enough to enforce these guidelines and to provide adequate training to workers and educate them about the importance of safe working conditions [4]. It is
important to realize that a safety program’s success lies in the involvement of all concerned parties: the contractor, the owner, the designer, and the workers. Designers have a role in construction safety by including safety considerations in design details and specifications, ensuring contractor’s submittals provide for worker’s safety, and inspecting construction sites for safety [5]. Owners, as well, are involved in securing the safest working conditions through the selection of safe contractors, including clauses related to safety in contracts, participating in safety management and governing project characteristics [6, 7]. And, contractors constitute a key element in providing a safe jobsite through implementing safe construction techniques, providing personal protective equipment, training workers and supervisors on site, and improving their safety practices through injury record keeping and learning from previous experiences.

This paper presents an assessment of the existence and implementation of safety practices in the construction industry in Lebanon through interviews conducted with several contractors practicing in the market today. This survey sheds the light on awareness of Lebanese contractors about the importance of safety on site, the common types of injuries and provided insurances, the adopted safety standards if any, the role of managers and employees in monitoring safety practice, and concludes by providing recommendations to promote a safety culture within the construction industry. This research is the preliminary phase of a wider scale study that is being conducted by the authors to survey all involved market constituents including designers, owners, developers, insurance providers, and concerned public authorities in order to form a clear idea about the main challenges to safety implementations. This study will be eventually used to develop a simple and cost-effective construction safety program tailored to the Lebanese market risks and conditions which can positively influence industry professionals and convince them of the necessity of a safe construction site.

2 The primitive safety culture in Lebanon and the surrounding region

The term “Safety culture” was first discussed by the International Atomic Energy Agency as being one of the causes of the Chernobyl accident in 1986 [8]. Safety culture was then defined in various ways, however all definitions focus on the individuals’ attitudes towards the safety system [9]. Construction safety culture is the combination of the beliefs, norms, attitudes and work practices aimed at reducing the vulnerability of workers and the public to accidents in the construction industry [10]. Commitment to safety, awareness about its importance, and enforcement of corresponding legislative measures varies significantly among different construction industries. In the United States, the Occupational Safety and Health Administration (OSHA) sets and enforces construction work standards to ensure safety of the workplace and to protect health of the workers. The agency with the help of the United States Environmental Protection Agency (EPA) monitors the compliance with OSHA regulations through regular inspections of construction sites and punishment of violations. As for the European Union member states, the European Union law enacted construction safety principles in a main framework in 1989 (Directive 89/391/EEC) to encourage both public and private sectors to improve occupational health and safety and enforce a dual responsibility on the employer to provide safe working conditions and on the worker to adhere to safety instructions and report potential hazards [11].

However, this is not the case in the Middle East region in general and Lebanon in specific where the construction industry is booming and providing lot of employment opportunities to labor force without any considerable attention to the safety of the site nor to the health of workers. In the UAE, for instance, the federal law applicable to all seven emirates addresses health and safety in general but does not include any specific legislation dedicated to construction safety. A recent study by the UAE University found that two-third of hospitalized injuries occurring on building sites resulted from lack of or inappropriate personal protective equipment [12]. Similarly, the construction industry in Kuwait was characterized by lack of safety standards and weak implementation of labor safety rights through examining data for accidents on construction sites for the last ten years. Although Chapter XV of the Kuwait private sector labor law gives public officials the right to inspect and enforce the law, most contractors agreed that such inspections never took place [13]. Another study conducted about the Egyptian construction market showed that workers on site do not undergo any training programs or safety meetings. It further showed that most companies do not include safety in their planned budgets while very few others do but only within the range of 0.5 to 1 % of the annual work volume [14].

Lebanon is a developing country that is still recovering from a 15-year civil war that ended in 1990 and after which the construction industry started booming exponentially. However, occupational health and safety was never given importance in that dynamic and hazardous field, and monitoring systems and relative legislation were never established by the government to control the inherent risks of that
profession. The Lebanese workers’ compensation law allows an injured worker on site full medical care, 75% of his daily salary since the time of injury and compensation in case of permanent disability or death. However, few construction companies purchase insurance tailored to protect workers and cover their injuries but rather the majority opts to provide an insurance policy for the whole worksite without naming individual workers or to pay from their accounts at the time of injury [15]. And here it is important to note that big portions of the labor force in the Lebanese construction sector are non-Lebanese who are hired on a daily basis which contributes to the negligence of employers in insuring them. From here stems the interest of this study in surveying current safety practices in the Lebanese construction industry through interviewing contractors to know more about their approach and commitment to an enhanced safety culture.

3 Survey Scope and Findings

The purpose of this paper is to assess the adoption of construction safety programs in the Lebanese construction industry. For this aim, a qualitative survey was conducted through interviews based on a questionnaire with fifteen construction companies. The interviewees occupied either the position of project manager or safety manager in their companies. All construction firms conduct work in Lebanon of which 33% have projects abroad as well. The fifteen interviewed companies have their headquarters based within the Lebanese territories. All firms are specialized in residential and commercial buildings with only three of them executing transportation projects as well.

The questionnaire was divided into nine parts which are the main safety areas identified by the Construction Industry Institute [16]: (1) Company Information: gathering background information about the company such as main scope of work, number of employees, whether or not it adopts a safety and health management system (SHMS), potential incentives to adopt a SHMS if adopted; (2) Management Commitment: assessing if top management participates in safety operations; (3) Staffing for Safety: allocated human resources for construction safety purposes; (4) Safety Education: the existence or lack of training programs for workers; (5) Worker Involvement: participation of workers in enhancing safety regulations and training new workers; (6) Punishment/ Reward System: assessing if any rewarding system is in place to promote good safety practice and punish non-conforming behavior (7) Subcontract Management: Assessing whether safety performance is a criterion in the qualification and subsequent supervision of subcontractors. (8) Accident Investigation: keeping accident records and investigation strategies (9) Drug and Alcohol Testing: determining whether workers are tested for alcohol or drug abuse.

3.1 Safety and Health Management System (SHMS)

As shown in Figure 1, 53% of the interviewed contractors adopted a SHMS. All large companies and 40% of the medium sized companies developed a SHMS. As it was expected, none of the small contractors has developed a SHMS. This may be explained by the fact that small contractors spend on safety a higher proportion of their profits than medium and large size contractors do [3]. 85% of the companies with no SHMS have less than 2 projects per year while 62% of the companies with a SHMS have more than 5 projects yearly. This proves again the direct relationship between a company’s annual turnover and adopting a safety and health management system.

![Figure 1. Distribution of companies by size and adoption of SHMS](image)

By classifying the companies based on their number of service years, 54.5% of the companies with more than 10 years in the market have adopted a SHMS, as well as 50 % of the companies who have now conducted construction business for 5-10 years. This shows that there exists no relationship between the age of the contracting company and its application of SHMS. It may be concluded that new companies are not shifting towards increasingly adopting SHMS despite the rise of awareness towards construction safety worldwide. One further observation is that all contractors working on international projects within the Middle East area have a SHMS, and only 30% of the companies solely working on local projects have a SHMS, all being large companies.
3.1.1 Companies without SHMS

As reflected in Figure 2, 71% of the companies with no SHMS claimed the reason to be that it is not required by law, 14.5% of them justified the absence of the SHMS through the high cost of safety implementation, and the remaining 14.5% answered that safety is a time consuming process. All respondents considered that the main incentive that would push them to develop and implement a safety program would be a strict legislative action by the government. All companies agreed that a citation in the range of 1,000,000L.L to 4,999,999 L.L ($667-$3330) would lead them to adopt safe work practices as it would be more cost effective. It is worth noting that the government did release a decree in 2004 to deal with occupational safety, yet none of the interviewed contractors including the ones with SHMS had knowledge about.

Figure 2. Reasons for lack of safety system adoption by contractors

3.1.2 Companies with SHMS

Companies based their SHMS either on Occupational Safety and Health Administration (OSHA) standards, requirements of the British Standards Institution (BSI), ISO 9001, on previous experience or on a combination of some or all. The tendency of construction companies to follow the OSHA standards (50%) more than BSI standards (25%) may be explained by the fact that American standards are viewed as less strict.

Figure 3. Safety standards adopted by Lebanese contractors with SHMS

Companies were asked about the main cause behind developing a SHMS. All respondents considered SHMS as an integral part of the company policy as opposed to being required by the owner or consultant. This implies that a company with SHMS will apply safety procedures even if not bound by contract.

The majority of companies gave priority to saving workers’ lives (63%) over saving the costs associated with a worker’s injury or death, as shown in Figure 4. All these companies indicated that they continually update their safety and health management programs following changes in project phase, incident or accident on site, update in reference standards or new site requirements.

Figure 4. Reasons for adoption of safety standards by contractors

3.2 Management Commitment

Based on the responses of interviewees reflected in Figure 5, about half of them only spend more than 1% of their project budgeted cost on safety programs which may reflect a lack of full commitment and dedication to safe practice from the top management side. However, all interviewed companies having a SHMS organized regular site visits for top management and allocated corresponding resources such as meeting time, offices and tools for safety purposes. Even two among the firms who did not implement any SHMS had administrative personnel visibly participating through site visits. The administrative personnel’s task was however limited to ensuring that no obvious unsafe behavior is present rather than checking the implementation of a fully developed safety system. Based on the aforementioned, it can be deduced that top management in general is willing to commit to safe work practices in terms of time and work, yet, they are not ready to allocate enough budget for this purpose.

Figure 5. Percentage of project budget allocated to safety programs
3.3 Staffing for Safety

This part of the survey deals with evaluating whether companies are allocating the required labor resources to properly implement safety procedures as part of the SHMS. It is obvious that companies with no safety and health management system reported having no Safety Officer (SO). As for the companies with a SHMS, seven out of eight had at least one full time SO. The company with no SO reported that the Project Manager (PM) is able to replace the SO. In order to confirm that the SHMS is properly implemented, site inspections are necessary. For instance, OSHA 1926 recommends regular inspections on Personal Protective Equipment (PPE), Personal Fall Arrest Systems (PFAS) and other protective equipment in addition to inspection on hazardous conditions similar to open excavations. Seven of the eight contractors with SHMS organized regular safety inspection and had allocated this task as part of the PM’s job or staffed a safety inspector for this purpose. One company rented the services of a specialized inspection firm to ensure the complete implementation of its SHMS. The company with no safety inspector had a safety officer and considered that once the safety manuals are developed by the officer foreman, gangers, site engineers and PM will make sure of it being implemented without the need for inspections. Both companies who didn’t staff a SO or a safety inspector had their SHMS based on personal experience and not on international standards. This may show the limitation of a safety program that is mainly based on experience from previous projects.

As part of accident management, contractors were asked whether they ensured the presence of a nurse or physician on site or at a nearby location. Obviously, none of the companies with no SHMS had a nurse or physician. Furthermore, two out of the eight contractors with SHMS staffed a nurse on site. One of the eight companies adopted a different approach by having an emergency team consisting of trained workers. The other five contractors didn’t even have the knowledge nor considered the possibility of having a registered physician at a nearby location to promptly deal with accidents when they occur.

3.4 Safety Education

One of the most effective elements to reduce the risks of a certain activity is the safety and health training of all labor resources in a project [17]. Training falls within the safety education along with safety meetings and safety signs and regulations. The main aim of safety education is to make sure the workers are aware of the site hazards and to train them to effectively deal with those hazards. Seven of the eight companies with SHMS provided training for their employees. The company which doesn’t provide training for its employees had a SHMS based on its previous experience. This shows the necessity of having a national standard that companies should abide by rather than having a SHMS based on their own experience or on a collection of standards. Two of the companies with no SHMS did provide basic safety training for their workers. The training mainly covered either safe equipment operation or safe work procedures. Training was provided by foremen or site engineers based on their experience rather than safety manuals. As for companies with SHMS, the training covered a combination of hazards, work procedures, PPE, and safe handling of equipment.

Safety meetings are also an essential part of the SHMS. However, only seven of the eight companies organized safety meetings. Only five of the eight companies included workers in their safety meetings despite the fact that workers are an essential entity in safety meetings. All companies with SHMS posted safety regulations and signs in their job site to raise awareness towards hazards. None of the companies without SHMS had signs and regulations on site.

3.5 Worker Involvement, Evaluation and Recognition, Drug and Alcohol Testing

The questionnaire further assessed tradesmen involvement in the SHMS through participation in setting safety rules, accident investigation, and safety training. 75% of the companies with SHMS involved their workers in setting rules while only 50% involved them in accident investigation. In response to the inquiry about why workers are not involved in investigation, a contractor affirmed that workers lack objectivity as they may either fear top management or fully sympathize with affected worker. Three of the eight companies involved its workers in safety training of other employees; all three being large size companies. Moreover, two of the companies with SHMS recognized the safe performance of their employees through both monetary and verbal rewards; two other companies provided monetary reward while the remaining four didn’t reward employees for applying safety rules. Even among the seven companies who lack a safety program, three of them did reward their employees for safe behavior. In case of safety related misconduct, all contractors with SHMS took measures up to termination of work. This shows a strict compliance to SHMS. Also, contractors without SHMS took measures against unsafe behavior as well; five of which reached job termination, one adopted decrease in monthly pay, and the last was limited to verbal warning. Although none of the
interviewed companies tested its employees for drug or alcohol abuse neither upon employment nor during work, a written document banning drug and alcohol abuse existed in workers’ contracts in 50% of the companies with SHMS. This shows the lack of awareness among Lebanese contractors about the importance of drug and alcohol testing for employees.

3.6 Subcontract Management

Hallowell and Gambatese affirm that subcontract management is among the most effective tiers of SHMS [18]. Only three companies with SHMS used safety performance as a prequalification criterion for the selection of subcontractors. However, all companies with SHMS bounded subcontractors by contract to follow safety procedures. This is not enough though to encourage subcontractors to be proactive in developing a safety culture and to be committed to safety training and education. As for companies with no SHMS, none of the companies considered safety neither in choosing subcontractors nor in contract with the latter.

3.7 Accident/Incident Investigation

As accidents and incidents indicate deficiencies in the SHMS, record keeping is essential to continually improve safety program and performance. All companies with SHMS kept record of accidents and incidents: six of which kept records for more than five years beyond the project completion date while the remaining two kept them for the duration of the project only. As for the seven companies without SHMS, only two kept records for more than five years and one for the duration of the project. Besides, only one of all interviewed companies which had no SHMS did not conduct investigation for causes of accidents when they occurred. Mostly the parties responsible for investigation were mainly the insurance representative assisted by the PM (in companies with no SHMS) or by the PM and Safety Officer (in companies with SHMS). This shows that companies with SHMS are serious about gathering data for the continuous update of their safety systems.

In response to the question regarding the most frequent type of injuries, nails in the foot and cuts were reported six times; eye splinters and fall from elevation appeared four times and object falling on head appeared twice. Having the fall from elevation as one of the most frequent accidents reveals the necessity of immediate action to enhance safety in the Lebanese construction industry.

Regarding the insurances obtained by all interviewed companies, they varied between all-risk, worker’s compensation and third party liability insurances. Almost all companies provided a combination of these except three of the companies with SHMS who purchased all three types simultaneously.

Among the eight companies with SHMS, seven reported conducting work site analysis and updating it in order to deal with job specific hazards. This shows that the majority of those companies is not following a predefined template but is being proactive in preventing accidents. As for contractors with no SHMS, only 57% conducted work site analysis which shows that PMs are improvising in the situation where no safety program exists to limit on site hazards.

4 Conclusion

Based on the presented survey findings, it can be concluded that there is a promising effort on behalf of mostly large Lebanese contracting companies to address safety concerns in their organizational policies through following international safety standards such as OSHA and BSI with continuous follow-up and updates. They also reported training workers regularly, assigning safety officers and inspectors on site to make sure their safety programs are implemented correctly, posting safety regulations and signs in their jobsites, and keeping track of accidents and investigating their causes. However, and although these companies showed concern for construction injuries and commitment to safe practices, they had allocated a low budget for safety in their projects, missed some other safety measures such as rewarding workers for compliance with safety and running alcohol and drug abuse tests on workers. Besides, 25% of these companies based their safety programs on previous experience which calls for the importance of drafting and legislating national safety standards that can be tailored to the Lebanese construction hazards and that can be applied homogenously across all construction firms, which would facilitate the process of monitoring compliance with those standards and penalizing violations.

On the other hand, the majority of medium and small construction companies did not have any safety management systems implemented on their job sites, neither safety officers nor trainings or signs. These firms reduced the whole safety program to just mere visits by administrative personnel in order to check that no unsafe practices are occurring during the course of work. These firms, however, showed willingness to implement safety regulations if they were enforced by law. And this highlights the role of the public health sector and all other relevant public authorities in raising
awareness about construction safety among all industry constituents through establishing formal monitoring bodies whose function is to ensure implementation of safety practices on construction sites, to regularly inspect workers and keep track of work-related injuries, and to devise a set of significant violation fines that would push contractors to adhere to the safety law.

Last but not least, coordination and commitment of all construction parties including owners, designers, contractors, subcontractors, managers, site engineers, and workers is crucial to the success of any safety program and necessary to face any safety implementation challenges in the construction industry.

References


5D-BIM: A Case Study of An Implementation Strategy in the Construction Industry

Anoop Sattineni a and Jennifer A. Macdonald b

a Auburn University, Auburn, AL, USA
b University of Technology, Sydney, Australia
E-mail: anoop@auburn.edu, Jennifer.Macdonald@uts.edu.au

Abstract -
Several factors have contributed to the growth of BIM usage in the global construction industry, including availability of appropriate software and hardware tools, the opportunity to minimize errors, waste & cost and increasingly competitive markets [1][2]. An emerging approach in the construction industry is the use of 5D-BIM, by combining the traditional three dimensions of a BIM with the schedule as the fourth dimension and cost estimate as the fifth dimension [3]. This approach allows the contractors to better predict the cost of the project, the time-line of the project when these expenses are anticipated to occur, while simultaneously allowing the schedule to be optimized by considering the quantities of materials derived from the model and the productivity rate of construction crews. On the one hand 5D-BIM methods provide an excellent opportunity to connect processes in design, cost and construction methods; on the other hand they call for a significant shift in the way construction companies operate. It is unrealistic to commit extensive resources for a large construction company, towards a re-alignment of their internal processes, while continuing to be competitive and profitable. This paper considers how one construction company in the United States adopted the 5D-BIM methods, the challenges faced in implementing it within the company and the lessons learned in the process. A case-study method was used in an effort to understand the paradigm shift within the company in adopting 5D-BIM. Interviews with key personnel within the company were conducted and content analysis of the data was performed to describe the results. The results from this study show the intricacies of implementing a technology driven paradigm shift to a population of construction experts with a not-so erasable philosophy of success within the construction industry. The captains within the company communicated a great appreciation for the 5D-BIM concepts but were unwilling to delve into it whole-heartedly. Several reasons contributed to this attitude, including a reluctance to change by some, motivation to demonstrate immediate profitability and the lack of demand for such endeavours from the owner/designer community. However there were also personnel within the company that quickly adapted to the 5D way of thinking about construction and were keen to proceed with using the methodology on future projects. The advantages of using 5D-BIM and the problems encountered in implementing it are analysed in the results section of the paper.

Keywords -
Construction Management, 5D-BIM, Case Study, 5D-BIM Implementation

1 Introduction
There is much misconception regarding BIM and what it exactly entails. The term, Building Information Modelling (BIM) has been used as a label for the 3D geometric renditions created by modelling software such as Autodesk REVIT and ArchiCAD, among several others. In reality, the 3D geometry only represents a portion of what BIM really is and its full potential.

There are many definitions of BIM that have been developed through its evolution. The most descriptive of these definitions is from the buildingSMART Alliance through the National BIM Standard – United States. “A BIM is a digital representation of physical and functional characteristics of a facility. As such it serves as a shared knowledge resource for information about a facility forming a reliable basis for decisions during its lifecycle from inception onward”

In the design and construction phase, BIM can begin be understood as a new process of delivering information and knowledge of a project to all the relevant parties. The information that is assimilated into a Building Information Model is provided by consultants, designers, manufacturers and engineers. This creates a new level of collaboration that the construction industry does not currently operate at. Thus, BIM is creating a new workflow from preconstruction
to post construction and beyond through a project’s entire lifecycle.

There is general consensus in the academic and the AEC industry community that BIM is more than just a software tool. The true impact of BIM is the enabling of comprehensive, digitized and collaborative processes within the AEC industries. The American Institute of Architects (AIA) recognized this seismic shift in the process of designing and constructing facilities and has proposed a new delivery method, namely ‘Integrated Project Delivery’ (IPD) [4]. BIM is the cornerstone of realizing the benefits proposed in this new methodology for project delivery [5]. The use of BIM in the US construction industry and academic research regarding BIM has been on the rise and is predicted to continue to increase [6].

1.1 4D – BIM

The fourth dimension added to the traditional 3D-BIM model is ‘time’. This allows the 3D parts or assemblies inherent to a BIM model to be combined with a construction schedule. The resulting 4D model is capable of creating virtual simulations of building the BIM model [7]. These simulations can be used to communicate the sequence of work to avoid rework in the construction industry as well identify, ahead of time any areas that may be congested due to too many trades being in one place at the same time. Congestion of various trades in one location within a construction site is considered a safety hazard.

Coordination between trades has become very important in the construction industry due the specialization involved in each trade. One of the advantages of using BIM in the construction industry is to optimize the space in the building where the services such as HVAC, water, waste and electrical systems are located. This can potentially reduce the amount of space needed for the various systems and hence reduce the overall height of the building, reducing the overall cost of the building. Under such circumstances, contractors installing these systems have to carefully coordinated so that each team has sufficient time and space to install their respective materials. By simulating the sequences of construction 4D-BIM can help optimize the process and allow for minimal rework and changes. 4D-BIM can also allow construction managers to optimize the movement of materials, equipment and people on a construction site ahead of time. This is depicted by the image in figure 1 created by HOAR Construction Company in Birmingham, Alabama. Hu [8], declares that the use of the fourth dimension is fundamentally changing the project planning, design and construction management strategies of companies.

1.2 5D – BIM

The fifth dimension added to the 4D-BIM model is ‘cost’. A 3D-BIM model has all the geometrical information needed to perform a take-off of material quantities. A 4D-BIM model has all the activities needed to complete the project. By attaching a cost database to the 4D-BIM model and by assigning actual costs to materials, equipment and personnel, a 5D-BIM model can be created to provide the construction team with a useful tool[9]. The implementation tools for implementing a 5D-BIM model have only been available for less than a decade.

The resulting 5D-BIM model can be used by construction professionals to give faster feedback about the cost of a project, allowing the designer to adjust the project design to fit the budget. The opposite scenario is often seen in the construction industry in that the project cost is estimated to be over what the owner can afford and consequently resulting in significant design changes and also perhaps losing the most attractive elements of the design. The 5D-BIM model can provide the owner and the design team with greater transparency in seeing the contractor’s budget, building confidence within the various stakeholders of a project. Projects using 5D-BIM would also mean there would be a less of a need for a large contingency on behalf of the owner.

2 Methodology

HOAR Construction Company based in Birmingham, Alabama in the south eastern United States was chosen for the purpose of conducting this research. The ‘Virtual Design and Construction’ (VDC) team members were interviewed for the purpose of
conducting this research using a case study methodology.

A case itself can be defined as a single person, subject, group or organisation [10]. Case study research may involve the investigation of a single case or multiple cases and can be categorised as ‘descriptive’, ‘explanatory’ or ‘exploratory’ in nature. A descriptive case study is used to describe a phenomenon or a processes whereas an explanatory case study is usually theory driven and may be used to develop hypothesis in a large research project [11]. An exploratory case study is typically used to test hypothesis to come up with logical conclusions [12]. In this study a descriptive case study is presented in regards to the implementation of BIM at HOAR Construction Company.

Saunders et al. [13] draw distinctions between three types of interviews, namely ‘Structured Interviews’, ‘Semi-Structured Interviews’ and ‘Unstructured Interviews’. Structured Interviews use a single set of pre-determined questions and the interviewer would read each question and record the answers for each of those questions. Structured interviews are used to collect quantifiable data and are referred as ‘Quantitative Research Interviews’ [13]. ‘Unstructured Interviews’, also referred to as ‘In-Depth Interviews’, on the other hand, are informal conversations where the interviewer does not have a pre-determined list of questions but instead is guided by a topic and asks questions allowing the interviewer to answer freely about events, behaviour and beliefs [13]. ‘Semi-Structured Interviews’ can be considered as occupying the middle ground between ‘Structured Interviews’ and ‘Unstructured Interviews’. In this format, the interviewer is guided by a list of themes and questions, however the order of questions may be different and some questions may be eliminated altogether, while new questions may be required, depending on the context [13]. The interviewee in a semi-structured interview has more latitude to answer questions in an open manner, however the researcher may guide the whole processes in order to get responses to all questions fully and in a timely fashion. The use of interviews to collect data is an acceptable method in social research [14]. The VDC team members were interviewed for the purpose of conducting this research, in the semi-structured interview format.

Interview questions relating to 5D-BIM implementation at HOAR construction were asked of BIM-team members. Questions focused on a historical perspective of BIM at HOAR construction, current use of BIM, their foray into implementing 5D-BIM, support and push from upper management within the company for implementing 5D-BIM, reaction from field personnel and senior field operations management team towards implementing 5D-BIM, surprises and challenges of implementing 5D-BIM and the future of 5D-BIM implementation at HOAR construction company were explored.

The data from the interviews was transcribed verbatim and analysed using qualitative data analysis. ‘Content Analysis’ and ‘Thematic Analysis’ are two common methods of analysing textual data. Content analysis is a qualitative method of analysing raw data as is thematic analysis, however thematic analysis only considers the qualitative nature of the data whereas content analysis generally results in quantifying the data by counting the frequency of the emergent themes and codes [15][6]. At the core of both content and thematic analysis, data is arranged into themes or codes by categorising raw text under meaningful labels or codes. While content analysis considers the frequency of the occurrence of codes to describe the data, thematic analysis interprets the data by analysing the meaning of the codes within that context [16]. The data collected in this research from the semi-structured interviews was analysed using the qualitative technique of thematic analysis by creating codes and labels. The findings of the interview data are grouped under appropriate themes and presented in the ‘Results’ section of this paper. The ‘Conclusions’ section presents the complicated nature of 5D-BIM implementation in the construction industry.

3 Findings

With over 300 construction professionals, HOAR Construction Company had an annual volume of construction of approximately $470 million USD in 2013. In 2013, HOAR was ranked 127 in the list of top 400 construction companies in the United State. In 2012 the company was ranked 107 in the same list and over the past decade has consistently stayed in the list of top 400 construction companies in the United States. These rankings are provided by the Engineering News Record (ENR) magazine. ENR is considered the premier trade magazine for the construction industry in the United States. HOAR as a company works in the civic & government, retail & mixed use, office & commercial, education, residential & hospitality, concrete & heavy-civil, industrial & manufacturing and healthcare sectors within the construction industry. Apart from being a construction company, HOAR also has a program management side of their business in which they advise owners about issues related to construction projects. Hoar construction has a dedicated staff of six people in their VDC team.

The VDC team was interviewed for the purpose of
conducting this research. The findings of the interview results are analysed using thematic analysis and presented under appropriate themes within this section.

### 3.1 BIM Adoption & 5D-BIM Exploration

BIM was first explored by one of the current team members in 2006, using the Autodesk Revit software. The program was only used to consider a BIM model to count the number of various types of doors present in it. At this stage the adoption of BIM was also in the very early stages within the construction industry[17]. As the team started to receive more models from design teams and with the evolution of BIM software within the marketplace, the adoption of BIM increased at HOAR construction. The company also has a long history of working with the ‘Walt Disney’ group on its construction projects at Disney’s various theme parks. Construction projects at these theme parks were described as futuristic and the owners encouraged and often required the use of latest technologies in design and construction of its various facilities. This also turned out to be a key factor in the evolution of BIM within HOAR Construction Company. The first 5D estimate was created on a Disney project in 2008. However, the use 5D models did not immediately involve other projects within the company. In fact a VDC group did not emerge within the company till 2011. A full team of six professionals are now in the VDC group. As of 2013, they have had five years of engagement with 5D models and three of those years were described as dedicated engagement with a full-fledged team. Apart from the cost of personnel for the VDC team, it was described that over the last three years approximately $0.5 million USD were spent on software & hardware costs related to BIM. These include the costs for training the personnel within the VDC group.

![Figure 2. Use of 3D Clash Detection & 4D BIM co-ordination tools at HOAR Construction Company](image)

Prior to the full-scale adoption of 5D-BIM the company used the 3D clash detection tools to identify conflicts in the design of various systems such as HVAC, electrical, fire-proofing etc., as shown in figure 2. The company also used 4D-BIM tools to virtually simulate the construction sequences to identify and resolve any problems in the planning phases of most large projects. These 3D clash detection and 4D-BIM sequencing processes are still in use within the company. The company is currently starting to use 5D-BIM on more projects and there is general optimism within the company about changing to these new processes. However 5D-BIM is currently not used on all projects as a limited number of staff knowledgeable enough in 5D-BIM is available to implement it on a company wide scale.

### 3.2 Internal Support for 5D-BIM Adoption

The use of 5D-BIM in the company was prompted by the senior management within the company. It was described that the vice-presidents were hearing about these concepts at various venues and industry gatherings and were keen to explore them within the company. The formation of the VDC group was a clear indication by the management to make real investments to incorporate BIM based processes within the company. The managers were particularly keen to see real-time cost data as design progresses, real-time quantities as design changes and real-time updates to project budgets. They actively recruited talented young people who could be trained in this area as well moved some individuals within the company in to the VDC department. In an effort to incorporate the cost estimating processes in to the 5D-BIM workflow, some estimators also joined the VDC team. The company started with some pilot projects where the cost estimation was performed using both 5D-BIM tools and conventional tools. These pilot projects were done to identify the processes that would need to change within the company to fully implement 5D-BIM. At this stage there was some push back from the managers and it was decided to slow down the adoption process. This was done as the original timeline for adoption would have meant radical changes and might have resulted in some unforeseen situations.

One example that was cited was that of the estimating database. The estimating database was created based on the 1995 ‘Master Format’ of sixteen divisions, as specified by the Construction Specifications Institute (CSI) [18]. These sixteen divisions divide the materials and processes needed to build a facility with appropriate sections and sub sections, and allow all stakeholders involved to organize their data in a common format. However the software chosen for 5D-BIM, ‘Vico’, used ‘Unifromat’ which is
a format presented by the American Society for Testing and Materials (ASTM) [19]. ‘Uniformat’ is based on a system to enable construction project information to be classified from a life-cycle point of view. The two systems differ significantly from one another and adoption of 5D-BIM would have meant a new estimating database would have to be created. An estimating database is a key tool for any construction company to come up with an accurate cost estimate. The database is constantly updated with historical actual costs of construction and is often considered a confidential tool. It was decided by senior managers that it would be a risky move to switch databases as several projects were using the existing database, moreover they were not completely sure that all the information could be transferred in to the ‘Uniformat’ database. Overtime, however the VDC team was allowed to create a parallel ‘Uniformat’ database. The new database is currently being used for 5D-BIM projects whereas with the existing database is being used for projects that do not use 5D-BIM. It was described that the senior management within the company is starting to realize the benefits of using 5D-BIM and are again proponents of its growth within the company.

The reaction from the field personnel was described as mixed. The newer generation of field personnel were more open to using the new tools but there was some initial resistance from experienced field personnel. This resistance was described as an attitude of reluctance to change. However, over time it was described that everyone was on board with using 5D-BIM tools and adapt their processes.

As previously described, the estimating database was a key issue and several similar issues had to be encountered with making the switch to 5D-BIM. Another key issue was the use of ‘Location Based Schedules’ also called as ‘Flowline’ methodology for scheduling. This was a departure from the traditional ‘Critical Path Method’ (CPM) for scheduling construction activities. Flowline methodology for scheduling represents a significant departure from the CPM method of scheduling and field personnel had to be trained to adopt the new method. However, it was found that several field personnel within the company liked the new method of scheduling and quickly adopted it. Some personnel even went to the extent of using the Flowline method of scheduling that were not using 5D-BIM.

Overall the VDC team had to deal with waning enthusiasm from various quarters within the company but were able to successfully convince senior managers and several field personnel to adopt the new methods.

### 3.3 Challenges to Implementing 5D-BIM

The VDC team expressed several technical and human challenges that had to be overcome to implement 5D-BIM. There currently is no single software available in the market that can perform all functions that BIM is capable of enabling. Therefore while incorporating 5D-BIM the team also had to assist in several projects that were using for 3D clash detection and 4D simulation purposes. 5D-BIM projects generally cost more and take longer to prepare in the initial implementations as more time is spent by key personnel within the company, as compared to traditional projects. Due to the higher costs incurred, it was also more challenging to convey the value of adopting 5D-BIM to senior managers within the company. The VDC team had to overcome the attitudes of people towards change in general when incorporating 5D-BIM within the company.

The 5D software itself was described as robust but complicated and cumbersome to use. The software programs that enable 5D-BIM require a high level of technological expertise as well as expertise with construction processes. They described an occasion when an experienced team member unintentionally over-wrote the estimating database with no recourse to restoring the previous working version. On that occasion the team was forced to switch to the newer version of the database in a very short time period, while several active projects depended on the database.

There are also some gaps within the industry to implementing a successful 5D-BIM environment within the company. The ‘Level of Detail’ (LOD) used in the creation of a model by the design team may not be sufficient to extract accurate quantities for cost estimation purposes. The models may have errors in them with elements not touching or duplicate elements and models need to be checked to ensure that estimates can be accurately extracted. Initially in some cases designers were not willing to share models with the construction team.

A majority of the work on construction projects within the company was outsourced to subcontractors. Therefore the scheduling of construction activities depended on knowing the productivity rates for subcontractor crews. Some subcontractors were unwilling to share this information with the construction team. Furthermore there were no guarantees that subcontractors would send the necessary numbers of crew members to the site on a given day as the subcontractors themselves are often trying to meet obligations on multiple projects. This meant that updating the schedule and coordinating with various trades was complicated. However, it must be noted that
this practice continues in the industry regardless of whether it is a 5D-BIM project or otherwise.

Figure 3. Use of 5D BIM tools at HOAR Construction Company

3.4 Future of Implementing 5D-BIM

The VDC team expressed that the company was starting to reap the benefits of implementing 5D-BIM as they never before had so much data about their own internal processes. An evaluation of process had to take place so that 5D-BIM could be implemented and processes were streamlined to better fit the new workflow. One of the team members expressed that “It is possible to teach an old dog some new tricks and our estimators now have a higher opinion to the whole 5D-BIM concept”. The team expressed hope that the software would get easier over time. It was also remarked that 5D-BIM, like other software tools has to be web-enabled to ensure that field personnel are able to extract maximum benefit from it. The VDC team mentioned that they are able to provide much faster feedback to their design counterparts for cost planning purposes. The team is currently exploring the possibility of a tie-in between their 5D-BIM and their web-based project management software. The team has matured its processes to now account for ‘change orders’ that occur downstream in a project in the 5D-BIM environment. There is definite ‘no turning back now’ mood within the company about implementing 5D-BIM. The team itself could be re-aligned so that instead of working on several projects in a given year, each team member would get more immersed in one project and see it through key execution stages and consequently have no more than two or three projects in a year. As mentioned earlier not all projects within the company use 5D-BIM but it is hoped that more construction professionals within the company would be trained in it and would be using it on future projects.

4 Conclusions

Implementing 5D-BIM in a large construction company has several challenges associated with it. There exist technological challenges that have to be overcome to implement 5D-BIM. These projects generally require an experienced construction professional who also has a high level of talent on the technology side. There are significant hardware, software and training costs associated with adopting 5D-BIM processes. It is also challenging to overcome the human reluctance towards change on part of some personnel within a company. Training existing personnel who have a wealth of industry knowledge to adopt new methodologies for estimating and scheduling can be challenging. This issue is further complicated by construction companies having to remain profitable while exploring these new technology enabled processes. The process of implementing 5D-BIM involves a close inspection of internal processes with a company and can be beneficial to streamline them. Increasingly designers are sharing valuable information with the construction team by means of BIM models. 5D-BIM represents a clear departure from the existing methods of planning construction projects and can potentially be a transformative experience for a construction company.

References


Simulation of On-Shore Wind Farm Construction Process in Lebanon

E. Zankoul\textsuperscript{a} and H. Khoury\textsuperscript{b}
\textsuperscript{a,b} Faculty of Engineering and Architecture, American University of Beirut, Lebanon
E-mail: efz00@aub.edu.lb, hiam.khoury@aub.edu.lb

Abstract -

For the past thirty years, Lebanon has been experiencing cuts in electricity, thus compelled to import it and use noisy and extremely unhealthy generators. This happened at very high prices with even mediocre quality. Hence, an adequate solution to this shortage in electricity supply can be achieved through the use of renewable energies, in particular wind energy or power produced from wind farm turbines. However, the on-shore wind farm construction process can be a very complicated task due to several reasons such as the challenging topography of the rural areas and the absence of paved roads where they are typically constructed, as well as the impact of wind on the construction process. In order to address these complexities, this paper takes the initial steps and presents work targeted at efficiently designing and planning the construction process of an on-shore wind farm in the region of Falougha, Dahr El Baydar, Lebanon. The solution to the problem is described in detail using a discrete-event simulation model developed in AnyLogic. The developed work illustrates the different construction stages from rough grading, access roads construction, foundation and electrical works, to wind tower assembly and erection. The whole process is then optimized to mainly minimize the project duration. The components of the proposed model have been created and preliminary results highlighted the potential of using AnyLogic for simulating and optimizing complex construction processes offering unique challenges such as those found when constructing on-shore wind farms.

Keywords –
Construction Management, Wind Energy, Wind Farm, Simulation, Optimization, AnyLogic

1. Introduction

Since the 1970s, large interconnected wind-driven turbines and generators have been constructed in many countries in “farms” to generate electricity [1,2]. However, Lebanon is considered behind in terms of wind power energy [3]. Current wind energy sources available in Lebanon constitute mainly of some wind turbines and these are rarely functioning. It is believed that an extensive development of wind energy in Lebanon, i.e. wind farm construction, can play a crucial role towards resolving current electricity supply shortages [3]. While some of the issues associated with other types of construction projects are common to these facilities or plants, large wind farm construction is relatively new in Lebanon and offers unique challenges.

Therefore, this paper takes the initial steps and presents work aiming at efficiently designing and planning the construction process of an on-shore wind farm in Lebanon. This will be mainly achieved through the use of Discrete Event Simulation (DES) techniques. Previous research efforts have focused on simulating typical construction processes using DES tools, in particular Stroboscope [4, 5, 6]. Some have even attempted at simulating wind farm construction operations [7,8]. However, this was done for the case of Egypt presenting different challenges, taking into consideration different parameters and using the Stroboscope software [4].

In this paper, the on-shore wind farm construction process will be modeled as well, but for a specific challenging site in Lebanon, i.e. Falougha, using another simulation software, namely AnyLogic 6.9.0 (Educational Version) [9] and by setting up different conditions and parameters. The overarching goal is to develop a generic and dynamic discrete-event simulation model that can be adopted by contractors in order to plan and optimize the construction process of on-shore wind farms in Lebanon and elsewhere.
2. Methodology

In order to achieve the main objective, the paper addresses issues in two specific task areas. The first task consists of identifying the phases and activities of the whole wind farm construction process and determining the activities’ typical required resources and daily outputs. The second task comprises designing the simulation model of the construction process using AnyLogic [9] then optimizing the whole process.

2.1. Construction Process Activities

The construction process of on-shore wind farms can be divided into five main packages. [1,2,8] The first one consists of conducting a topographical survey to the land where the wind farm will be constructed. Next, the internal roads are constructed and prepared so that trucks and cranes can move around the site to complete the other construction steps. The third package consists of constructing the turbines’ foundations. The fourth one entails excavating the electrical trenches as well as loading the electrical cables. Finally, the last package involves erecting the wind turbines.

In the next subsections, daily output values and corresponding crews were first extracted from the RS Means Building Construction Cost Data (71st edition) book [10] then adjusted according to construction practices in Lebanon.

2.1.1. Topographical Survey

The resources required for this activity consist of a surveying crew with a laser transit and the daily output is 13,355 m²/day.

2.1.2. Roads Construction

The roads construction package can be divided into two main parts. The first one consists of cut/fill operations to obtain the required slopes and widths of the roads on which trucks will be operating between turbines. The activities involved are: cutting soil, hauling, filling soil and compaction with typical daily outputs of 92 Bank m³ (BCM), 240 BCM, 765 Loose m³ (LCM) and 990 Compacted m³ (CCM) respectively. A truck is needed for all cut, fill and hauling operations. Besides a truck, a cut operation requires an excavator while a fill operation requires a loader. Compaction’s only resource is a compactor. Note that all vehicle-type resources involve an additional resource, i.e. an operator.

The second part consists of constructing the roads once cut and fill is complete. The activities involved in this part are overlaying, grading, watering, aggregate base compaction and testing compaction with daily outputs of 570 m³, 335 m³, 2500 m³, 990 CCM and 32 tests respectively. A road crew is required for the first three activities, in addition to an aggregate truck, a grader and a water truck respectively. Aggregate base compaction requires a compactor which is the same one required for fill compaction.

2.1.3. Foundation Construction

Foundation construction consists of steel reinforcement installation, formwork and pouring concrete and curing. Their respective daily outputs are 3.2 tons, 85 m³ and 510 m². A foundation crew is needed for the first two activities. A concrete truck is additionally needed for the second and two common laborers for curing.

2.1.4. Electrical Works

This package consists mainly of two activities; loading electrical cables with a daily output of 110 m and excavation of electrical trenches with an output of 306 BCM. Loading cables requires a main crane and an electrical cable truck. Electrical trenches excavation requires an electrical crew as well as an excavator.

2.1.5. Turbine Installation

A modern wind turbine is made of three main parts: the tower, the nacelle and the blade hub connecting all three blades as shown in Figure 1.

![Wind Turbine Components](image)

Figure 1. Wind Turbine Components

The first activity of this package consists of placing the tower at the assigned turbine location according to the as-designed layout. It is worth mentioning that the success of wind farm projects depends on precise turbine placement because meters can mean megawatts.
The second activity consists of tilting the nacelle, lifting, positioning and bolting it. The same series of activities is then repeated for the hub connecting all three blades. Four towers can be loaded and four nacelles can be tilted daily. Lifting and positioning either the nacelle or the blade hub occurs at a daily output of 2, while bolting occurs only at 1. A main crane is required for each of these activities except for bolting which only requires a bolting crew. This crew is also needed for positioning the nacelles and hubs. In addition to those resources, a secondary crane is required for nacelle tilting and a tower truck for tower loading.

2.2. Simulation Model

The construction process was modeled using AnyLogic 6.9.0 (Educational Version) software. It is commonly known for being the only simulation tool that brings together Discrete Event, System Dynamics and Agent Based methods within one model development environment [9].

The five activities or packages previously described in Section 2.1 were modeled using the Discrete Event Simulation component of AnyLogic. Their processes are depicted in Figure 2 and referred to by 1, 2a-2b, 3, 4, and 5 respectively.
The main elements used to model the wind farm construction process are the following: Source, Sink, Service, Delay, Resource Pool, Seize, Release, Select Output, Batch, and Copy.

The Source generates resources (entities) and is used in our model to represent the main resource(s) for each phase, e.g. site, road, soil, foundations, cables, turbines, etc. The most important property of the Source element is the arrival rate. The software offers an option to inject entities to that source when a certain condition occurs. In the case of this model, the conditions depend on the activities' sequencing.

The Sink is used as a discharge point for the entities generated by the Source. In this model, sinks are found at the end of each phase. These elements do not have any specific property. However, they are important to place a condition on when the run should end. This is defined in the model as follows:

```java
if(CutFillDone.count() == CutCM+FillCM)
    phase2 = time();
```

Regarding sequencing, the model contains eight Sources. In order to know when to inject the required quantities in each Source, sequencing the activities is required. In fact, as mentioned above, a Source has a property called arrival rate and in this case it will be defined manually, i.e. conditions were put on how many resources are available and at what times. For instance, at time 0, the whole site area is injected into “Site1” (Part 1 in Figure 2). Once surveying is done, all cut and fill material is injected into “CutRoadSegments” and “FillRoadSegments” (Part 2-a in Figure 2) as follows:

```java
if(Site2.count() == SiteAcres)
    CutRoadSegments.inject(CutCM);
    FillRoadSegments.inject(FillCM);
```

Using a similar syntax, the rest of the activities were sequenced by setting, for each, different conditions. When all the soil is cut and/or filled as required
(“CutFillDone”, Part 2-a in Figure 2), the road construction begins (Part 2-b in Figure 2). On the other hand, foundation construction (Part 3), cables loading and excavation for electrical trenches (Part 4) can begin simultaneously when the roads are ready signaled by “FinishedRoad”. Each time a foundation is complete, i.e. “FinishedFoundations”, the installation of one wind turbine begins and proceeds with all aforementioned stages in Subsection 2.1.5 (Part 5).

3. Case Study and Results

3.1. Falougha Case Study

The site selected for the case study is located in the region of Dahr El Baydar, Lebanon, precisely in Falougha. The choice of this particular site at the coordinates 33°49'4.36", 35°45'41.20" was made due to several reasons, mainly because it is a public property and it is isolated thereby preventing the high voltage effects from harming people. It is also a non-vegetated open space with neither dwellings nor infrastructure. Above all, the wind power in the area is great.

As a first approach to the site investigation process, an examination of the geological map and the Google Earth view of the site were done. From the geologic maps of Lebanon and the geologic study done by the American University of Beirut on Lebanon [11], it was found that the site is located above lower-mid cretaceous formation which is mainly composed of a thick layer of limestone. The water table is at a depth greater than 200 meters [11]. It was thus decided that the site is adequate for wind turbine construction.

In order to determine the total site surveying area, the Falougha region was located on Google Earth then loaded into AutoCAD. The area was found to be 429 acres. Given an uninhabited and relatively large site, large turbines were selected namely Vestas V80 with a power of 2 Megawatts (MW) [12]. Each of its 3 blades is 39 meters long with an 80 meters rotor diameter. The tower is 65 meters high and the total weight of the turbine is 230 tons [12]. In this case study, 30 V80 turbines were distributed over the area in question and arranged according to the layout in Figure 3. The detailed analysis regarding the choice of turbines type, their number and their layout was carried out separately and is, however, not included in this paper.

Based on this turbine layout, road paths were then designed (Figure 3). For the roads, it is essential to start by drawing the horizontal and vertical alignments. In order to achieve that, certain specifications were followed. For instance, the maximum allowable longitudinal slope for such a project is 10%. [13].

372
As for foundations works, a typical foundation design for Vestas V80 wind turbines [14] was adopted and is shown in Figure 5. Using these dimensions, the volume of concrete and area of curing for each of the 30 foundations were calculated, 320 m³ and 256 m² respectively. To find the quantity of steel needed, a unit weight value of 130 kg/m³ (typical for heavy industrial projects [15]) was adopted. It was found that the weight of steel needed for each foundation is around 42 tons.

![Figure 5. Turbine Foundation Dimensions](image)

According to the total road length computed earlier, the length of electrical cables should be 6.6 km. Assuming that the trench is 2 m wide and around 2 m deep, the total quantity to be excavated is 28,165 m³.

### 3.2. Results and Evaluation

Based on computed daily outputs (Subsection 2.1) and Falougha’s specific construction works quantities (Subsection 3.1), the generic simulation model (Figure 2) was run, with only one resource of each type, and preliminary scheduling information (Figure 6) together with resources’ utilization rates (Figure 7) were obtained.

At a first glance, it is apparent that the total project duration (2,228 working days) is unreasonable and utilization rates of the different resources are very unbalanced. In fact, eight resources have a utilization rate equal or very close to 1 while, seven other resources have a utilization rate ranging between 0.019 and 0.109. This implies that some resources are idle most of the time while others are fully active. Therefore, the number of resources needs to be efficiently chosen according to site and operations requirements. For instance, the resources “Excavator” and “Loader”, with a relatively low utilization rate, depend on the resource “Truck” (utilization rate>0.92). In other words, these resources are idle for long periods of time due to the lack of trucks. Increasing the number of trucks will thereby increase the utilization rate of “Excavator” and “Loader”.

![Figure 6. Gantt chart of the Falougha’s Wind Farm Construction Process (in working days)](image)

It was thereby imperative to optimize the process by minimizing time while keeping the utilization rate of all resources as high as possible. For instance, optimizing the number of resources used on more than one activity, such as “Main Crane”, is of paramount importance.

Using trial and error, the model was rerun to find ranges of optimum number of resources. AnyLogic 6.9.0 (Educational Version) offers a type of experiment called...
“Parameters Variation”. It allows the user to select ranges and a step for different parameters and get results for all possible combinations falling under these ranges. Independent phases were optimized separately by varying only their related resources while dependent phases or activities were optimized together, each time refining the ranges based on the results. For each variation, the duration and utilization rates were monitored. After several runs, the combination of resources leading to an optimal scenario was found to be as follows: six survey teams, 10 excavators, two loaders, 19 trucks, two compactors, nine aggregate base trucks, seven graders, one water truck, 16 road crews, 16 foundation crews, four concrete trucks, three electrical crews, three excavators for trenches, two electrical cable trucks, one secondary crane, five main cranes, two bolting crews for nacelles, two bolting crews for blade hubs, and one tower truck. This combination of resources led to the following Gantt chart (Figure 8) and utilization rates diagrams (Figure 9). The results have greatly improved; the duration has significantly decreased while the utilization rates have increased and become more leveled.

Figure 8. Optimized Gantt chart of the Falougha’s Wind Farm Construction Process (in working days)

Figure 9. Optimized Resource Utilization Rates of the Falougha Wind Farm Project

Table 1. Comparison between Initial and Optimized Duration Results (in working days)

<table>
<thead>
<tr>
<th></th>
<th>Surveying</th>
<th>Cut &amp; Fill</th>
<th>Road</th>
<th>Foundation</th>
<th>Loading Cables</th>
<th>Electrical Trenches</th>
<th>Turbine Installation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>130</td>
<td>1,584.28</td>
<td>58.27</td>
<td>450.5</td>
<td>67.78</td>
<td>92.1</td>
<td>427</td>
<td>2,228</td>
</tr>
<tr>
<td>Optimized</td>
<td>21.82</td>
<td>85.26</td>
<td>11.88</td>
<td>34.5</td>
<td>32.61</td>
<td>30.7</td>
<td>27.83</td>
<td>162.3</td>
</tr>
<tr>
<td>Ratio (I/O)</td>
<td>6</td>
<td>18.6</td>
<td>4.9</td>
<td>13.1</td>
<td>2.1</td>
<td>3</td>
<td>15.3</td>
<td>13.7</td>
</tr>
</tbody>
</table>

Table 1 depicts the initial scenario where only one resource of each type was applied in comparison to the optimized scenario. In this case, all activities durations are shorter and the resulting project duration is around 163 working days (7.5 months assuming 5 working days per week) compared to 2,228 working days (equivalent to approximately 8.5 years). It is important to note that the construction of the electrical substation and other related electrical activities were not included in this study and did thereby not contribute to the project duration.

On the other hand, project duration is not enough to assess and evaluate whether the performance was optimized. The utilization rate needs to be evaluated as
well. According to Figure 9, the majority of resources (11 out of 19) have a utilization rate greater than 0.9. The mean utilization rate in the initial scenario was equal to 0.53 and it increased to 0.83 after optimization. The standard deviation was 0.419 and it decreased to 0.198 which means that the use of the different resources has become quite more balanced.

Based on the aforementioned results, AnyLogic, when compared to other software, proved to be very effective in simulating construction operations using one of its modeling paradigms, DES. It does not require advanced skills in programming, is very flexible and can easily incorporate changes and updates, while allowing modeling of any desired situation with any level of details. It can display results in different formats (e.g. graphs, histograms, etc.) at runtime and can easily be linked to other software (e.g. MS Excel, GIS, etc.) for data collection and analysis purposes. Most importantly, it includes optimization tools and allows the user to validate simulation results using built-in 2D and 3D animation features. This latter attribute led the authors to take unprecedented steps attempting at animating and visualizing in 3D the whole wind farm construction process for validation purposes. Preliminary results are depicted in Figure 10.

Figure 10. AnyLogic Animation Snapshots

4. Summary and Conclusions

This study modeled on-shore wind farm construction processes using a general purpose simulation tool, AnyLogic. A generic simulation model, based on construction activities’ typical daily outputs and crews in Lebanon, was created then was tested by adopting the site of Falougha, Dahr El Baydar, Lebanon. Initial results displayed a long project duration coupled with uneven resource utilization rates. Those were greatly enhanced and improved after optimization in AnyLogic was carried out.

Future work will include cost information and consequently will optimize both project cost and time. A time-cost tradeoff analysis is deemed necessary to reach the most optimal results. Additionally, more efforts will be channeled into 3D animation to further ensure the credibility and validation of the proposed simulation model.

Acknowledgments

The presented work has been supported by Munib and Angela Masri Institute of Energy and Natural Resources. The authors gratefully acknowledge Masri Institute support.

References


Abstract -
Successful implementation of construction projects depends on accurate cost estimation. Cost analysis of construction work items is essential to a cost estimating process for contractors. However, current cost analysis tasks cannot be implemented effectively in practices. Therefore, this study proposes an ontology-based representation framework that aims to address such practical problems. The framework can be used to establish a cost analysis knowledge base which can benefit the modeling and application of cost analysis knowledge and hence, improve the accuracy of cost analysis and estimation. This framework is developed by using the ontological modeling technique with which key cost items of cost analysis and relationships among these cost items can be modeled and subsequently be examined. Actual cost analysis information of eight cases were collected and used to demonstrate and validate the framework. The case study results show that the proposed representation framework can effectively model and store cost analysis knowledge from both historical data as well as current professional work. Furthermore, the modeled cost analysis knowledge can be reused in new cost analysis tasks, and the accuracy and efficiency of cost analysis and estimates can be improved by eliminating the possibility of leaving out necessary cost breakdown items. Future research is suggested on improving the framework by developing a reasoning mechanism that can automate cost analysis processes of reusing existing cost analysis knowledge.

Keywords -
Construction Management; Cost Analysis; Cost Estimation; Ontological Modeling; Ontology

1 Introduction
Accurate cost estimation of construction projects enables successful project implementation. Construction project cost commonly has three levels of estimates: project summary level, cost item level and unit price level [1]. Project summary level summarizes various cost categorizes; cost item level subdivides each category into smaller cost items; and unit price level calculates the cost required to complete a unit of work for a cost item (the cost item level and unit price level are respectively referred to as work item level and cost breakdown level in the herein study). Therefore, to accurately estimate construction project costs, cost estimators should have practical experience and knowledge to determine the cost of every work item by performing cost analysis in which the cost breakdown items (i.e., sub-work items and resource items) of a work item are specified and their costs (unit prices) are determined. The overall project cost can then be calculated by summing up the work item costs.

Cost analysis of construction work items is essential to a cost estimating process for contractors for two reasons. First, it specifies the necessary sub-work items and resource items and their respective costs. Second, cost analysis generates cost analysis sheets for major work items that usually become a part of a contract; therefore, work item costs presented in these sheets become the basis for the application for payment of a contractor and for determining the cost of newly added work items during the construction stage.

Although commercial cost estimating software is available, they mainly support cost estimation on work item level instead of cost breakdown level. In practice, cost estimators usually use spreadsheet software to perform cost breakdown because using spreadsheets makes performing cost breakdown more flexible especially when the cost items of a construction project are plentiful and diverse. However, the major drawback of using spreadsheet is that estimators’ cost analysis knowledge is not well-organized and hence cannot be easily reutilized in future cost analysis tasks.

In recent years, Taiwan has encountered difficulties in establishing an accurate estimation of construction costs, further causing disputes. One of the reasons for
such a problem is that cost analysis for work items is not effectively implemented. According to practitioners and professionals in the construction industry, three causes hindering the effective implementation of cost analysis were identified. First, time frame for a construction project is tight. Limited time is allotted to the planning, design and bid preparation phase; as a result, cost estimation tends to be inaccurate. Second, cost analysis generally relies on cost estimators’ practical experience and knowledge of various disciplines; however, there is a lack of such competent personnel. Finally, cost analysis knowledge of experienced cost estimators is not structurally organized and stored. The inability to resolve these limitations makes it nearly impossible to fully exploit the knowledge of experienced cost estimators to initiate a new cost analysis task or conduct educational training on cost analysis for inexperienced estimators. Thereby, the herein proposed ontology-based representation framework aims to model and store cost analysis knowledge based on the work of contemporary cost estimators and historical cost analysis data. Moreover, in addition to facilitating the utilization of this stored knowledge in new cost analysis-related tasks or educational training, the proposed framework further aims to improve the accuracy of cost analysis and estimation by eliminating the possibility of leaving out necessary sub-work items or resource items of work items.

This paper is organized as follows. Section 2 reviews the essential literature of ontology and ontology language. Section 3 discusses the proposed representation framework and how to deploy the framework to establish a cost analysis knowledge base as a cost analysis ontology. Section 4 presents a case study that tests and validates the feasibility and application of the framework. The final section summarizes the findings of this study and future research directions.

2 Related work

2.1 Ontology and ontological modeling

Ontology was first defined by Bunge in 1977 as “the basic characteristics of the real world.” [2] Ontology originated from the philosophy domain and was commonly defined as “an explicit and formal specification of a conceptualization” [3]. Noy and McGuinness [4] further pointed out that the goal of ontology is to share common understanding of the structure of information among people or software agents in order to enable the reuse of knowledge in a field. In short, Ontology is a normative model, which represents concepts of a knowledge domain and the relationship between the concepts. Ontological modelling, therefore, can be viewed as a systematic approach for modelling concepts and relationships into ontologies [5].

Ontological modeling is widely used in fields such as knowledge management and organization, semantic web, web commerce, database design, natural language processing, agent-based systems, and software engineering [6]. In the field of construction engineering and management, ontological modeling has progressively been applied to different research areas, such as construction safety knowledge representation, reasoning and retrieval [5] [7] [8].

The basic elements of ontology include classes, attributes, relationships, and instances, each of which is introduced as follows [4]:

1. Class: A class is a category representing an entity of a certain domain, using one word or a combination of words to allow users or computer systems to understand the meaning of the category. A class can be subdivided into many subclasses with more detailed descriptions. For example, a class “Concrete Worker” can be defined as a subclass of the class “Concrete Labor.”

2. Attribute: Attributes are descriptions of the characteristics and features of each class in an ontology. Ontological information framework can be built and useful data can be provided through attribute definitions. For instance, an attribute “unit price” can be defined for the cost breakdown class “Concrete Worker,” representing the daily wage of a concrete worker.

3. Relationship: Relationships are semantic connections between classes, including association, generalization, equivalence and disjointedness. In the present study, relationships are defined to connect work item classes to their cost breakdown item classes. For instance, the work item concrete has sub-work items, such as concrete curing, and multiple resource items, such as concrete labor, 140kgf/cm² Type I concrete, and concrete pumps. Association relationships exist between the work item concrete and its sub-work item and resource items.

4. Instance: Instances are implementation of classes, similar to objects in object-oriented modeling. Therefore, an instance resulting from a class has specific values for all the attributes of the class. For example, the instance of a concrete labor class shall have a unit price of US$60. In other words, instances are implemented to more clearly express a class.

2.2 Ontology Language

An ontology requires a standard language to express domain knowledge. Therefore, different ontology
languages were developed and used to create ontological models so that a computer system can understand a model by interpreting the ontology language for the model. Many ontology languages have been developed in recent years. These ontology languages began with Extensible Markup Language (XML) syntax as the basis for ontology development (Figure 1). Other important ontology languages include RDF (Resource Description Framework) and RDFS (RDF Schema), DAML+OIL (Darpa Agent Markup Language + Ontology Interchange Language) and OWL (Web Ontology Language). Among these languages OWL has stronger reasoning capabilities over others and is the most popular one in recent years for developing ontologies. Therefore, this study adopts OWL as the ontology language for establishing cost analysis knowledge base.

3 Cost analysis ontology establishment

In this study, the representation framework is designed for the main structural work items of building construction (i.e., reinforce bar, formwork, and concrete). Figure 2 provides an overview of the representation framework. The representation framework provides a structure of an ontological knowledge base for storing cost estimators’ cost analysis knowledge and historical cost analysis data. Figure 2 also shows a knowledge base, which simply signifies archived historical data from actual project cases and knowledge of cost estimators. The following sections explain the framework and knowledge base respectively with examples.

3.1 Representation Framework

The purpose of the representation framework is to provide a structure for storing cost analysis knowledge and historical data as a cost analysis ontology. That is, the cost analysis information of past projects can be properly classified and stored through the representation framework. To achieve this purpose, three aspects should be considered in the framework: classification of work items and cost breakdown items, associations between these items, and attributes of cost breakdown items.

In this study, the representation framework consists of two major classes: “Bill of Quantity Items” and “Cost Breakdown Items” (as shown in Figure 2).

1. The class “Bill of Quantity Items” collects the main structural work items and categorizes them into three subclasses: “Steel”, “Form” and “Concrete”. For example, a work item 140kgf/cm² concrete using type I Portland cement is represented as a class “Structural Concrete Ready Mixed 140kgf/cm² Type I Cement” in the representation framework, and this new class is defined as a subclass of the class “Steel”. Furthermore, two new classes “Common Forms” and “Natural Forms” are defined as subclasses of the class “Form” in the representation framework to respectively represent two work items, common forms and natural forms.

2. The class “Cost Breakdown Items” collects cost breakdown items of bill of quantity items and categorizes these breakdown items into two subclasses: “Sub-Work Items” and “Resource Items”. The class “Sub-Work Items” represents the work performed in a work item whereas the class “Resource Items” indicates the resources used in a work item. Four classes, “Labor,” “Equipment,” “Material,” and “Miscellaneous Work” are defined as subclasses of the class “Resource Items” to respectively represent labors, equipment, materials and miscellaneous items used in a work item. In addition, each of these classes is then specialized by defining its three new subclasses that can further specify a resource item for steel, form or concrete work. For instance, a class “Concrete Labor” is defined to represent labor resources for concrete work while “Form Labor” is defined to represent material resources for formwork. Both classes are subclasses of the class “Labor.”

All the defined classes under the major class “Cost Breakdown Items” form a classification that can represent cost analysis information of bill of quantity work items. For example, for the work item, 140kgf/cm² concrete using type I Portland cement, its cost breakdown includes a sub-work item, i.e., concrete curing, and resource items, i.e., concrete operating worker (labor), concrete vibrator (equipment), 140kgf/cm² Type I concrete (material) and tool loss (miscellaneous work). New classes representing these sub-work item and resource items then can be defined as classes in the representation framework. For instance, a class “Concrete Vibrator” is defined as a subclass of the class “Concrete Equipment” and represents the equipment used in the work item 140kgf/cm² concrete using type I Portland cement.
In addition to the two major classes, five association relationships are also required in order to represent the connections between a bill of quantity work item and its cost breakdown items. Therefore, five association relationships, “hasSub-Work,” “hasLabor,” “hasEquipment,” “hasMaterial,” and “hasMiscellaneous,” are defined as object properties in the representation framework, which can be used to link a work item to its sub-work items and resource items.

Lastly, cost breakdown items own attribute information of unit, quantity, and unit price. Therefore, three attributes “unit,” “amount,” and “unit price” are defined as data type properties for the class “Cost Breakdown Items” in the representation framework to allow instances created from the class to take values on these attributes.

This representation framework provides a formal structure for categorizing concepts identified from actual project cases. Section 3.2 illustrates the steps of establishing a cost analysis knowledge base, i.e. cost analysis ontology, using the representation framework.

### 3.2 Knowledge base

This study collected cost analysis data from eight historical cases, seven of which were used to establish the knowledge base while the other case was used to test and validate the proposed framework. The work item concrete ready mixed 140kgf/cm² type I cement is used as an illustrative item (referred to as Case 1 in this study) to demonstrate the steps of establishing a cost analysis knowledge base. This study uses Protégé, an ontology editor developed by Stanford University [10], to establish the knowledge base.

1. Create classes and instances for work items and their cost breakdown items: Classes are first created in the representation framework for work items and their cost breakdown items. For example, a class “Concrete Ready Mixed 140kgf/cm² Type I Cement” is defined for the illustrative work item in Case 1 (Figure 3). Then, instances are created from the classes, and actual attribute values are assigned to the instances. For example, an instance “Concrete operating worker-Case1” is created from the class “Concrete Operating Workers” to represent that a labor item concrete operating worker is included in the work item concrete ready mixed 140kgf/cm² type I cement in Case 1. Furthermore, this instance have actual attribute values as follows: “labor” for the unit attribute,
“0.02” for the amount attribute, and “NT$2,500 (2,500 New Taiwan Dollars)” for the unit price attribute (Figure 3).

2. Connect classes with association relationships:
Each of the bill of quantity items is connected to its cost breakdown items using the five association relationships. For example, the bill of quantity item class “Concrete Ready Mixed 140kgf/cm² Type I Cement” is connected to two equipment item classes “Concrete Vibrator” and “Concrete Handling Equipment” with the association relationship “hasEquipment” (Figure 4). The purpose of this step is to represent the semantic relations between bill of quantity items and their corresponding cost breakdown items; then, the instances of cost breakdown items of past cases can be retrieved by specifying the bill of quantity item which is of interest to cost estimators.

3. Define equivalent classes: Classes which have the same semantic meaning are set equivalent (Figure 5). For example, two labor item classes “Production Physical Labor” and “Unskilled Worker” are used interchangeably in cost analysis process and therefore are defined equivalent to each other. Through the definition of equivalent
classes, cost breakdown items which have the same semantics can be formally represented and therefore, cost estimators will not ignore those cost breakdown items with the same meaning but in different texts when performing cost estimation.

Figure 5. Define equivalent classes

4 Case Study

This case study illustrates the use of an established cost analysis knowledge base, i.e., cost analysis ontology, and also validates the capability of the proposed representation framework. A work item class “Common Forms” is taken as an example to demonstrate how to use the knowledge base to perform a cost analysis task for the work item. Figure 6 shows the contextual scheme of the case study.

Figure 6. Scheme for performing a cost analysis with a cost analysis ontology

1. Start cost analysis: A cost estimator initiates a cost analysis and first decides to analyze the work item common forms.

2. Navigate the cost analysis ontology: The cost estimator can identify the class “Common Forms” for the work item common forms from the “Bill of Quantity Items” class hierarchy in the cost analysis ontology. The ontology displays those sub-work items and resource items used in the past projects as shown in Figure 7, allowing the ontology user to examine and evaluate the cost of those work items. For example, cost estimators can identify carpenters and unskilled workers as two labor items of the work item common forms because the cost analysis ontology shows that the classes “Carpenter” and “Unskilled Workers” are connected to the class “Common Forms”.

3. Estimate a work item cost: Cost estimators then assess what cost breakdown items, i.e., sub-work items and resource items, should be selected for the work item. In addition, the historical cost data of all the instances for the selected cost breakdown items can be examined; estimators can refer to and use these cost (unit price) information in the new cost analysis. For instance, two instances are found for the resource item class “Carpenter”, i.e., “Carpenter case6” and “Carpenter case8” (Figure 8). If estimators adopt the cost information for carpenters of the Case 8, the unit price for carpenters in this cost analysis is NT$2,400.

4. Meeting: After estimators use cost analysis ontology to perform cost estimation, they can list out the cost breakdown items and complete the cost breakdown form, as shown in Table 1, which includes all the cost breakdown items identified from the cost analysis ontology and their unit prices retrieved from the ontology. This can be used as a reference during the cost assessment meetings for project team members to evaluate
whether any cost breakdown item is missed or unit prices for these items are reasonable.

Figure 7. Identify cost breakdown items for the work item

Figure 8. Retrieve the unit price of the instances of a cost breakdown item for the work item

Table 1. Cost breakdown form for estimating the work item cost

<table>
<thead>
<tr>
<th>Work items</th>
<th>Cost Breakdown Items</th>
<th>Unit</th>
<th>Amount</th>
<th>Unit prices</th>
<th>Item cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common forms</td>
<td>Sheet materials loss for form</td>
<td>M²</td>
<td>1.000</td>
<td>56.34</td>
<td>56.34</td>
</tr>
<tr>
<td>Common forms</td>
<td>Cut material loss for form (contain support)</td>
<td>M²</td>
<td>1.000</td>
<td>52.82</td>
<td>52.82</td>
</tr>
<tr>
<td>Common forms</td>
<td>Cup hooks iron members</td>
<td>M²</td>
<td>1.000</td>
<td>17.61</td>
<td>17.61</td>
</tr>
<tr>
<td>Common forms</td>
<td>Carpenter</td>
<td>Labor</td>
<td>0.100</td>
<td>2,400.00</td>
<td>240.00</td>
</tr>
<tr>
<td>Common forms</td>
<td>Miscellaneous material and labor</td>
<td>Lump Sum</td>
<td>1.000</td>
<td>14.09</td>
<td>14.09</td>
</tr>
<tr>
<td>Common forms</td>
<td>Total (New Taiwan Dollars)</td>
<td>M²</td>
<td>1.000</td>
<td>380.86</td>
<td>380.86</td>
</tr>
</tbody>
</table>

5 Conclusion
This study proposes an ontology-based representation framework for establishing cost analysis knowledge base for work items. This ontology-based framework can store cost analysis experiences and knowledge of cost estimators from previous cases. It
can also store historical cost analysis information of main structural work items. Additionally, this study utilizes the characteristics of an ontology, such as establishing association relationships between classes and defining equivalent classes, to establish the framework for cost analysis. The study result show that cost estimators can identify the cost breakdown items of a work item and retrieve their unit prices when performing cost analysis through the representation framework and steps demonstrated in this study to establish a cost analysis ontology.

The proposed representation framework and the steps of developing a cost analysis knowledge base is part of an ongoing research, which aims to integrate an ontology-based cost analysis knowledge base and building information modeling to assist in selecting construction methods. This study still has some limitations to be improved in the future research. First, this study only considers cost analysis components for structural construction in the framework. The framework should be expanded to consider other work item types, such as temporary work and demolition and decoration constructions, to establish a comprehensive cost analysis ontology. Second, the herein proposed framework does not support automated identification of and reasoning about cost breakdown items. Such goals can be achieved by developing an automated reasoning mechanism in future research using ontology reasoning languages, such as SWRL (Semantic Web Rule Language) or programming language, such as Java. The reasoning mechanism shall be able to facilitate cost analysis process and therefore shorten the time for estimating construction project costs.

6 Acknowledgments

This research is financially supported by the National Science Council of Taiwan (Contract No. NSC 102-2622-E-009-010-CC3 and NSC 102-2221-E-008-102). We are also grateful for the valuable information and experience offered by the engineers and professionals involved in the case studies.

7 References


Estimating the Costs, Energy Use and Carbon Emissions of Concrete Recycling Using Building Information Modelling

A. Akbarnezhad and Z.S. Moussavi Nadoushani

School of Civil and Environmental Engineering, the University of New South Wales, Sydney, Australia
E-mail: a.akbarnezhad@unsw.edu.au, z.moussavinadoushani@student.unsw.edu.au

Abstract -
The economic and environmental benefits achievable through concrete recycling depend on many parameters including, but not limited to, travelling distances between demolished building and concrete recycling plant and/or landfills, prices of natural and recycled aggregates and the desired quality of the recycled aggregates which itself depends on a number of other parameters including the properties of parent concrete and the recycling procedure used. The overlapping effects of such parameters makes decision making about the selection of the concrete recycling strategy and appropriate level of recycling difficult. This paper presents a framework for assessment of costs, energy use and emissions incurred by adopting the concrete recycling strategy in a particular building project using the information made available by building information models. The estimated costs, energy use and emissions may be helpful in decision making about selection of the concrete recycling strategy in a particular project. An illustrative example is presented to highlight the potential benefits of the proposed method.

Keywords -
Building Information Modelling, Concrete, Recycling, Cost, Carbon

1 Introduction

Concrete recycling is considered as a sustainable alternative to traditional demolition and landfilling. Concrete recycling eliminates the need for the costly and energy intensive transportation to the usually remote landfills and reduces the need for the extraction of natural aggregates by converting the concrete debris to recycled concrete aggregates (RCA) [1]. However, an important fact that is often overlooked is that the recycling process involves a number of energy intensive operations including transportation between demolition site and recycling plant, breaking the large concrete chunks into smaller pieces that can be fed to the crushers, removal of non-concrete impurities and reinforcing bars, multiple crushing stages, possible beneficiation stages and sieving which may result in significant costs, energy use and carbon emissions.

The energy and carbon implications of concrete recycling have been investigated in a number of recent studies [2]. However, the average embodied energy and carbon reported in such studies have been mainly estimated through life cycle analysis of a particular recycling procedure and may not be a good representative of the recycling process required to achieve a particular RCA quality in many projects. The recycling process used in practice varies widely from a recycling plant to another and depends highly on the expected quality of the RCA products. There is always a tradeoff between the quality of RCAs and the costs and energy use of the recycling process. The quality of RCAs and thus their future application depend on a number of parameters including the properties of parent concrete and the recycling procedure [1]. As a result of the variation in the nature, sequence and number of the operations required to achieve a particular desired quality of RCA, any generalization on the costs, carbon and energy implications of the recycling process should be avoided.

Besides the costs and emissions incurred by the recycling operations, the economic and environmental impacts of concrete recycling strategy are influenced by many other project specific parameters including the distance to the landfills available, distance to the recycling plant, price and embodied energy of the alternative sources of natural aggregates available, amount of the concrete debris, etc. Therefore, assessment of carbon and energy implications of concrete recycling using the average values reported in available literature seems unrealistic. The objective of present study is to develop a framework and methodology for estimating the costs, energy use and associated emissions of the concrete recycling and conventional landfilling strategies by considering the project specific conditions and requirements. The estimated costs, energy use and emissions may be helpful in decision making about selection of the concrete recycling strategy in a particular project. A computer application was developed to automatically assess the economic and environmental impacts of the “concrete recycling” and conventional “demolition and landfilling” strategies using the information imported from building information models (BIM). An illustrative example is presented to highlight the
potential benefits of the proposed method.

2 Methodology

2.1 Identifying the Required Recycling Process to Achieve Desired RCA Quality

All the operations involved in concrete recycling process consume energy and thus lead to carbon emissions. The important parameters (operations) contributing to the economic and environmental impacts of the “concrete recycling” strategy are summarized in Figure 1a. As shown, besides the effects of the demolition operation which is common between the landfilling and recycling strategies, the impacts of concrete recycling strategy can be

Figure 1. Important parameters to be considered in estimating the economic and environmental impacts of a) concrete recycling strategy b) demolition and landfilling strategy

(a)

(b)
generally divided into two general groups; impacts of the recycling operations and impacts of transportation. The costs, energy use and emissions incurred in the recycling process depend basically on the operations used in a particular recycling plant. The type and sequence of the operations involved in the recycling process are basically a function of the quality of the RCA needed as well as the properties of the concrete debris available. The parent concrete properties and information on the availability of various local recycling technologies can be used to identify the recycling procedure needed to achieve a desired RCA quality. In this study, a computer application was developed to identify the required recycling operations to achieve the expected RCA quality for a given set of parent concrete properties (input by user or imported from BIM), the expected density (water absorption) of the recycled concrete aggregates or the expected maximum reduction in strength of the new concrete to be made with RCAs (referred to as recycled aggregate concrete (RAC)). The algorithm used was based on the experimental results reported in references [1, 3]. Similarly, to estimate the economic and environmental impacts of conventional demolition and landfilling strategy for comparison and decision making purposes, the main parameters affecting the costs, energy use and emissions in this strategy should be identified. These parameters are summarized in Figure 1b. These are on the top of the economic and environmental impacts of demolition.

2.2 Estimation of the Costs, Energy use and Emissions of Individual Operations

In the next step, the costs, energy use and associated emissions of individual recycling operations are estimated. The procedure used is summarized in Table 1. The input data for estimating the actual costs, energy use and associated emissions of a particular operation include the price, embodied energy and embodied carbon of the equipment, service life of the equipment and the estimated production rate of recycling plant. The embodied energy and embodied carbon of equipment refer to the energy use and the associated carbon emissions, respectively, incurred during the manufacturing process, transportation and installation of the equipment and can be estimated through life cycle analysis. Present study assumes that such information is made available by equipment manufacturers or can be obtained from available life cycle inventories. In the present study, a database was created to store price, energy, carbon, and service life information for the machineries commonly used by the local recycling plants and demolition contractors. The costs, energy use and carbon emissions incurred in the recycling process were then estimated using the methodology described in the following.

The energy use of the recycling process is estimated as:

\[ \text{ER} = \left( \sum_{i=1}^{n} \text{EU}_i + \text{DEE}_i \right) \times W \]  

(1)

Where, ER is the total energy consumed in production of W ton of RCA, EU$_i$ is the operation energy used by equipment number i, n is the total number of the recycling operations and DEE$_i$ is the predicted depreciation in the embodied energy of the equipment i when used to produce 1 ton of recycled concrete aggregate. Similarly one could derive similar expression for the associated carbon emission and costs of the recycling process as follows:

\[ \text{CR} = \left( \sum_{i=1}^{n} \text{C}_i + \text{DEC}_i \right) \times W \]  

(2)

\[ \text{CoR} = \left( \sum_{i=1}^{n} \text{C}_o + \text{DEV}_i \right) \times W \]  

(3)
Where, CR and CoR are respectively the carbon emissions and costs incurred by the recycling process to produce W tonne of RCA, C_i and Co_i are respectively the carbon emissions and costs incurred to operate the equipment/machine i in the recycling plant (per tonne production of RCA), DEC_i and DEV_i are respectively the predicted depreciation in the embodied carbon and economic value of the respective equipment when used to produce 1 tonne of recycled concrete aggregate, i is the number associated with each equipment and n is the total number of the equipment used to produce RCA in the recycling plant.

Various depreciation methods may be used to estimate the depreciation in the embodied energy, embodied carbon and economic value of the equipment per ton production of RCA. The method used in the present study is based on the estimated production capacity of the equipment during its service life, the salvage value (economic value, embodied carbon and embodied energy of the salvaged equipment) and the initial value (economic value, embodied carbon and embodied energy) of the equipment. The depreciation in the embodied energy value, embodied carbon value and economic value of equipment due to production of 1 ton of RCA may be estimated using the following equations:

\[
DEE_i = \frac{(EE_i - SEE_i)}{TSLP_i} \tag{4}
\]
\[
DEC_i = \frac{(EC_i - SEC_i)}{TSLP_i} \tag{5}
\]
\[
DEV_i = \frac{(EV_i - SEV_i)}{TSLP_i} \tag{6}
\]

Where, \( EV_i, EE_i \) and \( EC_i \) are respectively the initial economic value, embodied energy and embodied carbon of equipment i and \( SEV_i, SEE_i \) and \( SEC_i \) are respectively the salvage economic value, embodied energy and embodied carbon of equipment i. TSLP_i is the estimated total amount of RCA producible through the service life of the respective equipment (in tons). The embodied energy, embodied carbon and economic values of various equipment can be estimated using various local and international inventories available and are occasionally made available by the manufacturers. The salvage embodied carbon and embodied energy of equipment is estimated by taking into account the potential use for the salvaged equipment at the end of its service life.

Once the costs, energy use and emissions of individual operations have been estimated, the respective total values for recycling strategy (and/or demolition and landfilling strategy) can be then estimated by adding up those incurred in the individual operations involved. The methodology used in the present study to obtain the required information to perform such calculations is presented in the following section.

\[
\text{Figure 2. The four main components of the concrete recycling analyzer unit}
\]

3 Data Gathering and Analysis

The main components of the framework and the economic and environmental assessment tool developed in the present study are shown in Figure 2. The Information Modelling Unit is responsible for making available the information required for the economic and environmental analysis. The main components of this unit include building information models (BIM), costs, energy and carbon database, location database and user interface (Figure 3). The subunits of the information modelling unit are briefly described in the following.

**Structural BIM subunit:** Typical structural detailing BIM models created in state-of-the-art structural BIM software contain information that can
be useful in economic and environmental analysis of concrete recycling strategy. These may include properties of the parent concrete, amount of the recyclable concrete available, location of the building and any other information related to the special considerations such as potential hazards and contamination. Such information can be added manually by building designers or imported automatically from components libraries used frequently by the design team.

Cost, Energy and Carbon (CEC) Database: This database provides the information required to estimate the overall cost, energy consumption and emissions incurred by the recycling operations and transportation of debris to the recycling plant. The information stored in this database includes the unit cost, energy use and emissions incurred by various recycling operations including crushing, conveying, sieving, etc. In addition, this database includes information on the present market price of RCA and natural aggregates and the costs, energy use and emissions incurred by various modes of transportation.

Location and Technology (LT) Database: The framework and tool developed in the present study uses the geographic coordinates of the building defined in the building information model and the geographic coordinates of the recycling plants from the location database to calculate the travelling distances required to estimate the costs, energy use and emissions incurred by transportation.

In addition, the location and technology data base used in the present study includes information on the equipment available in the various local recycling plants listed in the database.

The data processor unit developed in this study imports the selected data from the structural information models, costs, energy and carbon databases and the location database, sorts out these data and creates the required data structure which can then be used by the analyzer unit to perform various economic and environmental analyses. In this study, a pilot Analyzer Unit (programmed in Visual Basic) was developed and used to perform the economic and environmental analyses (Figure 4). In the first step, the analyzer unit uses the input information on the required quality of RCA, strength of the parent concrete and size of the aggregates in the parent concrete to estimate the number of the crushing stages required to achieve the required quality. At the same time, based on the information available in the “location and technology database”, the possibility of using the RCA beneficiation technologies based on the availability and various operation alternatives are identified. In the second step, the analyzer uses the information imported from the building information model including the dimensions of the elements, number of the similar elements in the model and the size and number of the reinforcing rebar to estimate the total amount of the recyclable concrete and recyclable reinforcing rebars available. If the strength of parent concrete varies in the model, the quantities of concrete are grouped into different groups requiring different recycling operations to achieve the required quality.

Figure 3. Components of the Information Modelling Unit

In the third step, the analyzer unit uses the information made available by the CEC database including the unit costs, energy use and emissions of the individual operations (as imported from inventories), the information on the recycling operations required as determined in the first step and the estimated amount of the recyclable concrete and steel calculated in the second step to estimate the overall costs, energy use and emissions associated with the recycling process. In the fourth step, the analyzer uses the information made available by the location and technology database to estimate the transportation distance required and to identify the mode of transportation based on the options available. Next, the unit costs, energy use and emissions of the selected mode of transportation are identified from
the CEC database. These values and the estimated amount of the recyclable materials estimated in step 2 are used to estimate the economic and environmental impacts of the transportation. The overall impacts are then estimated by adding up the impacts of the recycling process and the transportation. To estimate the costs, energy use and emissions of the landfilling operation, the transportation distance is calculated using the information provided by the LT database. The amount of the materials as estimated in the step 2 and the unit values of costs, emissions and energy use of transportation looked up from the CEC database are then used to estimate the economic and environmental impacts due to the transportation to landfill.

4 Case Study

The case project considered in the present study is a commercial-residential building with the total site area of 25170.398 m² constructed using a fully precast concrete structure system with the 3D model shown in Figure 5. The structure was modelled in Tekla Structures BIM software. It was assumed that all the structural concrete used had a 28 compressive strength of 30 MPa and were made using granite aggregates with a maximum size of 20 mm. As shown in the proposed framework such information are crucial to estimate the quality of the recycled aggregates produced using a particular recycling process. It was also assumed that owners require any recycling strategy to produce RCAs which are suitable for use in structural concrete at a NA replacement percentage of at least 50% while limiting the strength reduction of the new concrete with respect to the concrete made with 100% natural aggregates to less than 10%. This information was...
used by the computer algorithm developed in this study to identify the recycling operations required to achieve the desired quality. The transportation distances to the recycling plant and landfills were assumed to be 10 km and 50 km, respectively. The transportation was assumed to be performed by a fleet of 10 ton and 5 ton motor Lorries.

The results of the economic and environmental analyses performed based on our proposed framework are presented in Figures 6, 7 and 8. Figure 6 shows the costs incurred during various stages of “recycling” and “landfilling” strategies as well as the earnings through the sale of recycled products in the concrete recycling strategy. The earnings were calculated by estimating the total volumes of coarse and fine recycled aggregates producible by considering the estimated yield of the recycling process and multiplying the estimated volume of each size fraction of RCA by its estimated market price. The steel rebars were assumed to be sold at the market price for the steel scraps. The demolition costs were estimated by calculating the total volume of concrete to be demolished using quantity takeoffs from BIM and multiplying the latter by the unit cost of demolition using the user-selected demolition method. Similar method was applied to estimate the energy use and emissions incurred during the building demolition.

As can be seen in Figure 6, the costs incurred by various concrete recycling operations including breaking, crushing, sieving and conveying accounted for about only 12.5% of the total costs of the recycling process whereas demolition and transportation accounted for about 59.5% and 28% of the total recycling costs. This indicates that the recycling process is not the main cost factor even when relatively more sophisticated recycling operations are used to produce higher quality RCAs (suitable for up to 50% replacement in structural concrete). Therefore, production of high quality RCAs through the use of relatively costlier recycling operations should be considered as a potential strategy.

As can be seen in Figure 6, while the demolition costs are relatively similar for concrete recycling and landfiling strategies, the longer travelling distance to the landfill than to recycling plant, as assumed in this case study, resulted in considerably higher transportation costs in the landfiling strategy. As shown, while the transportation costs accounted for almost 28% of the total costs in the recycling strategy, this figure was about 70% in the landfiling strategy. In general, transportation is also one of the most important factors contributing to the economic and environmental impacts of the concrete recycling and landfiling strategies. Therefore, transportation requirements may be considered as an important decision criteria for selection of the optimal deconstruction strategy for dealing with the concrete debris.

As shown, for the case project considered in the present study, the recycling process seemed to result in about 50% lower costs compared to the demolition and landfiling strategy. However, it should be noted that the costs of the recycling strategy can increase considerably if the concrete debris have to be transported for long distances due to the unavailability of local recycling plants.

Another interesting point to be noticed in Figure 6 is the considerably high revenues achievable from the sale of recycled and recyclable products including the RCAs and steel rebars. As shown, such revenues can easily exceed the costs of recycling and may serve as a source of considerable income to the project owners. As shown, in this case study, the significant earnings from the sale of steel scraps and RCA outweighed the
costs and resulted in a positive total net economic impact for the concrete recycling strategy.

Figure 7. Energy use incurred by concrete recycling and landfilling strategies.

Similarly, Figures 7 and 8 clearly show the benefits of the recycling strategy in terms of energy and carbon implications as compared to the demolition and landfilling strategy. Again, transportation was observed to be a determining factor affecting the overall energy use and associate emissions of both strategies. As a result of the relatively high emissions of land transportation, the contribution of the transportation to the overall energy use and carbon emissions of the recycling strategy in our case study was estimated to be about 65% and 81%, respectively. These figures are considerably than the 28% contribution of the transportation to the costs. This again highlights the importance of the transportation as an important decision parameter in selection of the concrete recycling strategy. The contribution of concrete recycling operations to the energy use and emissions of the recycling strategy were even lower than its contribution to the costs and were about 4% and 9.5%, respectively. This again suggests that producing high quality recycled aggregates through the use of more sophisticated recycling operations may make environmental sense in many projects.

Results presented in Figures 7 and 8 show that a significant amount of energy and carbon can be retrieved in the form of recycled products. As shown, by considering the embodied carbon and embodied energy value of the recycled products achievable, the benefits of the concrete recycling strategy in terms of energy and carbon emissions are appealing. The carbon and energy retrievable by recycling of aggregates were calculated by considering the amount of the RCAs producible at a particular quality level and multiplying the latter by the respective embodied energy and carbon values of original aggregates of relatively similar quality.

The results of this case study clearly show that the carbon costs, energy use and carbon emissions of concrete recycling depend on a variety of project specific parameters. Therefore, the results suggest that the decision about adoption of concrete recycling strategy and the level and degree of recycling should be made after detailed analysis based on all influencing parameters. The framework proposed in this paper may serve as an efficient method to perform such analyses.

5 Conclusion

The BIM based framework and tool presented in this paper can be used to provide decision makers with useful information about the costs and emissions incurred by adopting the recycling strategy in a particular project. The results of these analyses together with other important criteria can be used in multi-criteria decision making to select the optimal level and degree of concrete recycling based on owner preferences.

References


An Integrated 5D Tool for Quantification of Construction Process Emissions and Accident Identification

J.K.W. Wong, H. Li, G. Chan, H. Wang, T. Huang and E. Luo

Abstract -

The environmental and safety performance of construction sites are increasingly regarded as critical factors that need to be monitored for the successful completion of construction projects. Research has also repeatedly highlighted the need to minimise the carbon footprint of the construction process and enhance the capacity of the project team and on-site workers in detecting and avoiding potential construction site hazards. However, a multi-dimensional visualisation technology that would allow project teams to simulate potential carbon emissions from construction plant and equipment and to detect potentially 'dangerous' locations on a construction site is currently lacking. This paper illustrates an integrated 5D model that uses virtual prototyping technologies to quantify carbon emissions, simulate the pattern of emissions from the overall construction process and identify potential 'black spots' of site hazards at the planning stage. The proposed 5D BIM based pro-active construction management system (PCMS) can help to detect potential sources of danger to on-site workers and provides pro-active warnings to prevent fatal accidents caused by falling or being struck by moving objects. A public housing project developed by the Hong Kong Housing Authority is used as a case study to demonstrate the integration of the emission prediction visualisation and accident detection tool into the BIM. The proposed tool demonstrates the utilisation of BIM technology to promote pro-active carbon mitigation and safety performance strategies.

Keywords -

Carbon emissions; construction process; virtual prototyping; construction accidents

1 Introduction

Environmental and safety performance, together with 'cost', 'time' and 'quality', are currently considered to be five key indicators of construction project performance [1]. In recent decades, the construction industry has also been seen to play an increasingly important role in mitigating greenhouse gas emissions due to the ‘fuel-intensive nature and large share of carbon emissions of the industry’ [2]. However, construction sites are also regarded as the most risky and accident-prone workplaces. A poor site safety performance can result in legal liability for the contractors and clients as well as project financial loss and contract delay. In Hong Kong, for example, two major causes of injury on construction sites are striking against or being struck by moving objects, and being struck by moving vehicles [3 and 4]. The capacity to detect and avoid potential hazards will help to improve the safety performance of construction sites.

The current construction boom in Hong Kong presents challenges with regards to the potential for increased carbon emissions and on-site accidents [5]. Accordingly, there is a growing need to reduce the carbon emissions and enhance the safety of the working environment in the sector. A better visualisation of the carbon emissions from construction activities and potential on-site accident ‘black spots’ would help improve the environmental and safety performance of the industry. This paper reports the development and application of a virtual prototyping (VP) based 5D tool (i.e. three-dimensional model, emissions data and site real-time location data) for estimating the possible emissions from construction projects and detecting potential on-site accident black-spots. A public housing project in Hong Kong is used to demonstrate the application of the tool
centres in the US [6-17]. Despite these efforts, most of the existing emissions visualisation and quantification models are still in the early stages of development and are limited to regional applications. Much of the existing research also focuses on specific construction trades or activities, such as concreting, earthwork and lifting. Moreover, limited research has focused on developing tools that provide a more holistic estimation of emissions from all of the construction activities in a project [2]. Hajibabai et al. [10] have highlighted the need for a more comprehensive tool to analyse and visualise the carbon emissions from construction sites.

With regards to site safety detection, advanced positioning systems such as the radio frequency identification device (RFID), global positioning system (GPS), ultra wide-band (UWB) and wireless local area network (WLAN) allow real-time monitoring of the location of construction workers, equipment and materials [18]. The purpose of these positioning technologies for safety management is to prevent workers from entering hazardous areas such as floor openings, floor edges and equipment operation areas [18-24]. However, because of their varying levels of accuracy, the different positioning systems have the potential to generate false alarms.

The visualisation tool presented in this paper is implemented in four steps [2]: i) collect the general project and equipment data; ii) develop the plant operation plans; iii) identify the predicted emission quantities and setup the emission estimation model (PEEM); and iv) construct a four-dimensional virtual prototype and import the emissions data. First, a series of activities, each of which have a defined duration, are linked with the construction plant, components and resources [2]. Information, including the operating hours of the equipment and plant based on the site equipment operation plan, is then acquired to predict the emissions from the construction process [2]. Details of the VP emissions visualisation model can be found in [2]. By linking the 3D models (Revit-based software) and the construction project schedules (MS Project files) using Autodesk NavisWorks, the tool is able to model the 4D construction schedule and allows real-time and whole-project simulation. The 5D BIM tool also includes a pro-active construction management system (PCMS), which can assist construction site workers in detecting potential sources of danger and provide pro-active warnings on potential hazards. The PCMS comprises two sub-systems: a real-time location system (RTLS) and a virtual construction simulation system (VCS). Figure 1 depicts the typical three-tier web-based application structure (presentation layer + business layer + data layer) of the PCMS.

The real-time location system (RTLS) can be divided into two parts: the real-time location network and the real-time location engine. The ‘network’ is constructed using small hardware devices which serve as tags, which are designed to be mounted onto helmets and moving objects, and anchors, which are designed to be fixed in static locations to serve as reference points. The anchors locate the tags. The system uses the time of flight (TOF) based location schema. The tags also help to alert construction workers by vibrating and/or emitting a specific sound when they are exposed to a particular danger. An important task in location-based construction safety risk monitoring is to define the relevant dangers (e.g., static dangers and dynamic dangers) in the models and to calculate the relative distances between workers. The real-time location engine is designed with three functions: managing the location network, calculating the tag locations and sending alert signals to the tags. A network may be composed of dozens of tags and anchors. When the ranging results are received, the location engine uses an effective algorithm to calculate the tag positions and sends the positions to the application server for the virtual construction simulation and safety management.

The virtual construction simulation system (VCS) comprises the application server (i.e., the virtual construction engine), the client end, the web server and the database server. The application server handles the business logic of the PCMS by monitoring three possible sources of danger, namely, a person falling from a height, striking against or being struck by moving objects, and being struck by moving vehicles. The system monitors the relative distances between the workers (represented by the positions of the tags installed on the helmets) and potential sources of danger (represented by the tags installed on the moving objects and danger zones that are dynamically defined in the 3D model of the construction site). If the detected distance between a worker and a nearby source of danger is equal to or less than an allowable value, a warning signal will be triggered and sent to the real time location

![Figure 1. System architecture of the PCMS](image)
engine, which will then relay the signal to trigger the warning device on the tag installed on the worker’s helmet. Other functions of the application server include synchronising the user ends to simulate the construction processes and storing and retrieving tag positions to enable the construction processes to be replayed. The user client is a web-based application for visualising construction processes, tracking people and equipment and replaying construction processes. The user client also implements administration features such as managing the danger zones, configuring the anchor positions, managing the relations between tags and tag carriers and managing the virtual construction models.

3 Application of the 5D model

A public housing project in Hong Kong is used to demonstrate the applicability of the VP-based emissions prediction and the PCMS. The housing project involves the construction of a 34-storey residential building and a public car-parking area. To calculate the potential CO2 emissions from the project, a number of discussions and meetings were held with the main contractors and material suppliers to ascertain the equipment to be used during the construction process and the likely fuel consumption rates of all items of equipment. The total number of hours used and the total amount of equipment required were then determined after the details of all of the items of equipment, including the type, engineer tier and nature of activities involved, were collected. The emissions data were then imported into the simulation model using Autodesk NavisWorks. In the simulation, different construction activities are presented in a 4D virtual reality environment. The simulation displays the amount of total emissions and the emission variations. The amounts of CO2 emitted by various types of equipment are presented in graphic format in the lower right corner (Figure 2).

The simulation can visually represent the operation of any items of construction equipment or plant at a particular stage. The construction team members and contractors can identify any activities that have high predicted emission rates and find a solution, such as reducing the idling time of the equipment [2]. The simulation can also enable project team members to communicate and identify strategies to minimise unnecessary emissions and set up an appropriate environmental management plan. The simulation predicted that the housing project would generate a total of 700,000 kg of CO2 emissions from all items of plant and equipment. To reduce the environmental impact of emissions, the Hong Kong Housing Authority wishes to reduce the total emissions of the project by fifteen percent (i.e. 105,000 kg). The project team reassessed the construction programme and the equipment schedule, and identified any unnecessary tower crane and excavator operations to reduce energy consumption.

To detect any potential hazard black spots in the housing project, the location-based virtual construction was developed by integrating the virtual construction technology with the RTLS. This enabled the virtual models to be immediately connected with realistic construction situations, in particular through integrating the static virtual models and dynamic dangers. Two toolkits, Unity and SmartFoxServer, were used to develop the location-based virtual construction for this project. Unity helps generate 3D video games, architectural visualisations and real-time 3D animations, and was used to build features, including visualising the construction process and defining the static and dynamic dangers, for the user client. SmartFoxServer, which is a massive multiplayer game server, was used to help construct the application server. A server object extension will be developed based on the SmartFoxServer to drive and synchronise all of the user clients with the real construction situations. After testing several location technologies, the project team finally selected CSS (Chirp Spread Spectrum) as the ranging technology. The CSS uses TOF to estimate the physical distance between two devices, although it has higher precision than other TOF methods, e.g. greater receiving signal strength.

A collaborative localisation schema was adopted to construct the real time location system (RTLS), which does not require synchronisation between the infrastructure nodes, and is believed to be suitable for construction site environments. The location tags perform ranging with the location anchors, hence the distances are known. The coordinates of a location tag can be estimated based on these distances and the known anchor coordinates. Based on the application requirements, the system is designed such that the...
positions of the location tags are calculated by the location engine. Approximately 100 tags were installed on-site for around four months to evaluate the technical feasibility and usefulness of the PCMS. The construction site was separated into eight zones approximately 30m x 30m in size for the PCMS. Ranging results were sent to the location engine through the CSS wireless network. An example of an image captured by the PCMS system during the trial run for this project is shown in Figure 3. The red spot represents the location of the site operators while the blue spots represent the locations of the hook of the tower crane. The simulation results indicated that the project had no obvious site hazard black-spots.

In summary, the CO₂ emission prediction tool presented in this paper can help contractors to identify the sources of emissions and to quantify the amount of emissions generated. The tool also promotes a proactive environmentally conscious construction approach and the best practices for sustainable development. The tool can assist builders/contractors to forecast activities with excessive emissions and identify suitable mitigation strategies, such as replacing old plant and equipment with energy-saving models and reducing idling time. The PCMS also provides a platform for the construction project team to reassess their site safety plan. The tool provides pro-active warnings to site workers and helps them to detect surrounding sources of danger, such as height hazards and materials being moved by the tower crane.

Figure 3. Location of the site operators (red spot) and the hook of the tower crane (blue spots) captured by the PCMS system

4 Conclusion and Future Research

This paper outlines the development and application of a 5D visualisation tool to support project teams in estimating and visualising the CO₂ emissions from construction activities and predicting potential hazard black spots. Nonetheless, the 5D model is still in its preliminary stage and the tool needs to be applied to different construction projects of varying scale and nature. A comprehensive carbon footprint assessment tool is also required to predict the total embodied energy (including the carbon emitted from embodied energy and the building assembly process) of the project. An integrated life-cycle analysis (LCA) with BIM will be developed to monitor the embodied carbon [for example, 25 and 26]. A BIM based tool that can provide support for managing construction and demolition waste is currently lacking. The PCMS presented in this paper is also at the trial run stage, and the tool requires further testing and validation before it can be widely adopted on-site. These limitations will be tackled in future studies.

Acknowledgement

The authors would like to acknowledge the support of the Hong Kong Housing Authority in conducting this research.

References


Teizer, J., Cheng, T., and Fang, Y. Location tracking and data visualization technology to advance construction ironworkers’ education and training in safety and productivity. Automation in Construction, 35: 53-68, 2013


Employing Ant Colony for the Optimal Reduction of Project Risk Severity

N.Y. Zabel\(^a\), M.E. Georgy\(^b\) and M.E. Ibrahim\(^c\)

\(^a\)WorleyParsons Arabia, Al-Khobar, Saudi Arabia
\(^b\)School of Property, Construction and Project Management, RMIT University, Australia
\(^c\)Faculty of Engineering, Cairo University, Egypt
E-mail: nael.zabel@worleyparsons.com, maged.georgy@rmit.edu.au, moheeb.elsaid@eng.cu.edu.eg

Abstract -

Efforts undertaken in identifying, analyzing and assessing project risks are only made good use of when proper risk treatment strategies are decided upon and pursued. Based on the criteria established by senior management, the risk management plan goes about defining how each risk is to be handled. There are options to that end, including acceptance, avoidance, transfer and mitigation. Whilst these strategies are known to all in the industry, the decision-making process is far from easy. A research was undertaken to optimize risk treatment in construction projects, where both costs and benefits are balanced out at the project level. The paper particularly introduces Ant Colony Optimization (ACO) as a capable algorithm for the balanced selection of risk treatment strategies; that is to reduce the overall risk severity in a project at the minimum cost possible. ACO resembles the real life behavior of ants in their intelligent and guided search for food. The research is being applied in the pipeline construction sector and made use of professional knowledge and project records from a big construction company in the Middle East. The paper further presents an example project to demonstrate how ACO explores the risk treatment alternatives in a project and chooses the optimal set of strategies in such context.

Keywords -
Risk management; Risk treatment; Risk mapping; Optimization; Ant Colony

1 Introduction

Project risk management (PRM) is crucial and indispensable to the success of projects. Indeed, risks in the complex projects of nowadays have magnified in terms of number and global impact. Projects are more than ever exposed and averse to risks, and stakeholders are asking for more risk management to cover themselves against financial or legal consequences [1].

Efforts have been undertaken over the years to help us better identify and analyze risky events in projects. Yet, little has been done to address the decision-making component during the risk handling/treatment stage [2, 3]. A review of the literature revealed that research on risk handling/treatment is mostly opinion- or case-based and, as such, it offers scant guidelines for making the decision [2].

Only recently have researchers realized the need to address risk treatment in more depth. Chapman and Ward [4] recommended balancing the cost of treatment actions with the consequences of the associated risks. Quantitative approaches were then adopted to optimize and/or simulate the risk treatment strategies in light of the set project objectives [1, 2, 3].

2 Challenge and Research Approach

The aforementioned researches are difficult to apply in construction projects, as they depend on numerical variables difficult to estimate in real world practice. Furthermore, the poor and inefficient record keeping, which is not uncommon in some construction companies, will complicate the matter further. Accordingly, the authors developed a model that employs indices for the risk treatment decision-making [5, 6].

2.1 Optimizing the Risk Treatment Actions

Optimizing the risk treatment in a project involves identifying the actions with the highest benefit-cost (B/C) balance to that project. A risk treatment index, \(I_{RT}\), is devised to measure the B/C balance, as follows:

\[
I_{RT} = \left( \frac{(RM_b - RM_a)}{CRT} \right) \times RM_b \tag{1a}
\]

\[
I_{RT} = \left( \frac{(P_d J_b - P_d J_a)}{CRT} \right) \times RM_b \tag{1b}
\]

where \(RM_b\) is the risk magnitude prior to the risk treatment action, \(RM_a\) is the risk magnitude after the
risk treatment action, \( P_b \) is the probability of risk occurrence prior to applying the risk treatment action, \( I_b \) is the risk impact prior to applying the risk treatment action, \( P_a \) is the probability of risk occurrence after applying the risk treatment action, \( I_a \) is the risk impact after applying the risk treatment action, and \( C_{RT} \) is the cost associated with the risk treatment action.

As noted, the B/C ratio is multiplied by the term \( RM_b \) so as to factor in the relative significance of the project risks, which is a fundamental aspect in the succeeding optimization process. The authors’ approach allows using either qualitative or quantitative data sources [6]. Utilizing the model in case of qualitative data is made possible via the use of numerical rating scales that correspond to the qualitative terms.

2.2 Why Ant Colony?

The optimal reduction of project risk severity requires comparing potential actions for treating individual risks. The literature has noted the difficulties associated with using mathematical optimization on large-scale problems [7]. This has contributed to the development of alternative optimizers, such as genetic algorithms, ant colony, particle swarm, etc. The study by El-Beltagi et al. [7] compared these alternative optimizers in an attempt to identify the ones with the better performance. Study noted ant colony to perform superiorly in discrete optimization problems besides being the least demanding in regards to the computer processing time. As a result Ant Colony Optimization (ACO) was the evolutionary algorithm of choice in this study. The paper focuses on this element of the research.

3 Ant Colony Optimization

ACO was developed by Dorigo et al. [8] based on the fact that ants are able to find the shortest route between their nest and a source of food. This is done using pheromone trails, which ants deposit whenever they travel as a form of indirect communication, figure 1. When ants leave their nest to search for a food source, they randomly rotate around an obstacle, and initially the pheromone deposits will be the same for the right and left directions. When the ants in the shorter direction find a food source, they carry the food and start returning back following their pheromone trails and still depositing more pheromone. New ants at the nest will choose the shortest path with the more concentrated pheromone. Over time, this positive feedback (autocatalytic) process prompts all ants to choose the shorter path [9].

Implementing ACO for a certain problem requires a representation of \( S \) variables for each ant, with each variable \( i \) having a set of \( n_i \) options with values \( l_{ij} \) and associated pheromone concentrations \( \tau_{ij} \). As such, an ant is consisted of \( S \) values that describe the path chosen by the ant, figure 2 [10].
solution improves. Therefore, for minimization problems, equation 3 shows the pheromone change as proportional to the inverse of the fitness. In maximization problems, on the other hand, the fitness value itself can be directly used.

Once the pheromone is updated after an iteration, the next iteration starts by changing the ants’ paths (i.e. associated variable values) in a manner that respects pheromone concentration and also some heuristic preference. As such, an ant $k$ at iteration $t$ will change the value for each variable according to the following probability [8]:

$$p_{ij}(k, t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}{\sum_{l} [\tau_{il}(t)]^\alpha \cdot [\eta_{il}(t)]^\beta}$$

where $p_{ij}(k, t)$ is the probability that option $l_{ij}$ is chosen by ant $k$ for variable $i$ at iteration $t$; $\tau_{ij}(t)$ is the pheromone concentration associated with option $l_{ij}$ at iteration $t$; $\eta_{ij}$ is a heuristic factor for preferring among available options and is an indicator of how good it is for ant $k$ to select option $l_{ij}$ (this heuristic factor is generated as per the problem characteristics and its value is fixed for each option $l_{ij}$); and $\alpha$ and $\beta$ are exponent parameters that control the relative importance of pheromone concentration versus the heuristic factor [10]. Both $\alpha$ and $\beta$ can take values greater than zero and can be determined by trial and error.

Based on the previous discussion, the main parameters involved in ACO are: number of ants $m$; number of iterations $t$; exponents $\alpha$ and $\beta$; pheromone evaporation rate $\rho$; and pheromone reward factor $R$.

4 Proposed ACO Model

The index $I_{RT}$ can be employed to quantify the suitability of potential risk treatment strategies in a project. However, deciding upon the optimum risk treatment strategy for a project can be more challenging than it appears. Any project would have a number of risk treatment options constituting the decision points. Each of these risk treatment options can possibly associate with and positively influence multiple risk factors. Another dimension is the inter-dependency of risk factors. Development of a given risk can give rise to other risks. As such, risk mapping is indispensable to modeling such inter-dependency. The authors, in an earlier research [5], have addressed the development of risk maps in construction projects. The pipeline construction sector was used to exemplify the risk mapping process, where 9 risk groups were identified. Each group had a set of potential risks relevant to that group and denoted by GxRy. In this context, $x$ refers to the group number and $y$ refers to the risk code.

![Figure 3. Sample dynamic risk treatment pattern (DRTP)](image)

Obviously, each risk treatment action has a cost associated with it, and when enacted will affect the risks in concern in a certain way. Let us now assume the treatment strategy for the project, i.e., the set of treatments for reducing the project’s risk severity, is represented by an ant. Each treatment $i$ has a total of $n_i$ options. Finding the optimum set of treatment actions then follows as per figure 4.

![Figure 4. ACO process for risk treatment optimization](image)
In a given iteration, the process starts with initializing potential treatments based on the options available for each. Such scenario would correspond to a certain pheromone concentration $\tau_0$. An artificial ant is launched for the 1st treatment strategy and proceeds, i.e., pseudo-randomly walks, till the last treatment as shown in figure 5.

Each ant $k$ would generate a solution. Having completed a cycle, the pheromone value of the selected option $\tau_j$ as generated by the ant is updated according to the pheromone updating rules. After all ants finish their travels, their fitness values are evaluated as per equations 1a and 1b and the best ant solutions then selected.

The approach adopted in this study allows the pheromone updating to be performed according to both the local and global updating rules. The local updating rule implies that the updating is performed after each solution is completed, i.e., when an ant has traveled from 1st to last treatment. On the other hand, the global updating rule involves updating the pheromone after an entire iteration is over, that is to say, after all ants have completed their travels. When the iteration is completed, the pheromone values associated with options belonging to the best solution in that iteration (i.e., inter-best solution) are updated.

As for the stopping/termination criteria, a maximum number of iterations is used in the proposed model due to its convenience and popularity [12]. The algorithm loops back for another iteration until the maximum number of iterations is reached.

5 Illustrative Application in Pipeline Construction Example Project

Substantial research was carried out by the authors to identify, assess and find means to treat the risks that have potential to influence pipeline construction delivery.

Identification and assessment of risks benefited from: (1) the literature, (2) unstructured interviews with selected experts in the field, and (3) a questionnaire survey to a large pool of qualified experts in the Middle East region, where some of the authors work. Full details can be found in an earlier publication by the authors [5].

Survey highlighted 47 risks to exist in the pipeline construction context, tables 1a and 1b. It further revealed the prior probabilities and impacts of these risks. A DRM was then developed to model the interdependencies amongst the risks in reference. Part of such DRM was presented earlier in figure 3.

In a similar effort, risk treatment actions for pipeline construction projects were identified and associated with the risks from the previous research step. Fifty two treatment actions resulted, as illustrated in tables 2a and 2b. Full details can be found in another publication by the authors [6].

A computerized ACO engine was developed using Visual Studio 8 to perform the required optimization process. An example project was used to demonstrate the functioning of ACO. Given the magnitudes of all risks, a total sum of 167 was recorded as indicator of the project’s risk severity. When risk treatments are adopted, the project’s risk severity is reduced. Obviously the dilemma is to find the most effective risk treatment pattern while accounting for the costs associated with each. The latter are represented via cost indices, table 2b.

Assume that four risk treatments RT1, RT2, RT3 and RT4 were available to pursue with a target of reducing the project’s risk severity by 10%. The objective of the optimization process is to maximize the reduction of the risk severity at the least cost possible.

One can think of different patterns comprising the four treatments RT1, RT2, etc. To find the optimum risk treatment pattern, the process starts with initializing the ACO parameters and proceeding with the steps depicted in figure 4. A number of ants (solutions) are created, each of which represents a scenario of using available risk treatment option(s). The evaluation and pheromone updating continue till the termination condition is met.

To exemplify let us consider 16 patterns for risk treatment. Four patterns concern the adoption of only one of the risk treatments RT1, RT2, RT3 and RT4 while the rest comprises combinations of two risk treatments, e.g., RT1 and RT2.

As per table 3, results show the pattern leading to the optimum reduction of project’s risk severity to be pattern 16. This risk treatment pattern consists of RT4 and RT3. It provides the greatest reduction in risk severity compared to the costs invested. Also it satisfies...
the original target of 10% reduction in risk severity. Any further reductions will require other risk treatment patterns that will apparently incur additional costs.

Table 1a Risks in pipeline construction

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1R1</td>
<td>Inability of the owner to finance the project</td>
</tr>
<tr>
<td>G1R2</td>
<td>Delay in progress payments</td>
</tr>
<tr>
<td>G1R3</td>
<td>Inefficient decision making by the owner</td>
</tr>
<tr>
<td>G1R4</td>
<td>Owner’s refusal or questioning of the compensations</td>
</tr>
<tr>
<td>G1R5</td>
<td>Changes in owner expectations</td>
</tr>
<tr>
<td>G1R6</td>
<td>Delay or inability of owner to give full possession of site</td>
</tr>
<tr>
<td>G1R7</td>
<td>Delay or inability of owner to proceed with final acceptance</td>
</tr>
<tr>
<td>G1R8</td>
<td>Owner’s high expectations for quality beyond standards</td>
</tr>
</tbody>
</table>

Table 1b Risks in pipeline construction

<table>
<thead>
<tr>
<th>Code</th>
<th>Magnitude*</th>
<th>Related Risks**</th>
<th>Related Treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1R1</td>
<td>2.61</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>G1R2</td>
<td>4.57</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>G1R3</td>
<td>3.85</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>G1R4</td>
<td>4.12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>G9R3</td>
<td>2.48</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

* Based on surveyed probabilities and impacts. ** Relations defined via the DRM, whether directly or indirectly.

Table 2a Risk treatments in pipeline construction

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT1</td>
<td>The contractors should study the owner’s financial position, and his ability to finance the project for its entire duration</td>
</tr>
<tr>
<td>RT2</td>
<td>Contractors should study &amp; analyze the effect of inflation and devaluation on the project’s costs and consider them in its cost estimate</td>
</tr>
<tr>
<td>RT3</td>
<td>The contractors should obtain a large advance payment, as possible</td>
</tr>
<tr>
<td>RT4</td>
<td>The contractors should ensure receipt of advance payment</td>
</tr>
<tr>
<td>RT52</td>
<td>Cultural and commercial awareness training for management and key personnel who may have to deal with corrupt officials</td>
</tr>
</tbody>
</table>

Table 2b Risk treatments in pipeline construction

<table>
<thead>
<tr>
<th>Code</th>
<th>Cost Index*</th>
<th>Related Risks**</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT1</td>
<td>2.33</td>
<td>4</td>
</tr>
<tr>
<td>RT2</td>
<td>2.80</td>
<td>4</td>
</tr>
<tr>
<td>RT3</td>
<td>2.17</td>
<td>6</td>
</tr>
<tr>
<td>RT4</td>
<td>1.70</td>
<td>6</td>
</tr>
<tr>
<td>RT52</td>
<td>6.47</td>
<td>9</td>
</tr>
</tbody>
</table>

* Based on surveyed costs. ** Relations defined via the DRTP.

6 Conclusion

Decisions made about risk treatment actions are sometimes too subjective. The primary contribution of the research at hand is to devise means that can facilitate making informed decisions about risk treatment in projects. With a sound decision-making process, one can justify why a given set of actions are adopted rather than others.

The paper attempted to develop an optimization algorithm that utilizes ant colony for the balanced selection of a project’s risk treatment strategy. In this context, the benefits and costs associated with the project’s risk treatment strategy are balanced out. Due
to the complexity of projects and the inter-dependency of risks, the algorithm made use of Dynamic Risk Maps (DRMs) which were introduced by the authors in an earlier publication. The novelty of this algorithm lies in its multilevel evaluation process, which accounts for not only the direct impacts but the indirect ones as well.

Table 3 ACO solution

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Ant’s Path</th>
<th>Risk (before)</th>
<th>Risk (after)</th>
<th>Change in Risk Severity</th>
<th>Cost Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RT1</td>
<td>167.00</td>
<td>163.00</td>
<td>2.40%</td>
<td>2.33</td>
</tr>
<tr>
<td>2</td>
<td>RT2</td>
<td>167.00</td>
<td>159.00</td>
<td>4.80%</td>
<td>2.80</td>
</tr>
<tr>
<td>3</td>
<td>RT3</td>
<td>167.00</td>
<td>156.00</td>
<td>6.50%</td>
<td>2.17</td>
</tr>
<tr>
<td>4</td>
<td>RT4</td>
<td>167.00</td>
<td>155.40</td>
<td>6.90%</td>
<td>1.70</td>
</tr>
<tr>
<td>5</td>
<td>RT1-RT2</td>
<td>167.00</td>
<td>160.00</td>
<td>4.19%</td>
<td>5.13</td>
</tr>
<tr>
<td>6</td>
<td>RT2-RT1</td>
<td>167.00</td>
<td>156.00</td>
<td>6.58%</td>
<td>5.13</td>
</tr>
<tr>
<td>7</td>
<td>RT1-RT3</td>
<td>167.00</td>
<td>152.00</td>
<td>8.98%</td>
<td>4.50</td>
</tr>
<tr>
<td>8</td>
<td>RT3-RT1</td>
<td>167.00</td>
<td>154.00</td>
<td>7.78%</td>
<td>4.50</td>
</tr>
<tr>
<td>9</td>
<td>RT2-RT3</td>
<td>167.00</td>
<td>151.00</td>
<td>9.58%</td>
<td>4.97</td>
</tr>
<tr>
<td>10</td>
<td>RT3-RT4</td>
<td>167.00</td>
<td>155.00</td>
<td>7.18%</td>
<td>4.97</td>
</tr>
<tr>
<td>11</td>
<td>RT4-RT3</td>
<td>167.00</td>
<td>154.00</td>
<td>7.78%</td>
<td>4.03</td>
</tr>
<tr>
<td>12</td>
<td>RT3-RT1</td>
<td>167.00</td>
<td>152.00</td>
<td>8.98%</td>
<td>4.03</td>
</tr>
<tr>
<td>13</td>
<td>RT1-RT4</td>
<td>167.00</td>
<td>150.00</td>
<td>10.17%</td>
<td>4.50</td>
</tr>
<tr>
<td>14</td>
<td>RT4-RT2</td>
<td>167.00</td>
<td>150.80</td>
<td>9.70%</td>
<td>4.50</td>
</tr>
<tr>
<td>15</td>
<td>RT3-RT4</td>
<td>167.00</td>
<td>156.00</td>
<td>9.58%</td>
<td>3.87</td>
</tr>
<tr>
<td>16</td>
<td>RT4-RT3</td>
<td>167.00</td>
<td>147.32</td>
<td>11.78%</td>
<td>3.87</td>
</tr>
</tbody>
</table>

The study showed ACO to work fairly well. It is understandable that other optimization engines could be used for such step, however, earlier studies proved ACO to be superior in this particular context. Despite that, further research on parameter selection may be conducted to further improve the robustness of the ACO model.

References


Incorporating Uncertainty into Project Schedule Crashing: An Algorithm

E.S. Subhy, M.E. Georgy and M.E. Ibrahim

Faculty of Engineering, Fayoum University, Egypt
School of Property, Construction and Project Management, RMIT University, Australia
Faculty of Engineering, Cairo University, Egypt
E-mail: ess1@fayoum.edu.eg, maged.georgy@rmit.edu.au, moheeb.elsaid@eng.cu.edu.eg

Abstract -

The uncertain environment in which construction projects are executed poses a challenge to project managers and planners alike as they go about planning and controlling these endeavours. Statistical and simulation models have been devised over the years to assist in estimating the time and cost of construction activities while accounting for uncertainty. However, most of these models fail to comprehensively associate these uncertainty-mindful estimates with the varied resource patterns/configurations that may undertake the work. As a result, when performing schedule crashing, one could end up with a very different strategy for project execution when accounting for all uncertainties compared to the deterministic counterpart. As such this paper presents a dynamic simulation algorithm for project schedule crashing. The devised algorithm incorporates computer simulation into the resource-time-cost triad at the activity and project levels. To perform the dynamic simulation, successive schedule simulations are created with each corresponding to a particular resource pattern/configuration that could possibly be used for executing project activities. After each simulation run is completed, parameters such as the project’s completion time and direct cost are estimated against a certain degree of confidence. The combination of resource patterns for project activities that delivers the minimum project cost is then utilized to produce the optimal-cost project schedule. The computer-automated algorithm is exemplified via a simple project scenario. Results are compared to the traditional approach for least cost scheduling, which show how ignoring the uncertainty dimension could result in strategies far from being optimal.

Keywords -
Project scheduling; Crashing; Construction resources; Simulation; Optimum construction cost

1 Introduction

The reliable estimate of resources, durations and costs for the various project activities is fundamental to its successful planning and control. The deterministic representation of such parameters, also known as single-point estimating, is the norm in developing project schedules. However, the environment in which construction projects are executed is quite dynamic. A plethora of events can influence the construction work, whether positively or negatively. And hence, the confidence in the deterministic estimates is occasionally questioned.

A number of researchers attempted to depict the project parameters stochastically. In this context, multiple values are considered for each parameter into consideration. Probability distributions are utilized to signify the chance each value can occur in reality. With such representation, computer simulation can follow [1] to simulate the project behavior and the many scenarios that can possibly occur.

The stochastic simulation, classified as either discrete-event or continuous, is defined as:

“a technique to make prediction of system performance and to understand its behavior” [2].

Generally, Monte Carlo simulation is the more commonly used platform for the stochastic scheduling and costing in projects. In construction, specific simulators have been devised over the years such as CYCLONE [3], STROPOSCOPE [4], PICASSO [5], among others.

2 Research Need and Approach

In the deterministic analysis, concepts of schedule crashing are normally applied to satisfy the project’s requirements and completion date constraints. Such concepts are sometimes referred to by the term least-
cost scheduling as they aim to arrive at the least possible cost that satisfies the project’s requirements and time constraint. The literature is quite rich in discussing these concepts based on the deterministic representation of resources, durations, and costs [6, 7, 8].

Realizing the importance of stochastic modeling and due to the lack of researches into stochastic schedule crashing, a study was initiated by the authors.

The paper at hand presents a new schedule planning and updating technique, referred to as the Dynamic Simulation for Optimal-Cost Scheduling. Technique establishes comprehensive algorithms for both project planning and the more sophisticated schedule updating. Paper particularly focuses on project planning and introduces the algorithm used for such purpose. It explains how to account for the various resource patterns/configurations used for job execution. Finally, a comparison of the typical least-cost scheduling process and the dynamic simulation approach is presented via an example construction project.

3 Modeling the Resource-Duration Relationship

The productive resources assigned to a construction activity determine the cost and length of time it takes to complete it. As known, the more resources allocated to a job, the less duration and more cost it has. When accounting for the uncertainty in work execution, a given resource configuration/pattern will cause an activity to be executed in a duration that possibly ranges from a low/minimum/optimistic value \( o \) to a high/maximum/pessimistic value \( p \). Within this range, a most likely estimate \( m \) corresponds to the duration having the highest probability of occurrence amongst all values possible.

Let us assume the minimum and maximum levels of productive resources to be \( R_{min} \) and \( R_{max} \), which correspond to durations \( D_{min}(p,m,o) \) and \( D_{max}(p,m,o) \). A triangular distribution is the distribution of choice in this study. The unavailability of abundant data in construction companies makes this distribution a more suitable choice in real-world practice. Further, its usage has been promoted in the process simulation literature [9,10]. With the triangular distribution in place, the relationship between the resource usage and activity duration can be represented in the 3D view illustrated in figure 1. The human resource order refers to the level/number of human resources assigned to the construction job. The least order of human resources corresponds to the longest duration and vice versa. Given the manpower wage rates, the human resource order can be converted into dollar values.

The triangular probability distribution can be converted into the cumulative probability distribution illustrated in figure 2. This distribution is fundamental to the randomization process and the choice of resource patterns and duration values for the subsequent simulation runs.

![Figure 1. Resource-duration relationship (case of unlimited resources)](image1)

![Figure 2. Cumulative probability distribution (case of unlimited resources)](image2)

The representation in figures 1 and 2 assumes no limitation on the productive resources and/or materials used for work execution. However, in reality, some activities in the project will typically have such limitations, figures 3 and 4. As seen, the limitation imposes a minimum possible duration of \( C_i \) on activity \( i \). Consider for instance insufficient materials or construction tools that prevents completing the work in less than \( C_i \). Even if the manpower assigned to work can complete the work in less time, the possibility is eliminated.

![Figure 3. Resource-duration relationship (case of limited resources)](image3)
4 Algorithm for Simulating the Project Schedule and the Crashing Process

A project’s activities need to be executed in a certain order so as to deliver the project on time and on budget. If a certain set of resource configurations/patterns are selected for the various project activities, computer simulation can then be utilized to estimate the possible completion times for such project. Typically a probability distribution for the total project duration is developed. Using this distribution, the probability of completion by a certain date is estimated.

In a sense, this stochastic analysis determines how risky the entire project is, with regard to its completion by a pre-set date. Also, the corresponding project cost can stochastically be estimated for the selected resource configurations/patterns. Time and cost in this regard represent the performance criteria that govern the development and approval of project plans.

When one examines the process of schedule crashing in case of stochastic analysis, it becomes obvious that the process is rather sophisticated. Simulation turns into a successive process, whose parameters change with the change of resource patterns, activities, critical path(s) in the project network, etc. To address such an aspect, a strategy has been established in this research to approach the crashing process.

The stochastic analysis starts with the so-called “Base Simulation”. Minimum productive resources (maximum duration) are assigned to each activity as shown in figure 5. This simulation run is denoted with Sim0.

Using the minimum-resource probability distributions for all project activities, the project’s probabilistic time and cost are developed. These cumulative probability charts provide the basis for moving to the next step. It is understandable that each subsequent simulation, i.e., Sim1, Sim2, etc., will associate with a different resource pattern. Results of the various steps are organized in vectors and matrices as shown in figure 6.

The third vector in figure 6 is particularly important as it tracks the criticality of each activity in the project (i.e., the number of times the activity was recorded as critical in a given simulation). This guides the immediately subsequent simulation of the project and which activities to shorten.

The dynamic simulation process proceeds as per the following steps:

1. Activities not utilizing the maximum possible resource configuration, i.e., there is a window to assign more resources to reduce the duration of the activity, are identified.
2. A preferential analysis is performed to select activities worthy of being assigned more resources in the crashing process. This preferential analysis is designed to take into account four aspects; they are:
   - Utilizing the information of the preceding simulation to progress forward;
   - Observing the criticality of all activities having potential for crashing;
   - Observing the cost increase per unit time for each of the above activities; and
• Observing the multiplicity and parallelism in critical paths.

In conclusion, for any valid-for-reduction activity $i$, the expected increase in the project direct cost per one unit reduction in the project completion time, $ECI_{i, \text{sim}(j)}$, can be calculated as follows:

$$ECI_{i, \text{sim}(j)} = \frac{(DC_{i, \text{sim}(j+1)} - DC_{\text{sim}(j)})}{(DUR_{\text{sim}(j)} - DUR_{i, \text{sim}(j+1)})}$$

Where $DC_{\text{sim}(j)}$ is the direct project cost at sim $j$, $DC_{i, \text{sim}(j+1)}$ is the expected direct project cost at sim $j+1$ resulting from manipulating activity $i$, $DUR_{\text{sim}(j)}$ is the project duration at sim $j$, and $DUR_{i, \text{sim}(j+1)}$ is the expected project duration at sim $j+1$ resulting from manipulating activity $i$. Both $DC_{i, \text{sim}(j+1)}$ and $DUR_{i, \text{sim}(j+1)}$ result from the preferential analysis cited above. Apparently, the minimum $ECI_{i, \text{sim}(j)}$ will point out which activity whose duration should be reduced.

3. The activity selected via preferential analysis will be allocated the additional resources needed for the project’s time reduction. A new CPM-based simulation, according to the new configuration of resources, is performed to obtain the values of $DUR_{\text{sim}(j)}$ and $DC_{\text{sim}(j)}$ for the subsequent step of the simulation process.

4. The new value $DUR_{\text{sim}(j)}$ is evaluated to judge whether further actions are needed to meet the project’s time constraints. If the deadline for project completion is met, the simulation process stops at this stage and the resource patterns/configurations utilized in the last simulation are used in the project plan.

5 Illustrative Example: Least-Cost Scheduling vs. Dynamic Simulation

A simple construction project is used to exemplify the presented algorithm and compare with the more traditional least cost scheduling. The project comprises the fabrication of steel template with gratings, where unlimited resources exist. The activities of this project are (1) preparation of shop drawings and materials take-off; (2) delivery of materials; (3) issuance of shop drawings for fabrication; (4) cutting of template slots; (5) clearing of template slots; (6) rolling of template slots; (7) cutting of gratings; (8) clearing of gratings; (9) rolling of gratings; (10) painting of template slot; (11) galvanization of gratings; (12) inspection; and (13) delivery to site.

5.1 Least-Cost Scheduling

Performing a least-cost scheduling exercise for the project results in the values illustrated in table 1. It is to be noted that the same action is consecutively taken to reduce project duration from 45 to 42 days. Each day in this reduction adds $120 per day. This accumulates into $360, i.e. 3 times $120, when accounting for the 3-day duration reduction.

<table>
<thead>
<tr>
<th>Days Shortened</th>
<th>Project Duration</th>
<th>Direct Cost</th>
<th>Cost Increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>47</td>
<td>$36,355</td>
<td>$0</td>
</tr>
<tr>
<td>1</td>
<td>46</td>
<td>$36,415</td>
<td>$60</td>
</tr>
<tr>
<td>2</td>
<td>45</td>
<td>$36,505</td>
<td>$90</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>$36,865</td>
<td>$120 x 3</td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>$37,065</td>
<td>$200</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>$37,285</td>
<td>$220</td>
</tr>
<tr>
<td>8</td>
<td>39</td>
<td>$37,585</td>
<td>$300</td>
</tr>
<tr>
<td>9</td>
<td>38</td>
<td>$37,905</td>
<td>$320</td>
</tr>
<tr>
<td>10</td>
<td>37</td>
<td>$38,428</td>
<td>$523</td>
</tr>
<tr>
<td>11</td>
<td>36</td>
<td>$39,356</td>
<td>$928</td>
</tr>
<tr>
<td>12</td>
<td>35</td>
<td>$40,411</td>
<td>$1,055</td>
</tr>
</tbody>
</table>

5.2 Dynamic Simulation Scheduling

Using the same information, the project was input into a computerized Dynamic Simulation system developed based on the principles and algorithms presented in the paper. First, minimum resources for project activities are assigned, i.e., establishing the basis for running the Base Simulation where days shortened is equal to 0. Project was simulated using 750 runs, and the results of project schedule are illustrated in figures 7. Figure 7 shows that a 54.55-day project schedule has a degree of confidence of 85% (or a 15% chance of being exceeded). When compared with the results obtained from the deterministic least-cost scheduling previously performed, a 47-day project schedule (obtained when days shortened=0) has a degree of confidence of 12% (or an 88% chance of being exceeded). Thus, based on the stochastic duration output, a 47-day project schedule provides the contractor with only a 12% chance that the project can be completed using this schedule duration and resource configuration.

Following to the analysis of the base simulation, the computerized Dynamic Simulation system was used to shorten the project by 1 to 12 days. Analysis of simulation results was performed for each shortened day. Results of the Dynamic Simulation are summarized in table 2.
While the deterministic least-cost scheduling technique was shown to produce less than satisfactory estimates, it still provides a reasonable guideline for the crashing process steps. The same set of activities was selected though the sequence differed between the two approaches. However, in conclusion, the dynamic simulation approach provides a more comprehensive view of the crashing process that is hard, even impossible, to visualize using the traditional approach.

The reader may have noticed that the analysis have not explicitly addressed indirect costs. Overheads in construction projects are without doubt a fundamental element of the cost. However, the indirect costs are typically tied to and considered a function of project duration [11]. Since both the least cost and dynamic simulation scheduling algorithms follow a step-by-step process for project duration reduction, the indirect costs for each step will be the same for both algorithms. In other words, the indirect cost will not create a different answer when comparing the two algorithms. However, it is understandable when the optimum total project cost is pursued—which is not the focus of this particular paper—incorporating the indirect costs becomes a necessity.

6 Conclusion

The actual times and costs for the constituent activities of a project cannot be determined until these activities are carried out in reality. When performing the project planning function, the deterministic techniques are simply lacking. They consider one possible time/cost value for each activity. And as such, they fail to portray the big picture during the planning phase.

A new stochastic technique is developed to better identify the activities that should be shortened in the schedule crashing process (or schedule relaxation if the opposite is sought). This new technique, referred to by the term Dynamic Simulation, can be used in project planning as well as project control. Both cases of unlimited and limited resources are accounted for. This paper focused on the algorithm used for planning purposes.

The proposed algorithm performs successive discrete-event simulations using multiple configurations of resources and their corresponding duration and cost estimates. Each simulation run, with a unique resource configuration, provides a basis for the progressive crashing processes. At the end, the Dynamic Simulation algorithm not only arrives at an optimally-crashed schedule but also provides the optimum level of resources needed to achieve this goal.

The example project demonstrated that less informative decisions can be made when accounting for the deterministic results alone. The dynamic simulation...
approach stochastically accounts for the different resource patterns, job completion times, and the associated costs. The final decision on how to perform project crashing and whether it is possible to complete project by a certain date becomes handy when the latter approach is adopted.

Finally, the paper utilized a triangular distribution for the illustration of the devised algorithm. The adoption of triangular distributions was due to their practicality. However the dynamic simulation algorithm can function with any distribution of choice e.g. uniform, normal, beta, etc. It comes down to which distribution is more appropriate in a given construction application.

References


Automation for Mobile Crane Motion Planning in Industrial Projects

Zhen Lei\textsuperscript{a}, SangHyeok Han\textsuperscript{a}, Ahmed Bouferguène\textsuperscript{b}, Mohamed Al-Hussein\textsuperscript{a} and Ulrich Hermann\textsuperscript{c}

\textsuperscript{a}Hole School of Construction Engineering, University of Alberta, Canada
\textsuperscript{b}Campus Saint-Jean, University of Alberta, Canada
\textsuperscript{c}PCL Industrial Management Inc., Canada

Corresponding author email: zlei@ualberta.ca

Abstract -
Mobile cranes are used to lift heavy modules in industrial projects. These heavy modules are often prefabricated and each project consists of a large number of lifts (e.g., a typical industrial project may have between 150 and 1000 modules to be lifted). To ensure the safety and efficiency of these lifting activities, crane motion planning is needed. However, in practice, most of the heavy lift studies are, at best, performed semi-automatically and still require significant manual work. In addition to being inefficient, this approach is also characterized by a high error rate, especially in the context of congested construction sites, not to mention its slow response to changes in work order or project scope. For instance, if a module is delayed, the crane motion planning may need to be redesigned (at least partially) since the configuration of the obstacles on the construction site no longer conforms to what had been assumed in the original planning. This paper thus proposes a generic model for mobile crane motion planning that can be implemented in industrial projects. An industrial project with more than 100 modules is selected for validation of the proposed method. 3D visualization is also developed to demonstrate the lifts in a 3D Studio Max environment.

Keywords -
Mobile crane; motion planning; industrial projects; 3D visualization

1 Introduction

Mobile cranes are widely used in today’s construction projects to lift heavy objects, particularly in modular heavy industrial projects. In Canada, industrial modules are often fabricated offsite and shipped to the site as complete objects for installation; (Figure 1 shows an example of a piperack module). This offsite construction of components for industrial projects not only saves time, but also reduces the associated costs and increases reliability throughout the construction process. When modules are installed onsite, mobile cranes are used to perform the lifts. Considering the frequency of using mobile cranes, (e.g., in a typical industrial project, mobile cranes are used to lift as many as 150 modules or more), as well as the cost of mobile crane rental and operational crew, appropriate planning of these mobile crane operations is critical to the success of industrial projects. Mobile crane planning and management usually involve many components, such as: (i) crane type and operation location selection \cite{6, 7, 16, 17, 19}; (ii) crane lift path planning \cite{1, 2, 3, 11, 15, 18}; (iii) crane productivity improvement and equipment design \cite{4, 5}; and (iv) simulation and visualization of crane operations \cite{8, 9, 12, 13}. Among these crane planning components, the crane motion planning and analysis aims to determine whether or not the crane can successfully perform the lift under the given site constraints. This process is particularly important in industrial projects: the more appropriate the crane motion, the smaller the chance of encountering safety issues during the lift and the more efficient and productive the operation. Optimal economy of the project cannot be achieved without analyzing the crane motion and consequently determining the optimal crane motions. Thus, in this paper we introduce an automated system for crane motion planning and analysis that is currently used by PCL Industrial Management Inc. This system assists practitioners in rapidly and automatically planning crane motions and determines the overall economy of the crane operations of projects.

2 Mobile Crane Motion Planning and Challenges

Mobile crane motions can be categorized into two main types: (i) crane lifting without walking (or pick-and-swing), where the mobile crane settles at one
particular location and lifts the module from its pick location to its set location; and (ii) crane lifting with walking (or crane walking), where the mobile crane picks an object, walks with the load, and finally sets it in its resting location. In this context we note that construction sites are typically congested with spatial constraints which encumber crane motions. In industrial projects, the site layout is even more complex and congested than in other construction sites; (Figure 2 shows a typical industrial project). Typical constraints for mobile crane lifting consist of: (i) the geometry of the mobile crane (e.g., the boom of the crane or its tail-swing equipment); and (ii) the constructed structures; (for industrial projects, examples of constructed structures may be the foundations or the lifted piperack modules).

The current practice for crane motion planning is to use heavy lift study, which involves manually checking the feasibility of the given crane lift at the module’s pick and set locations. This process begins with selecting the potential crane location for the lift based on the crane’s capacity; (the crane manufacturer’s capacity chart is often used to determine the distance of the crane body to the lifted module). Following this task, the selected crane location is used for boom clearance checking to ensure that there is no conflict for the boom at the module’s pick and set locations. In the case of crane walking operation, the crane’s walking path needs to be checked to ensure sufficient clearance for the movement. However, the challenges of the current practice are that (i) the current manual process of motion planning is tedious and prone to errors; (ii) the planned results are not sufficiently responsive to site changes; and (iii) industrial projects involve a large number of lifts, which makes the planning process difficult. Therefore, a motion planning methodology is needed in order to ensure rapid and accurate motion planning. In order to achieve this, automation and integration are required: the crane motion planning process needs to integrate different planning components, such as crane location selection and lifting path/walking path planning. Also, the entire process needs to be automated so as to involve the smallest human component possible.

3 Proposed Algorithms

Over the past few decades, a number of research studies have been devoted to crane motion planning. In this context, one of the most important contributions has been to identify the similarity between crane motion planning and path finding in the field of robotics. Many algorithms have been borrowed from the field of robotics and adapted to crane lifting. Lozano-Pérez, for instance, has proposed a configuration space approach by which to deal with 2D object path finding [14]. Based on this idea, Latombe has introduced the methods that are used to plan paths based on the configuration space approach: roadmap, cell decomposition, and potential field [10]. Here the shape of an object is simplified and represented as a representing point and corresponding configuration space obstacles (C-Obstacles). Following this initiative, construction engineers and researchers have begun to build virtual models for both mobile cranes and tower cranes. A recent work by AlBahnassi and Hammad has used the configuration space method to model mobile cranes in a 3D space environment [1]. Four degrees-of-freedom (DOFs) are defined in the model, based on which a Rapidly-exploring Random Tree (RRT) method is used as the search engine to plan the lift path. The configuration space method has also been adopted by Ali et al. to create a search space for cooperative crane lifting, and the genetic algorithm has been used for the purpose of path optimization [2]. Both single- and dual-crane models have recently been developed using the configuration space method by Chang et al. [3]. Although the configuration space method is not novel to crane path planning, rarely has a study combined this method with a heavy lift study method, and crane walking planning has not garnered much
attention among researchers. Therefore, in this paper we introduce a motion planning methodology combining the configuration space method with heavy lifting practice, focusing on the *pick-and-swing* and *crane walking* operations. A mathematical model is built for checking the feasibility of both crane motions, and a detailed planning algorithm for crane motions is developed and implemented in a 3D Studio (3ds) Max environment.

3.1 Pick-and-swing Analysis

In the pick-and-swing scenario, the mobile crane performs the lift without any movement of the crane body. This type of crane lift is analyzed based on the *lifting range*. The **lifting range** can be defined as the range within which the crane can perform the lift without exceeding its lifting capacity. The lifting capacity chart is often used as the reference to determine the **lifting range**. Given any **lifting range**, the furthest distance the crane can reach from the crane location considering the lifting weight is $R_{\text{max}}$, while the closest distance for the crane is $R_{\text{min}}$, which can be expressed as Equation (1).

$$LR = \{(x,y) | R_{\text{min}} \leq d \leq R_{\text{max}}\}$$

Where $LR$ is the *lifting range* that is the set of points that can be reached by the crane based on (i) its capacity as defined in the manufacturer’s lifting chart and (ii) the total lifting weight, including the module (i.e., payload) and the components of the rigging system; and $d$ is the distance between $(x,y)$ and the crane location calculated using Equation (2).

$$d(x,y) = \sqrt{(x-x_c)^2 + (y-y_c)^2} \quad \forall (x,y) \in LR$$

Where $(x_c, y_c)$ is the crane location. The calculated $LR$ can be represented as two circles in a 2D view (see Figure 3). However, on the construction site, the $LR$ needs to be further modified to avoid clashes between the boom/tail-swing and the surrounding environment, satisfying Equation (3). (An example of boom clash is shown in Figure 4.)

$$MLR = \{(x,y) \in LR | BE \cap SOs = \emptyset \textrm{ and } TSE \cap SOs = \emptyset\}$$

Where $MLR$ is the modified lifting range; and $BE$, $TSE$, and $SO$ denote boom envelope, tail-swing envelope, and site obstructions, respectively. The $C$-Obstacles are then generated based on the SOs and the geometric shapes of the modules. Since the modules are often constructed as cubes, the geometric center of the module is selected as the *representing point*. Given the total number of SOs on site, $k$, and any SO from its top view that has $n$ number of vertices, with its corresponding coordinates of $(x_{SO}^i, y_{SO}^j)$ ($i=1 \textrm{ to } k$; $j=1 \textrm{ to } n$), and a module having $m$ number of vertices with its corresponding coordinates of $(x_u^M, y_u^M)$ ($u=1 \textrm{ to } m$) and the coordinates of its geometric center $(x_{GeoC}, y_{GeoC})$, the points of $C$-Obstacle $(P_u(x,y))$ are calculated using Equations (4) and (5). These points are removed from the $MLR$ in order to obtain the crane’s final operation range (FOR), satisfying Equation (6). The module’s pick point $(x_{\text{pick}}, y_{\text{pick}})$ and set point $(x_{\text{set}}, y_{\text{set}})$ should be within the FOR, in which case a lifting path exists. Otherwise the given crane location $(x_c, y_c)$ does not have any lifting path (Equation (7)), which necessitates crane walking analysis.

![Figure 3. Crane $R_{\text{min}}$ and $R_{\text{max}}$](image)

![Figure 4. Example of boom clash](image)
3.2 Crane Walking Analysis

In cases for which the pick-and-swing is not possible for the selected crane location, the crane must walk to complete the lifting task. Three criteria must be satisfied when planning these walking paths: (i) the crane must have sufficient space to pick up the module at the start-walking point (SWP); (ii) at the SWP the crane must have clearance for its crane body and tail-swing equipment; and (iii) the crane should have a collision-free walking path. To satisfy criterion (i), the module’s envelope must not conflict with the SOs, while (ii) requires conflict checking between the crane body and the SOs. The area in which the crane can pick up the modules and also remain collision-free is denoted as the crane feasible pick area (CFPA). Thus, a crane walking envelope is created based on the crane track width (CTW) and crane track length (CTL), and the length of the crane envelope is kept infinitely long and rotated counter-clockwise, as illustrated in Figure 5(a). In Figure 5(a), the crane needs to perform the lift at the crane location (CL); there are three SOs on site and the crane has two CFPA\(s\). When the crane envelope starts rotating, at some angles the crane walking envelope clashes with the SOs, as in the example in Figure 5(b). This means that, at this angle, the crane cannot reach the CFPA without overlapping with SO\(1\); and thus the walking path does not exist. Figure 5(c), alternatively, shows a scenario in which the crane can reach the CFPA, and which, consequently, can be considered as a potential walking path. All the areas that are possible for crane walking are then generated as the shaded areas shown in Figure 5(d).

![Figure 5. Crane walking envelope and walking planning](image)

3.3 Crane Motion Detailed Planning

After checking the feasibility of the pick-and-swing and crane walking motions, the detailed motion planning is performed for the purpose of visualization. This planning consists of: (i) pick-and-swing; (ii) crane walking; and (iii) instant clash detection and responsive action mechanism.

(i) Pick-and-swing Detailed Planning

Given that the crane sits at one fixed location, \(CL(x_c, y_c)\), the lifted module (object) needs to be lifted from its pick point, \(PP(x_p, y_p)\), and placed at its set point, \(SP(x_s, y_s)\). In mobile crane operation
practice it is considered that, during the lifting process, the lifted module should remain as close to the ground elevation as possible. Thus, the angle $\theta(PP \rightarrow SP)$ can be calculated using Equation (8).

$$\theta(PP \rightarrow SP) = \cos^{-1}\left(\frac{\overrightarrow{CP} \cdot \overrightarrow{CS}}{|\overrightarrow{CP}| \cdot |\overrightarrow{CS}|}\right)$$  \hspace{1cm} (8)

Where $\overrightarrow{CP}$ is a vector $(x_p - x_c, y_p - y_c)$ and $\overrightarrow{CS}$ is a vector $(x_s - x_c, y_s - y_c)$. The crane lift path, $\varphi_1(PP \rightarrow SP)$, is checked by means of the instant clash detection and responsive action mechanism.

(ii) Crane Walking Detailed Planning

In addition to the $PP$ and $SP$, the crane walking involves the crane start walking point, $SW(x_{sw}, y_{sw})$. Thus the $\varphi_2(SW \rightarrow CL)$ must also be checked using the instant clash detection and responsive action mechanism.

(iii) Instant Clash Detection and Responsive Action Mechanism

As mentioned, in both the $\varphi_1$ and $\varphi_2$ scenarios, clash detection needs be conducted. Once a clash is detected, an action, $u$, is needed to adjust the crane motions. The variable, $u$, denotes all the actions that can be performed by the mobile crane, including: (i) boom up; (ii) boom down; (iii) hoist up; (iv) hoist down; (v) rotate following $\theta(PP \rightarrow SP)$ direction; and (vi) rotate the module clockwise. Meanwhile, we define that in the 3D environment there is a finite state, $X$, of the mobile crane, and that each instance of $X$, denoted as $x$, represents a collision-free configuration of the mobile crane and its lift. When a clash occurs, $u$ must be applied in order to avoid the clash, and $x$ changes to $x'$, as in Equation (9).

$$x' = f(x, u)$$  \hspace{1cm} (9)

Ultimately, the objective is to obtain a sequence from $X_f$, namely, the initial state of the crane, to its $X_g$, which is the goal state. A detecting mechanism based on the ray tracing method is developed for instant checking of the potential clash (see Figure 6). In Figure 6, there is a moving object (MO) at any state, $X_m$, which needs to be checked for clashes with the site obstruction (SO). Given that the MO has $n$ vertices, $P^{MO}_i$ ($i = 1$ to $n$), and the SO has $m$ vertices and surface geometric center point, $P_{j}^{SO}$ ($j = 1$ to $m$). The set of distances ($D$) from each $P^{MO}_i$ to $P_{j}^{SO}$ is then calculated, based on which the shortest distance, $Min.D$, can be found. The found $Min.D$ should be less than the minimum allowed clearance (e.g., 1 m); otherwise, $X_d$ is detected as a clash. During implementation, the MO is replaced and checked for clashes with (i) the lifted module; (ii) the crane body and boom; and (iii) the crane rigging and other attached components.

Figure 6. Ray tracing method for clash detection

3.4 System Automation

The above proposed algorithms cannot achieve high efficiency without automation. In this research, automation is achieved by embedding the proposed methodology into a programming environment that can loop through the entire lifted module and crane configuration automatically. The implementation markedly reduces the human component involved in the planning process. A pseudo code for the automation implementation is shown in Figure 7.

4 Implementation

In Alberta, Canada, numerous industrial projects are constructed for the extraction and processing of oil and gas resources. Modular construction, whereby modules are prefabricated off-site and shipped to the site for onsite installation by mobile cranes, is a widely used method by which to complete this type of project. Meanwhile, such projects are difficult to complete due to the fact that they (i) have a large number of lifts; (ii) are usually complex; and (iii) due to the lengthy construction period, involve frequent project changes. These characteristics make crane motion difficult. Mobile crane operations, moreover, can be expensive if not planned well.

Facing this challenge, PCL Industrial Management Inc. in Alberta, Canada has been working with researchers at the University of
Alberta’s Hole School of Construction Engineering seeking solutions for planning crane motions. This research has been developed and implemented in actual projects built by PCL Industrial Management Inc. A case is selected to demonstrate the designed system and the planning results. The pick-and-swing analysis is built in a VB.Net® programming environment and shares information with an internal company database that stores project and crane data. The detailed crane motion planning algorithm is implemented in 3ds Max by coding in MAXScript (a built-in platform available to developers in 3ds Max).

MOTION PLANNING AUTOMATION

forall crane \in C

CLs \leftarrow UCLs

forall CL \in CLs

Pick-and-swing()

if Pick-and-swing.result = SUCCESS

Crane-Motion-Detailed-Planning()

else crane-walking()

if crane-walking.result = SUCCESS

Crane-Motion-Detailed-Planning()

else at CL \rightarrow lift impossible

end if

end if

*Crane = one crane configuration; C = set of all crane configuration; CLs = crane locations for one Crane; UCLs = universal crane locations for C; CL = one crane location belonging to CLs.

Figure 7. Pseudo-code for crane motion planning

Figure 8 shows a site layout presented in the built VB.Net environment. The rectangular boxes are the lifted modules with the site boundaries. In the presented project, there are a total of 127 modules. The pick-and-swing and crane walking analyses are automated, and all the possible crane locations are analyzed. A total of 9,770 crane locations are analyzed for four different crane configurations, with the entire process taking 6 hours to complete. Figure 9 shows a sample of the analysis results for one crane configuration (Demag-CC 2500 with Superlift equipment), where the red points are the crane locations that do not have lifting solutions and the green points indicate locations from which the mobile crane can perform the lifts. According to the overall results, it is found that the Superlift equipment limits the rotation of the mobile crane, thereby leading to the failure of many crane locations, particularly in congested areas.

Detailed motion planning is conducted in 3ds Max. Figure 10 shows an example of clash detection. Here it can be seen that an obstacle is encountered during the crane swing; in order to avoid the conflict, the crane must lift up the module. Another example is presented in Figure 11, in which a crane walking path is planned in detail.

5 Conclusions

This paper has presented a crane motion planning methodology that can be applied generally to industrial projects built based on an off-site modular construction method. This research has built on existing studies focused on crane lift path planning by considering two previously unexplored aspects: crane walking algorithms and entire site motion planning. The research presented in this paper has used the configuration space method developed in the field of robotics, and the built system has been proven efficient for entire site motion planning and visualization. We are presently expanding the work scope such that the future system will be able to analyze the cost (penalty) of motion plans, and...
thereby will seek optimal plans for crane operations.

![Figure 10. Clash detection](image1.png)

Figure 11. Crane walking planning

![Figure 11. Crane walking planning](image2.png)

6 Acknowledgements

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development (CRD) program. The support of our partner for this research, PCL Industrial Management Inc., is appreciated. Mr. Jonathan Tomalty’s assistance in editing is also much appreciated.

References

CONSTRUCTION MANAGEMENT


A Study on Significance of System Dynamics Approach in Understanding Adoption of Information Technology in Building Construction Projects

Vinay Mathews, Koshy Varghese, and Ashwin Mahalingam

Building Technology and Construction Management Division, Civil Engineering Department, Indian Institute of Technology Madras, India
E-mail: vinaymathews30@gmail.com, koshy@iitm.ac.in, mash@iitm.ac.in

Abstract –

Today several activities of a construction project can be supported with information technology tools. However, there have been many cases of technology adoption failure. Lack of understanding about the process of technology adoption is often pointed out as the reason. This paper is based on a study to explore BIM technology adoption process and attempt to model and evaluate the process using system dynamics. The paper has its base on literature survey and five month field studies of three construction projects where BIM was being introduced. Insights from literature survey and case studies were used to develop a core system dynamics model of technology adoption. To this core, project and technology specific data was added to develop a preliminary system dynamics model. The research indicates that system dynamics is a promising approach in improving our understanding of technology adoption in construction projects.

Keywords –
System dynamics; Construction projects; Information technology; Adoption

1 Introduction

Usage of Building Information Modelling (BIM), Enterprise Resource Planning (ERP) and other information technology tools are rapidly growing in construction industry. Suitability and effectiveness of such tools in a given construction project are uncertain.

Technology developers can guide construction companies in understanding the advantages of using technology. But the complexities that emerge in the learning phase of technology adoption are still unclear.

Acceptance and adoption of information technology has been a popular research issue [1]. Over the past two decades, researchers have developed many frameworks to help decide on the suitability and acceptability of technology. Technology Acceptance Models (TAM) are popular in predicting and explaining technology use [2, 3]. But they concentrate on technology adoption by individual users. TAM models are not suitable to study adoption of collaborative technology in construction projects, where the use by an individual user is not significant.

Recently, there have been studies that looked into organisational aspects of technology adoption in construction industry [4-6]. But they are yet to find place in any of the IT adoption models.

Many studies were taken up to understand technology adoption specific to construction industry. Most of these studies focus on identifying new factors that plays a crucial role in technology adoption. While these exploratory studies are relevant in identifying influencing factors, there is no guidance suggested to a project manager on which factors to focus on while selecting and adopting a technology [7]. Further, correlation between these factors is not known and the alignment of factors required for successful technology adoption is also not tested.

System Dynamics (SD) is an approach used to understand behaviour of complex systems over a period of time [8]. The use of SD to model construction projects is not new [9]. SD has also been used in fields like agriculture to understand technology adoption [10]. In this paper, a preliminary evaluation of system dynamics to model technology adoption in construction projects is explored.

The paper is a part of a larger research that looks at improving technology adoption in construction projects. The first author spent 5 months as an observer on three construction projects that were initiating the use of BIM. Several observations of the BIM adoption process were made during this field study phase. This paper attempts to structure these observations as well as aspects reported in literature to understand the process of technology adoption and represent this using SD.

The next section of this paper presents the findings
from reviewed literature to understand the current state of knowledge on technology adoption. The details of three field studies conducted and the observations made are presented in Section 3. The insights from case studies have been detailed in Section 4. Section 5 discusses the development of the system dynamics model based on the insights on technology adoption from literature and field studies. Section 6 discusses the utility of the system dynamics model in understanding technology adoption process.

2 Literature Review

Among the many studies on information technology adoption, Technology Acceptance Models TAM, TAM2, UTUAT are among the most popular.

Unified Theory of User Acceptance and Use of Technology (UTUAT) [11] is a highly cited work on technology adoption. The paper summarises eight user acceptance models and also proposes a unified theory bringing together the best features from all those.

Eight models that are unified in the paper are Theory of Reasoned Action (TRA), Technology Acceptance Model (TAM), Motivational Model (MM), Theory of Planned Behaviour (TPB), Combined TAM and TPB (C-TAM-TPB), Model of PC Utilisation [12], Innovation Diffusion Theory (IDT) and Social Cognitive Theory (SCT).

TRA, proposed by Fishbein and Ajzen [13] suggests that the behaviour of an individual is affected by one’s own attitude and subjective norm. Subjective norm is an individual’s perception of what behaviour is expected of him from others.

TAM, which has its basis on TRA, included constructs like usefulness and usability of a technology in predicting the chances for successful adoption of technology [14]. TAM initially did not take the concept of social norm from TRA. However TAM2 included subjective norm to it after studies proved its importance.

MM describes intrinsic and extrinsic motivation for behaviour. Intrinsic motivation refers to the motivation that drives a particular behaviour to enjoy the behaviour itself. External motivation refers to the motivation which is driven by the end result. It also tries to explain the interrelationship that intrinsic and extrinsic motivation has [15]. Social Cognitive Theory (SCT) by Albert Bandura looked in to learning from observation and action. The concept of self-efficacy was also brought out in SCT. Self-efficacy is the self-belief of a person in completing a given task [16]. TPB is an extension of TRA, with the idea of self-efficacy borrowed from SCT. C-TAM-TPB combines the best of TAM and TPB [11]. IDT describes the process of diffusion of an innovation in a group [17]. MPCU tries to understand the factors that affect the use of personal computers (PC). Social norm and expected consequences of using PC were found to be the major factors that influence PC usage.

Venkatesh et al’s [11] summary of all acceptance models could be given a pictorial representation as shown in Figure 1.

![Figure 1. Underlying concept of all acceptance models](image)

TAM models, with their origins in psychology, concentrate on adoption of technology by individuals. While it is good to have an understanding of technology adoption by each user, system level understanding of technology adoption would enable decision making on selection and improvement of technology adoption. Details on technology adoption by individual users in a time bound project setting may not be of much use.

Technology developers embed a process to the technology. The closeness of this process to organisation’s work practices has a significant influence on the ease of implementation. Many of the research works agree on the need to select technology that is closely aligned with the existing work practices [18, 19].

If both organisation and technology were to be rigid, selection of proper technology would have been the only area that researchers could look into. Building on the existing literature of structuration theory [20], many other theories have come out which endorses the fact that technology and organisational structures are not rigid and can actively engage in restructuring one another [21, 22]. This brings out the importance of activities surrounding technology development and technology implementation in determining the acceptability of technology. Whyte [6] elaborates on hybrid practices that emerge when digital technology is introduced in to an institutionalised organisation. Dossick and Neff [5] discuss the importance of messy talk for effective communication and problem solving in projects involving multiple teams. They define messy talk as "conversations neither about topics on meeting agendas, nor on specified problems or specific queries for expertise". Technology can create trading zones, which increase collaboration within various teams and thereby becoming an integral part of the organisation [4, 23]. Many of these recent findings have not yet found place in any of the frameworks for successful technology adoption.
There have been exploratory studies to find both technical and nontechnical factors that affect technology adoption in construction industry. Sargent, K et al extended UTUAT by adding the element Top Management Support [24]. Peansupap et al conducted an exploratory study to come up with factors that influence adoption of IT in Australian construction organisations [25]. These factors were 11 in number, classified to four categories – Individual, Environment, Management, and Technology. Nikas, A et al [26] describes the importance of considering antecedents of technology adoption separately from drivers of technology adoption.

In the above studies, the authors have either stated the explored factors as independently or positioned them as extensions of any of the existing technology acceptance models. While it is essential to have exploratory studies in the initial stages, it is equally important to have a framework that can comprehensively accommodate their results. A framework is essential to develop models that can be used to make sense of the entire system. Currently, there are no frameworks that are specifically developed to understand technology adoption in projects.

The data from field study were used to understand the dynamics of technology adoption in a construction project. This understanding was later used to develop a framework for technology adoption model suitable to construction industry.

3 Field Study

Field studies were conducted in three construction sites located in three different cities in India over a period of 5 months. The purpose of these field studies was to observe the technology adoption in these construction sites.

The first author spent five months in these three sites. The inputs for these studies mainly came from interviews, observations in construction sites, attending weekly meeting on BIM and MEP coordination meetings.

The client for all the three projects was a leading IT company in India. All projects were executed on a fast track basis, meaning – construction process was concurrent with design process. In each of these sites, a Project Management Consultancy (PMC) was appointed for overall planning and coordination of the works. Separate organizations for architecture, structural design, MEP design were a part of the project team. A main contractor (MC) did the civil works. Specialised MEP sub-contractors were nominated by the client, but reported to main contractor.

BIM Agency (BIMA) was appointed in all the three cases. Their responsibility included developing 3D model from design drawings, clash detection and clash resolution of MEP drawings. They also had to prepare weekly status report of the construction activities based on inputs from construction site. In one of the sites, BIMA reported directly to the main contractor, whereas in other cases, BIMA reported to the owner.

Table 1 gives a brief description of the similarities of the cases studied.

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Building Owner</td>
<td>IT Building Leading Indian IT Company</td>
</tr>
<tr>
<td>Project Delivery Method</td>
<td>Design Bid Build</td>
</tr>
<tr>
<td>Time of Study</td>
<td>Construction Phase</td>
</tr>
<tr>
<td>Built up area</td>
<td>Approx. 2 M SQF</td>
</tr>
<tr>
<td>Location</td>
<td>Across India</td>
</tr>
<tr>
<td>BIM modules used</td>
<td>BIM Clash Detection, BIM Status Report</td>
</tr>
</tbody>
</table>

3.1 Clash Detection Using BIM

In all three cases, manual clash detection of MEP services was to be replaced with BIM based clash detection. Although not in a deliberate fashion, concepts from Socio-Technical Systems Design (STSD) literature [27] were utilised for study project design in all the three cases. STSD highlights the importance of considering people, technology and context in a holistic manner. In accordance with STSD, technology change was implemented with associated changes in people and organisational structure.

MEP designer continued to make designs in CAD and main contractor used CAD in execution of construction work. BIMA took drawings from MEP consultant, architect and structural consultant to develop BIM model from the 2D drawings. After clash detection, clashes were resolved in discussion with MEP consultant, architect, structural consultant, main contractor and sub-contractors. The proposed workflow seemed appropriate in the beginning. With least possible interference made with institutionalised Indian construction practices, BIM was to be used for clash detection and resolution. But, things did not happen as expected. Each of the three cases had different responses.

3.1.1 Case 1

BIMA had to make 20 revisions of first floor BIM model before releasing good for construction drawings.
The unexpected BIM revisions delayed the start of execution of first floor MEP works by 15 days. The reason for one such revision was that BIMA had left out AC duct insulation in the BIM model. The 2D drawings that were given by main contractor to BIMA did not have AC duct insulation. BIMA, whose job is to develop the BIM model of given 2D drawings, did not consider it necessary to model elements that are not present in the 2D drawings. Main contractor did not include insulation of duct in 2D drawings as that was the convention that the industry has been following. The main contractor believed that it was lack of technical knowledge of BIMA which lead to this problem. From an outsider perspective, the presence of a proper information sharing protocol could have avoided this.

The PMC on site insisted that MEP works could start only after BIM clash free drawings were released. PMC head said that he was under tremendous pressure to release drawings. None of the MEP sub contractors wanted their engineers and workers to stay idle in site, waiting for release of drawings. After the release of first floor drawings, decision was made to discontinue mandating BIM generated clash free model as a necessary step before execution.

3.1.2 Case 2

BIM clash detection and resolution began much earlier in this site. BIMA representative on site coordinated many clash resolution meetings on site. Members from design offices and BIMA participated in many of these meetings using teleconferencing and screen sharing. But clash resolution process was not completed as estimated. In one case, BIMA prepared the model with old versions of a few drawings. It was during clash resolution meetings that other parties pointed out this discrepancy to BIMA. PMC, who was in charge of forwarding design drawings to BIMA, blamed BIMA for not incorporating the changes. BIMA denied receiving these revised drawings from PMC. Despite exceeding the initial time estimates, clash free model of first floor drawing was released without causing any delay in MEP work execution. Having taken up the entire buffer for first floor, it was anticipated that BIM clash free drawing of subsequent floors would be delayed. So, it was decided that BIM clash free drawing will be used only if it does not delay start of execution of MEP works.

3.1.3 Case 3

Despite many difficulties, BIM clash detection continued to be a prerequisite before execution. The quality of the clash free model was not good enough. An onsite MEP engineer said, "BIM doesn’t give clash free model. It is in site that all the clashes are resolved. “Another MEP engineer said, "The BIM coordinated drawing when compared to the old style of manual coordination does not bring any change. Not better not worse.”

There was a clear disconnect between PMC and MC in this site. Given below is a snippet from a site meeting called by PMC to communicate the necessity to have as built drawing.

PMC: "Whenever we say something you say against it. Every time you are talking negative....... We will have to cut your payment"

MC: "Every time you are cutting our payment. Show in contract document where it is written to submit as built drawings."

PMC did not react to what was happening in site. PMC continued to insist the need to have BIM clash free drawing.

3.2 Status Report Using BIM

In all the three cases studied, weekly project status report was prepared by BIMA from beginning of construction phase. BIMA used baseline schedule prepared by Project Management Consultancy and the updates from contractors for preparing the graphical status report. The report was used by client team to monitor the delays closely. The site team, actively engaged in construction process, did not need a BIM report to understand progress. Sitting in his site office, a PM gave the following remark on the usefulness of status report using BIM, "I move these curtains and I know the status of this bloody building".

This monitoring was not effective because the level of detail of the 4D BIM model was inappropriate for weekly monitoring in all of the sites. In original base line model, concreting of all 100 columns in a floor was represented as a single activity. The BIM status report could only show the progress of work for the entire set of 100 in a single state. So, even when a single column of the 100 was not completed, all 100 shown to be in progress. This miscommunication of actual status resulted in frequent queries on slow progress by management. Hence the BIM report required frequent explanation for the differences in BIM model and the actual progress of work. To resolve this miscommunication it was decided that baseline schedule would be expanded to include individual column-wise details. This ensured that BIM status report was up-to-date with actual construction to the detail of a single column. PMC, MC and BIMA had to do extra work to incorporate these changes. Project schedule such minute
details was new to all the project participants.

Although the site faced initial problems with the BIM report, there was no indecisiveness on the need to have BIM status report in any of the three construction sites. After appropriate level of detail was modelled, accurate BIM status report was generated in all the three sites. The client team, which was away from the site, could better understand the site progress.

4 Discussion on field study

This section tries to understand the dynamics involved in IT adoption in the three cases. The dynamics so identified has been used in Section 5 for developing a system dynamics framework of technology adoption in construction industry.

Two underlying features of technology adoption process in the cases: 1. constant conflict between existing and new work flows. 2. variation of factors across adoption time

4.1 Constant conflict between existing and new work flows

There are several uncertainties in the learning phase of technology adoption. Figure 2 is a pictorial representation of the resulting adoption trajectories. Trajectory 1 represents case 3 in clash detection section. Here, the effort to use a technology continues, without the desired outcome. Trajectory 2 represents the successful status report generation module of BIM in all the three cases. In this trajectory, all the complexities that came across were overcome. Case 1 and case 2 of clash detection module are represented by trajectory 3.

The figure reveals one feature that is often left out in most of the technology acceptance literature. There is always a tendency for a project to discontinue the usage of a new technology and go back to old technology. Institutional theory can be used to explain the bearing a project has from previous projects and experience of its participants [28, 29]. Though each project is unique and can be designed individually, there is always a power struggle between the comforts of the established norms and benefits of new design tools.

The observation matches with ‘Relative Advantage’ concept by Rogers [17] to an extent. Relative Advantage is the perceived advantage of an innovation when compared to the conventional one. The main difference of ‘Relative Advantage’ with the observation here is the dynamicity associated.

4.2 Variation of factors across time:

In the three cases listed in Section 3, BIM was initially used for clash detection. Among the three, two of the projects discontinued its usage later. This discontinuation indicates that there have been some changes in factors that affect technology adoption as project progressed.

![Figure 2: Three trajectories of technology adoption](image)

The Project Manager’s and field engineer’s commitment to use BIM is likely to diminish with time if the results are not useful for their core function. Another aspect which could change with time is the standards and protocols of information exchange.

5 System Dynamics Model

System Dynamics is a modelling and simulation technique to understand the behaviour of complex systems over the period of study. System Dynamics does not look into details of individual participant, but looks things at a system level. This makes SD suitable for developing a framework for technology adoption in construction industry.

In this section, the details of formulating a system dynamics model of adoption of BIM adoption have been discussed. The development consisted of two phases:

1. Build a framework for a model that could be used irrespective of the project and technology under consideration

2. Use project specific data to complete the model

Two features formed the core of the framework of system dynamics model. One based on literature and other on the case study:

a. A cyclic relation between intention to use a technology and actual usage at project level. (From Figure 1 in Literature Review Section)
b. Constant power struggle between new workflow and the existing work flow (from case study)

In figure 3, ‘Actual new workflow rate’ and ‘Actual conventional workflow rate’ have been used to represent the existence of two conflicting workflows. Also, variables ‘Actual new workflow rate’ and ‘Intention to use new workflow’ have been cyclically connected by variables ‘Total work done in new workflow’ and ‘Relative Advantage’.

Figure 3. Proposed framework for SD model on technology adoption

The framework is an abstract SD model that could be reused for any technology adoption process in construction industry. It should be possible to enter project and technology specific details to the abstract model.

A system dynamics model was developed from this abstract model to validate the extensibility of the abstract model. As development of system dynamics model is more at the discretion of modeller, a model developed at a particular site by the researcher could be biased in favouring the universality of the abstract model. To overcome this, the model was developed based on an international survey by Won et al [7] which looked into critical success factors (CSF) for BIM adoption. Twenty four CSF emerged from the study, which involved 52 expert responses from 4 continents. The study included weight of each of these CSF on a scale from 1 to 7. It was observed that the all the CSF were contributing either to ‘Intention to use new workflow’ or ‘capability to use new workflow’ or both. None of the CSF had to be left out from the model. The classification of these factors is given in Table 2.

Together with values from the study and a few assumed values for variables like ‘estimated new workflow rate’ and ‘estimated conventional workflow rate’, hypothetical SD models were developed for adoption of BIM clash detection and status report.

Table 2. Influencing Factors [7]

<table>
<thead>
<tr>
<th>Factor</th>
<th>IU*</th>
<th>CU^</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willingness to share information</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Master BIM Manager</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Effective Collaboration</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Organisational Structure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous Investment</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Senior Management Leadership</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Information Sharing Protocol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIM Training Program</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Technical Support for Interoperability</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Standardised Work Procedures</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>PM Interest</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Request from client</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Project Complexity</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Field Engineer’s interest</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Architect Firm’s use</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Result Demonstrability</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Expected ROI</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Company’s business strategy</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Usefulness of technology in project</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>BIM technology capability</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Known successful cases of BIM</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Software Interoperability</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Institutionalisation of BIM application</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Modelling ease</td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>

* IU – Intention to Use Technology
^ CU – Capability to Use technology

Figure 4. Representation of a CSF being accommodated in abstract model

6 Discussion on System Dynamics Model

As there are many available features in system
dynamics software, the insights that one can get from a SD model are many. Here, only two key benefits are discussed:

1. When the two hypothetical models were combined, the impact of timing of introducing the two modules – ‘clash detection’ and ‘status reporting’ could be analysed. This is an indicator that decision makers could use system dynamics models to understand the significance of timing of introducing technology modules.

2. Huge sets of inputs can be programmatically given to the system dynamics model by systematically varying the values of different variables. System dynamics software could use these inputs in generating corresponding set of project outcomes. Project designers can now fine tune their designs based on the required output, unchangeable inputs and changeable inputs.

7 Summary

A reusable system dynamics based framework (Figure 3) that can be used for modelling technology adoption process in construction industry has been proposed based on literature review and field study. The paper has demonstrated the feasibility of utilising system dynamics based approach to develop a model of technology adoption in construction projects.

References


[14] Venkatesh, V. and Davis, F.D. A theoretical extension of the technology acceptance model:


Learning Effect of Interior Finishing and Building Services Works for Multi-Dwelling Complex Project

N. Mine\textsuperscript{a}, A. Olanrewaju\textsuperscript{b}, and M. Takeuchi\textsuperscript{c}

\textsuperscript{a, b}Faculty of Engineering and Green Technology, University of Tunku Abdul Rahman, Jalan Universiti, Bandar Barat, 31900 Kampar, Perak D.R., Malaysia

\textsuperscript{c}Architect, Environment & Technical Solution Division, Shimizu Corporation, 2-16-1, Kyobashi, Chuo-ku, Tokyo 104-8370, Japan

E-mail: naoto@utar.edu.my, olanrewaju@utar.edu.my, m.takeuchi@shimz.co.jp

Abstract -

In the construction industry, many individual trades have to work simultaneously. This creates many potential difficulties on the construction site, primarily in relation to efficiency loss and construction period extension. To improve efficiency, a new interior finishing system for dwelling was developed. By using the system, both interior finishing and building services system are well combined. It is a well-known fact that the repetitive nature of construction produces a learning effect, which reduces both the overall working period and labour man-hours utilized. However there is little learning data available that is specifically related to interior works. Therefore, using the developed method, an observation study was conducted in order to collect learning data in an 11 storey housing construction project. The analysis demonstrates that a typical learning effect was found in 11 recorded repetitions.

Keywords

learning curve effect, labour productivity, interior finishing work, building services work, multi-dwelling complex

1 Introduction

This study aims to integrate interior finishing works and building services works in a Multi-dwelling complex construction. An interior finishing work consists of various trades’ activities and they are carried out simultaneously. This causes interferences among trades, which results in low productivity in a project. To solve the problem a new interior construction method has been developed in which a carpenter carries out most of the interior finishing as well as building services works such as electric wiring and plumbing. The productivity of the interior finishing works has been reported in a previous paper [1].

In the construction planning for a building project, the manpower plan is carried out by considering learning or experience curve effect of labours. There are many studies to confirm the learning curve effect [2][3]. However, such studies are lacking for Multi-dwelling complex interior construction. This paper discusses learning or experience curve effect on interior finishing works by repetition (the proposed system should encompasses both interior finishing and building services, but due to circumstances, this study is limited to only interior finishing). To obtain man-hours data of interior finishing workers, actual projects were observed by using a newly prepared activity recording sheet.

2 Objective and Scope

The authors developed an interior finishing and building services system, called “panel system”, for Multi-dwelling complex. The purpose of the system is to reduce manpower and construction time in an interior finishing work, and is explained in 3.3.1 in detail. This paper studies the system in an actual interior finishing work. Basic data such as man-hours for each activity, work process, and construction time are recorded and analyzed. Through the study, basic data to improve the construction system and planning of future project were obtained. Furthermore, we identified reduction of construction time and man-hours by learning curve effect as a sequel to repetition of work.

3 Observation Items and Methods

3.1 Observation Items

The authors conducted observations on interior finishing and worked on a Multi-dwelling complex construction project. The observation items are:
(1) Problems in the “panel system” during construction,
(2) Processes in interior finishing work,
(3) Changes in construction time and man-hours of each
activity in each floor.

3.2 Observation methods

3.2.1 Man-hours

To obtain accurate data during the whole construction
period, the authors prepared a specially designed activity
recording sheet which records dates, activities, and
locations every day. The recording sheet is not comparable
to time study in accuracy, but is useful to measure man-
hours for a long period of entire project term. In the sheet,
one day is divided into three time zones; morning,
afternoon, and overtime; the main activities in each time
zone are recorded. The activity in each time zone is to be
selected from several predefined ones, by drawing a circle
in the cell below the selected activity’s label; activities not
among the predefined ones are to be written in a cell under
the label “Miscellaneous”. Each sheet holds the record of
a week, which consists of six days (i.e. excluding Sunday).

In our study, recordings were performed by a selected
worker during rest in each time zone. At the start of the
observation period, the authors had several meetings with
supervisors, site engineers, foremen from sub-contractors,
and workers to initiate them to the objective and
significance of the observation. An appropriate worker
was selected after the meetings. The selected worker
recorded his activities during the entire construction
period; this did not burden him since each recording took
less than a minute. The authors visited the project site to
copy the records once a week. At that time, problems on
recording activities were discussed. Measures were taken
immediately for the resolvable problems. The records
were taken from the 2nd to the 11th floor inclusively.

Figure 1 shows an example of a completed activity
recording sheet.

3.2.2 Schedule

The project schedule was extracted from the records in
the activity recording sheet. The dates of the activity start
and end were read from the activity record, and the data
was processed by using scheduling software.

3.3 Subject of the observation

3.3.1 Outline of the construction method

The construction method was developed for the
purpose of improving construction productivity and
building services equipment maintenance works. To
achieve the purpose, dwelling interior finishing works and
building services works were incorporated properly. The
basic element of the system is a gypsum board panel
combined with wooden studs (in short “panel”). The panel
is used as an element of partitions in a dwelling. The
dimensions of the gypsum board are: 12.5mm thickness,
910mm width, and slab height. Wooden studs are 40mm ×
35mm section.

Figure 2 shows a standard panel. After installing
backing frames, panels are installed, and then electric
wiring and plumbing works, finally folded gypsum boards
are installed. Based on the panel system, other interior
finishing and building services works such as water
supply, sewage, and electric wiring works are integrated.
By integrating the construction method and introducing
multi-skilled worker, idle among work trades and work
process are eliminated, which are common problems in
most construction projects. Hence, work efficiency is
improved.

Figure 3 shows a corridor partition which is a basic
usage of panels. This type of plan is the most popular for
current Multi-dwelling complexes in Japan. As shown in
the figure, main building services such as water supply
pipes and electric wires are installed in the ceiling cavity
which allows simplification on piping and wiring to each
room. Although the system makes the distances of wiring
and plumbing longer than conventional methods, it also
makes work easier.
The system simplifies interior finishing and building services works for Multi-dwelling complexes. In the development of the method, the authors carried out two mockup experiments; through the experiments, technical problems were identified and solved before the system’s actual deployment.

### 3.3.2 Outline of the construction project

Table 1 shows the outline of the project to which the panel system was applied. The project is a middle height Multi-dwelling complex, located in a housing area. Before commencement of the project, the authors tried to introduce the developed panel system. However, due to contract conditions and maintenance of building services equipment, the building services works part of the panel system cannot be applied, and is left out of this study.

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building use</td>
<td>Multi-dwelling complex</td>
</tr>
<tr>
<td>Location</td>
<td>Chiba, Japan</td>
</tr>
<tr>
<td>Structure</td>
<td>Reinforced concrete</td>
</tr>
<tr>
<td>Storey</td>
<td>Sub-structure: 14 storey, Basement: 1 storey</td>
</tr>
<tr>
<td>N. of dwellings</td>
<td>137</td>
</tr>
<tr>
<td>Building height</td>
<td>40.7 m</td>
</tr>
<tr>
<td>Site area</td>
<td>14,892 m²</td>
</tr>
<tr>
<td>Building area</td>
<td>1,673 m²</td>
</tr>
<tr>
<td>Total floor area</td>
<td>16,340 m²</td>
</tr>
<tr>
<td>Average area of a dwelling</td>
<td>93 m²</td>
</tr>
<tr>
<td>Interior finishing method</td>
<td>Timber/steel backing frame &amp; gypsum board</td>
</tr>
</tbody>
</table>

### 4 Results of the observation

#### 4.1 Analysis of actual schedule

##### 4.1.1 Actual schedule

For the observation, a typical dwelling was selected and activities were recorded from the 2nd floor to the 11th floor, using the activity recording sheet. The actual schedule was also recorded using this sheet. The selected dwelling was located at a gable-end side. The object of the observation was a 93m² floor area consisting of three bed rooms, a living room, a dining room, a kitchen, and a
toilet/bathroom. The room setups from the 2nd to the 13th floor are the same, with the exception of special experiments at the 12th and 13th floor. Therefore, activity records were taken at 2nd to 11th floor. The work repetition was ten times. The numbers of repetition is enough to analyze learning effect.

Fig.4 shows an actual schedule drawn with scheduling software using source data obtained from the activity recording sheet. In Figure 4, the actual schedule for each activity is represented from top to bottom. The black bar with triangles at both edges shows one floor’s schedule from start to finish; it indicates the whole period of each floor including non-working days, that is, in “calendar days”. Hence, the entire construction starts from the marking for the 2nd floor until the marking for completion of 13th floor, giving a total of 184 days, with 109 clear days. Works on the 13th floor finished on 18th February 2009. The activities of the 2nd floor activities are shown at the top of the chart. Most of the activities were completed in August and September. However, curtain box, counter, and several other activities were carried out in October and November 2008. As a result, whole construction period became 63 calendar days and clear working day was 17 days. There is a big difference between the two.

4.1.2 Change of construction days for each floor

Figure 5 shows calendar days and clear days for each floor from the 2nd to the 11th floor. As shown in Figure 5, the calendar days for each floor’s construction were reduced according to the construction progress. They changed as follows, 63 days at the 2nd floor, 20 days at the 3rd floor, 15 days at the 4th floor, and 12 days at the 6th floor. However, each floor construction days increased from 14 to 26 days from the 7th to the 10th floor. From the 9th to the 10th floor, the calendar days include the Japanese New Year holidays. Hence, the calendar days for the 9th and 10th floor increased.

On the other hand, clear days for each floor reduced form the 2nd floor to the 8th floor according to the construction progress. For this reason, we consider that learning curve effect has occurred, being caused by repetition of same activities. However, the clear days for each floor increased after the 9th floor. Again, this is due to the holidays.

4.1.3 Learning effect of construction days

In this section, learning curves are calculated based on the above data. Generally, learning effect is represented by equation (1).

\[ A_c = t_1 x^{-n} \]  

\( A_c \): cumulative average days (day) 
\( t_1 \): number of days for the first repetition (2nd floor) 
\( x \): numbers of repetition (floor) 
\( n \): learning coefficient (slope of a line when plot on the logarithm graph)

Learning rate is a decreasing rate of time or man-hours. It represents decreased rate of time when repetition becomes doubled (from the initial value). The learning rate is represented by equation (2).

\[ P = (1/2)^n \]  

\( P \): learning rate 
\( n \): learning coefficient

According to past researches, the learning rate for building works are reported as almost 0.75 – 0.95. Figure 6 shows the learning curves for each floor construction days. As shown in the graph both calendar days and the clear days for each floor decrease floor by floor. The learning rates are calculated by using equation (2) as follows;

The learning rate for calendar days: 0.71, 
The learning rate for clear days: 0.86.
We can recognize a certain pattern in the work progress shown in Figure 4. As seen in the graph the work begins with marking in each floor, then panel assembling and installation, and door frame, counter, etc continue. There are some idle days after these activities. During these days the carpenters did their works at lower or upper floor; meanwhile, electricians did electric wiring, and plumbers did water supply piping in walls and ceilings. After these works have been done carpenter carried out the latter half of their activities.

The primary purpose of the panel system is to reduce this idle time between activities by introducing multi-skilled workers. However, multi-skilled workers were not introduced, thus the remaining works were done in the conventional way. In this project the idle days were filled in that carpenters did their work at other floors.

Figure 7 shows the process model of these series of activities. The carpenter’s work schedule on a specific location is often interrupted by the need for other preceding trades to be completed his work scope before he can complete.

4.2 Man-hours analysis

4.2.1 Man-hour for each activity

The average value of total man-hours for each floor is 10.53 man-day/dwelling. The mean values of each activity from the 2nd to the 11th floor are shown in Figure 8. The largest man-hours activity is gypsum board and its value is 2.35 man-day/dwelling. The next largest is backing frame, with a value of 1.58 man-day/dwelling, followed by panel assemble and install, with 1.56 man-hours/dwelling.

The main reasons of large man-hours for gypsum board are as follows.

(1) As the activity was carried out after panels are installed, that there were not enough space to treat large size gypsum boards.

(2) There were many on-site processing works.

4.2.2 Change of man-hours

Man-hours data collected by the activity recording sheet is summed up and calculated to obtain learning curve effect data. Figure 9 shows the change of total man-hours for each floor. The “individual man-hours” in the graph means the man-hours for each floor, and “cumulative average man-hours” means cumulative means for each floor from the 2nd to the 11th. As shown in the graph the man-hours for each floor decreased according to increase of floor. The cumulative average man-hours are noticeable in man-hours decrease. It is considered a result of learning effect. Generally, the cumulative average man-hours are suitable for calculating the learning curve effect. The actual man-hours change was explained in 4.2.1, thus it is omitted in this section. As mentioned above, there is a learning effect in total man-hours for each floor.
4.2.3 Learning effect of total man-hours for each floor

Figure 10 is the graph showing learning curve of total man-hours for each floor for 2nd – 11th floor.

Figure 10. Learning curve for total man-hours per floor

In the graph both cumulative average man-hours and individual man-hours are displayed. Watching the change in the cumulative average man-hours in the graph, the learning effect can be implied. The learning rate is calculated as 0.86.

Theoretically, the learning effect should continue without limit. However, the man-hours reduction converged to a limit, stabilizing at almost 10 times repetitions. Some part studies’ assessment on learning curve effect was derived from 10 times repetition of the same task [4]. As calculated above the learning effect for the panel construction is 0.86 learning rate, which means that the cumulative average man-hours for second repetition is 86 percent to the initial man-hours. As stated in section 4.1.3, previous studies on building construction works shows that the learning rate of many of the construction works are 0.75 – 0.95. Table 3 shows the learning rate for building works publicized by United Nations Committee on Housing, Building, and Planning [5]. The panel system is a finishing work, and it is comparable to category III in Table 3. Therefore, the learning rate is around 0.85, calculated value is 0.86 which is almost equivalent to the value in Table 3.

Table 3. Learning rate for building construction

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Learning rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Entire structure of ordinary complexity such as high-rise office building and tract housing</td>
<td>0.95</td>
</tr>
<tr>
<td>II</td>
<td>Individual construction elements requiring many operations to complete such as carpentry, electrical work, plumbing, erection and fastening of structural units, concreting</td>
<td>0.90</td>
</tr>
<tr>
<td>III</td>
<td>Individual construction elements requiring few operations to complete such as masonry, floor and ceiling tile, painting</td>
<td>0.85</td>
</tr>
<tr>
<td>IV</td>
<td>Construction elements requiring new operations and on assembly-line such as field fabrication of trusses, formwork panels and bar bending.</td>
<td>0.80</td>
</tr>
<tr>
<td>V</td>
<td>Plant manufacture of building elements such as doors, windows, kitchen cabinets, and prefabricated concrete panels.</td>
<td>0.90-0.95</td>
</tr>
</tbody>
</table>

4.2.4 Learning effect of each activity

Having confirmed the learning effect for total man-hours of each floor in the previous section, the learning effect for each activity is calculated here. However, the data was recorded by a worker who actually worked and at the same time recorded data at intervals of half a day. Therefore, highly accurate man-hours data cannot be expected for further analysis. Hence we analyzed only to confirm whether each activity was under the learning effect.

In the analysis activities such as “miscellaneous”, “rework” were eliminated because these activities are not recurrent. Figure 11 shows the relationship between floor (repetition) and cumulative average man-hours for each activity. Decrement of cumulative average man-hours can be distinguished from this graph. They are backing frame, gypsum board, curtain box, and counter. However, it is difficult to distinguish if an activity is under the learning effect if it has man-hours less than 1.00 man-days.
For backing frame, gypsum board, the man-hours curve show typical learning curve. Conversely, man-hours are decreased in such activities as panel assemble and install, and door frame. Furthermore, man-hours of panel assemble and install increased rapidly as 0.75 man-days/dwelling at the 2nd floor to 3.00 man-days/dwelling at the 3rd floor. As a result, the learning curve has a convex shape. We assumed that some data was missing in activity “panel assemble and install” at the 2nd - 3rd floor. In Figure 11, learning curves for small man-hours such as runner, washstand, and glass wool are difficult to distinguish.

From the graph, we observe that runner and glass wool activities show the learning effect. It is difficult to confirm the learning effect for counter because its man-hours are too small to distinguish. Thus, we plotted it in a graph of logarithmically scaled vertical axis in Figure 12. A gradual decrease is observed from the 2nd to the 4th floor, while an increase is observed at the 5th floor. Due to this large fluctuation of man-hours, we consider that counter did not have the learning effect. Obviously, man-hours of washstand increased by repetition. The possible causes are mistakes in record or man-hour increased in actual fact.

Based on the analysis, the learning rates for each activity are calculated and shown in Table 4. In the table, learning rates are not shown where the learning effect is not observed. Although learning effect is confirmed on total man-hours, some activities did not have learning effect. Figure 11 and 12 shows that the man-hours of each activity fluctuated differently. The learning effect should appear in total man-hours when large man-hours activity has learning effect. However, this is not the case here.

5. Summary

The authors have developed the panel system to integrate interior finishing and building services works which was applied to a Multi-dwelling complex. Actual application was limited to only carpentry work, excluding building services works such as electricity and plumbing. Through the application we obtained construction system improvement data for future projects. At the same time man-hours data was recorded of the interior finishing work. The summary of the study is as follows.
(1) Construction time for each floor reduced according to increase in floor which shows learning effect. The learning rate is 0.86.

(2) Total man-hours for each floor also show learning curve effect and its learning rate is 0.86. The result of the observation is in good agreement with past studies.

(3) Some activities show learning effect, but some do not. The learning rates were calculated for each activity in which activities learning effects were confirmed.

(4) Average man-hours for each activity are calculated. Man-hours of “gypsum board”, “backing frame”, and “panel assemble and install” are large.

Collected data and information were used to improve the system and also stored for future projects.

Acknowledgement

We would like to thank the people who contributed to the study. We would deeply thank all the Shimizu Corporation people who assisted with the construction site. We appreciate the efforts of the students in Mine laboratory in The University of Kitakyushu who processed and analyzed data. Finally we would like to thank deeply Mr. Tadashi Okamura who recorded activity data during his work.

References


A Knowledge-Based Framework for Quantity Takeoff and Cost Estimation in the AEC Industry Using BIM

S. Arama, C. Eastmana and R. Sacksb

aCollege of Architecture, Georgia Institute of Technology, USA
bCivil and Environmental Engineering, Technion - Israel Institute of Technology, Israel
E-mail: shiva_aram@gatech.edu, charles.eastman@coa.gatech.edu, cvsacks@techunix.technion.ac.il

Abstract

An important set of information provided through Building Information Modeling (BIM) platforms are quantitative properties of design elements and assemblies. The capability to extract or deduce such quantitative properties from explicit and implicit model information is essential for bidding, procurement, production planning, and cost control activities in the AEC projects. Current solutions for quantity take off (QTO) and cost estimation (CE) are developed based on the assumptions that the design models are suitable, contain adequate information to perform these tasks efficiently and accurately. In practice often these criteria do not exist in the models that cost estimators receive. Many estimators, engineers and managers distrust BIM operations as a result or find it difficult to adopt a BIM-based preconstruction process. This leads to a cumbersome, manual and error-prone QT and CE process currently used by most construction companies. In order to overcome these shortcomings, we have developed a framework for a knowledge-based system to perform model based QTO and CE. This framework includes domain, reasoning, task and interface layers. This paper reports on the progress on an ongoing research effort which so far mostly focused on developing a domain layer and rule libraries for the reasoning layer. The domain layer contains a knowledge base which along with rule libraries were developed by acquiring and representing domain experts’ knowledge. The rule libraries include modules of rules to infer knowledge about different product features. The inferred knowledge will enable providing and representing model information in a compatible format for QTO and CE tasks. It facilitates filtering, grouping and representing feature information provided in design models based on criteria that determines their true cost behavior. Finally, this knowledge will enable forecasting the properties of product features absent from design models. Examples are drawn from various fields inside and outside of the AEC industry, with a focus on the precast concrete industry.

Keywords
Knowledge based systems, knowledge inference, quantity take off, cost estimation, precast concrete

1 Introduction

Efficient and accurate quantity take off (QTO) and cost estimation (CE) are pivotal to a project’s success. They are knowledge-intensive [1]; they are the prerequisites to many other activities in a project from budgeting, bidding and contracting to value based design, production planning and budget control; they require extracting information based on the knowledge of domain experts about the rules and processes throughout the products and projects lifecycle. There are commercial software products available that attempt to semi-automate these tasks through augmenting the quantitative information elicited from design models, creating pre-structured yet customizable cost databases and reducing repetitive aspects of these tasks [2].

Based on our study, QTO software products need to maintain three conditions for their successful performance (i) architectural and structural design models to be readily suitable for quantity takeoff and cost estimation; (ii) all the needed information to be quantitative in nature; (iii) designers’ models to contain complete information needed for these tasks. In practice these conditions are rarely met. The focus here is not on users’ modeling practices and their use of correct
modeling methods. Yet even when designs are correctly modeled:

1. Categories of contained information in models developed by designers and constructors and the way the information items are modeled and represented are different, as these models serve different purposes. Two examples are Cast-In-Place (CIP) and precast reinforced concrete products where the units of quantity take off and cost estimation are each concrete placement breaks and a product piece, respectively. However, these units often are not distinguished in models. This difference leads to rework and often for QTO and CE purposes, different construction parties have to create their own models from scratch.

2. The main focus of these solutions are eliciting and enhancing a set of standard quantities like volume, surface area, etc. for different products. The problem is that (a) each product type needs elicitation of a specific set of design properties for QTO and CE which can only be determined based on that product’s supply chain, (b) sometimes the properties that impact cost of a product are not inherently quantitative. Current systems either don’t elicit information about these properties from design models or they are represented as raw data and can’t provide the user with the insight needed for decision-making. An example is product shape. Different shape parameters that impact the cost and in what value ranges their cost relationships and behavior change should be identified.

3. The detailed design with complete information for rigorous cost estimation are developed late in the project lifecycle and usually for fabrication and production of products. For instance due to high time and cost required, many features of reinforced concrete products like connections that are important for accurate cost estimation of reinforced concrete products are often designed and modeled after the companies are contractually bound to the project. Currently QTO and CE experts mostly rely on their judgment and rules of thumb which are developed based on historical data, information and knowledge to generate knowledge. Unlike traditional information systems they can act as decision makers and serve like an expert on demand [3, 4].

These issues create considerable technical drawbacks for efficient and accurate model-based quantity takeoff and cost estimation. Our studies have shown that currently the QTO and CE processes employed by most construction subcontractors, where a detailed QTO and CE is required, is generally manual and a majority of companies only use 2D drawings rather than 3D parametric models.

In an attempt to overcome these limitations, we outlined development of a framework for a Knowledge-Based System (KBS) to identify, define and retrieve the minimum set of model information required for quantity takeoff and cost estimation of building systems. The example building system that we have selected to implement a proof of concept is precast concrete. However, the developed methodology and structure of this framework have been defined to address broader applications and is adaptable to other building systems.

This framework is designed in a way that it addresses the three above mentioned shortcomings. We study, identify and embed the rules to provide and represent information in BIM platforms in a compatible form for QTO and CE purposes. The specific set of design features and their properties, both qualitative and quantitative, that impact the cost of a project are identified. The criteria to categorize and represent these features in groups are defined, based on parameters and their value ranges where their cost relationships change. Knowledge of domain experts is elicited, and will be codified and embedded to forecast the properties of design features required for QTO and CE tasks but absent from design models (e.g. connections) with acceptable accuracy. The complete method will provide estimators with the design-related information required to perform a model-based cost estimation in an efficient and semi-automated way.

2 Research on Knowledge-Based Systems for QTO and CE

Knowledge-Based Systems (KBS) have emerged from the Artificial Intelligence (AI) field and are employed for numerous purposes in various industries. KBS are systems that acquire, represent and process data, information and knowledge to generate knowledge. Unlike traditional information systems they can act as decision makers and serve like an expert on demand [3, 4].

Several research efforts [5, 6, 7, 8] have developed knowledge-based systems for product and project cost estimation purposes. Some of these systems were developed both as a decision-making support system for choosing the manufacturing process, machines and material of products and as a cost estimation solution based on the selected options. For example, Chan & Lewis [6] developed a knowledge-based system incorporating product design, process and cost knowledge into inference engines used for material and process selection and ultimately for cost estimation.

An example in the manufacturing industry is the system developed by Shehab & Abdalla [5] for modeling cost of machining components as well as molded components. The system’s inputs include a material, a mold and a processing database as well as geometric and feature data of the product design model. Domain knowledge was represented in an expert system
toolkit through frames and rules like material selection rules and manufacturing process and tool selection rules based on various characteristics such as material cost, product functionality and machine availability. Based on the system’s recommended process, the product’s manufacturing cost was estimated. While some product features like number of cavities and surface finish were factored in the estimated cost, it is not clear how qualitative aspects like shape complexity were contributed to the cost model.

A diverse team sponsored by the National Institute of Standards and Technology sponsored Advanced Technology Program (NIST ATP) developed the Federated Intelligent Product EnviRonment (FIPER) [9] knowledge-driven environment for concurrent engineering to reduce cost of product development. In FIPER product cost information is integrated with the knowledge base. Koonce et al. [10] developed a cost with the goal of providing an integrated web-based estimation tool in which they used the design data provided by FIPER at different stages of design completion. They integrated the design data with a cost engine consisted of Work Breakdown Structure (WBS) elements and element attributes that determine the cost of an element using a hierarchical structure for attribute inheritance.

Knowledge-based systems have been developed for various purposes for the Architecture, Engineering and Construction (AEC) industry as well. For the cost estimation domain, Staub-French et al. [1] proposed a reasoning process based on cost estimators’ knowledge to represent and apply their rationale about impact of design features on cost estimation. This process customizes the activities and allocation of resources to each activity to account for project-specific features. Lee et al. [11] developed a framework that uses an ontology designed for work conditions and work items in tiling and through reasoning rules automatically selects the most appropriate work item. The inference process is designed based on knowledge of an expert and the selected work items are then used for cost estimation. In both of these efforts the focus has been on developing an ontology to represent different design and construction conditions that affect the cost of a project.

The reviewed KBSs all assume that product models used for cost estimation include all the information about feature properties that impact projects’ cost and that the unit of products represented in product design models fit the cost units of manufacturers. In other words, they only extract information represented explicitly in design models, but cannot modify the design to reflect the fabrication and installation units critical for cost estimation. They do not anticipate product features missing from design until very late stages of a project nor attempt to enhance the information retrieved from design models to contribute to a project’ cost estimation.

Figure 1. Knowledge based systems structure
These systems would only work under ideal situations when late project information is available early in the project for design entities and is represented in design models, which is relatively rarely the case.

Many of the principles used in the previously developed KBSSs can be used in the AEC industry and we use them in our framework. The proposed KBS in this paper aims to build on those frameworks and modify and improve them to depict real work environments. This is achieved by designing a framework to adjust design models and make them suitable for cost estimation without the need for redoing the model. The key extension is to infer the knowledge critical for accurate cost estimation about missing design features. Thereby the proposed system attempts to enhance the knowledge extracted from design models and to automate the current mostly manual and time-consuming QTO and CE process.

3 Knowledge Based Systems Architecture

The two major components of a KBS architecture include a knowledge base and a reasoning engine [3]. Some researchers have also included a task [12, 13] and a user interface layer [14, 15] as essential and separate components of a KBS structure. Figure 1 illustrates structure of a knowledge-based system.

Domain Layer: Domain layer consists of a knowledge base which is a repository that represents the knowledge acquired from various domains and represented using different representation tools. Knowledge acquisition and representation deal with content and format of knowledge respectively and enhance availability and usability of knowledge [14]. Various textual, graphical and computer-interpretable knowledge representation conventions and tools have been developed to standardize knowledge modeling in different domains. Examples include UML and family of IDEF languages [16].

A knowledge base represents the acquired domain knowledge using an ontology. Ontologies, originally defined by Gruber [17] as “explicit specification of a conceptualization”, are fundamental for sharing and reusing knowledge. An ontology specifies a vocabulary - set of representable objects, their properties and relationships – for a universe of discourse. KBSSs model their domain of interest through explicit abstraction hierarchies and rules about their relations that comprise an ontology. Shared ontologies tie modules of a KBS and are essential for communication and reuse of knowledge among different modules of one knowledge base and for integrating knowledge base of separate KBSSs [13].

Reasoning Layer: The reasoning layer includes modules of rule libraries and inference engines. Reasoning processes in this layer are outlined by utilizing the concept of a Problem-Solving Method (PSM) which specifies the logics behind the reasoning processes. A PSM determines required inference actions, their dependencies and sequence as well as role of each acquired knowledge piece, namely observables, abstract observables, solution abstractions and solutions to reach a specific goal [12]. Notion of a shared ontology facilitates implementation of a modularized structure for the reasoning layer where different modules computationally work as an integrated whole.

Task Layer: While hierarchy and relations of tasks are defined in the reasoning layer, a finer decomposing of tasks to the goal, required input, expected output and the strategy applied to generate the output is provided in the task layer [18]. Decomposing a KBS in this way allows having several hierarchies of tasks where tasks can be mixed and matched and different task compositions can be built to solve various problems.

Interface Layer: User interface systems enable interactions of KBSSs with users [14]. For efficient communication, these interactions should consist of two main aspects of (a) receiving inputs from users that outline users’ organization preferences, limitations or requirements. These inputs are used during the reasoning process to refine problem-solving strategies and achieve a dynamic and customized solution based on users’ needs; (b) representing the outputs of reasoning and task layer based on users’ criteria for selecting, filtering and grouping outputs.

4 Designed Knowledge-Based System Framework for Quantity Take off and Cost Estimation

4.1 Framework Overview

We have developed a KBS framework to provide a streamlined, 3D parametric model based quantity takeoff and cost estimation for construction products. This framework is represented in Figure 2 and includes the 4 layers of domain, reasoning, task and interface, designed for the precast concrete products which comprises the area chosen to implement a proof of concept for this research effort. This is an ongoing effort and so far the focus has been on developing a knowledge base and rule libraries. Several precast companies have collaborated and provided their company standards, practice manuals and their historical project cost estimation information. The principal researcher of this effort co-located for a few weeks with company experts to collect information.
Figure 2. Developed framework for knowledge-based quantity takeoff and cost estimation
from estimators, structural engineers, plant managers and erectors; to observe their QTO and CE process; and to formulate the inference rules with the help of these experts. The knowledge base and reasoning rules are being developed both for architectural and structural precast concrete products.

The method involves the following steps:

- Studied different cost estimation conventions practiced in the precast concrete industry. Analyzed performance of different cost estimation methods and documented the results of the study in [19].
- Devised a combined feature- and function-based analytical cost estimation method [19] as the most suitable for the intended estimation level of detail and accuracy.
- Decomposed precast concrete products into their functional components and identified features required for each function.
- Developed a process map for quantity takeoff and cost estimation for each function and feature.
- Identified cost-driving attributes of each feature and specified the parameters required to measure the impact of each attribute on cost of a project. These variables comprise the information items necessary for precast concrete cost estimation.
- For the information items typically implicit design models or absent from models, defined the rules to infer knowledge about them and created a rule library for each function.

Expected results of implementation include enhancement of the design models to make them suitable for QTO and CE activities, extraction and representation of model information using the industry convention format and measurement units, and forecasting parameters of absent features by inferring new knowledge.

4.2 Domain Layer: Knowledge Base

The domains studied in order to develop an example knowledge base that guided the listed steps from product decomposition to process mapping and rule development included architectural and structural design, and supply chain analysis (fabrication, transportation and erection) of precast concrete products. The focus of knowledge acquisition was on those domain aspects that are interdependent with quantity takeoff and cost.

4.2.1 Cost estimation methods

Cost estimation methods can be categorized as intuitive and analogical methods used in early design stage CE and analytical methods for late design stages. A detailed study of these CE methods used in different project stages and analysis of the performance and shortcomings of each method is published by authors in a separate paper [19].

CE methods used in early stages of a project mostly can work only with a limited number of variables and provide a rough approximation of cost of a project suitable for budgeting. Considering this, they are not suitable for a more detailed CE process when there are more design information available and for instance geometry of building and different spaces within a building, type of building structure and location of structural elements are determined. Hence, for this research work, we use analytical CE method. Analytical methods decompose a product using operation-, tolerance-, feature-, and activity-based modeling.

We used a combination of activity- and feature-based product decomposition. We studied variety of design features that compose a specific product type, the supply chain process and activities that are required to manufacture each feature, and identified design variables that affect cost of each activity and therefore are important to be provided for cost estimators.

The main goal of the reviewed CE methods has been defining relationships of different design variables to cost of a project using historical data and applying various machine learning and optimization methods [20]. The focus in building the knowledge base of this framework is not to define cost relationships, rather to identify existence of those relationships between different variables and cost of a project and providing value of these variables to users, when they are not readily available in design models, through building a rule library and a rule processing engine. When the value of different variables are determined and provided to users they can then plug them in their formulas that are built based on their production process and local economic conditions.

4.3 Reasoning Layer: Rule Library and Inference Engine

As shown in Figure 2, we structured the reasoning layer by developing specific-purpose modularized rule libraries for various functions (e.g. connections, reinforcement, finishing, etc.) of different precast concrete product types (e.g. columns, beams, slabs, etc.). Rule libraries are being developed using different inference mechanisms to infer new knowledge for QTO and CE of different aspects of a product. These rules will be applied on the information extracted from 3D parametric design models as well as user inputs.
regarding company limitations and preferences. We will use a combination of generic inference tools many times found as off-the-shelf inference shells and specific purpose reasoning modules developed for domain applications.

These modules represent the rules and reasoning for three major purposes:

(i) Enhancing models using combination of implicit and explicit information to make them suitable for QTO and CE activities without the need to create new models from scratch: As explained earlier, while the unit of QTO and CE for precast concrete products is a precast concrete piece, in the architectural design models often pieces of precast concrete products are not distinguished. Through extracting the geometric and spatial relation information of products from models and applying panelization (modularization) rules developed based on the acquired domain knowledge, precast concrete model objects can be panelized to represent acceptable approximations of precast concrete pieces.

(ii) Extracting and representing explicit model information using the industry convention format and measurement units. This information mainly include dimension, surface area volume and weight measurements and properties like material.

(iii) Forecasting information about features absent from design models: Detailed design of many key features of precast concrete products like connections, reinforcement and form stripping and lifting inserts for the most part is performed by structural engineers who work for precast concrete contractors. The process is costly and time consuming and normally is performed after winning the bid and securing the project. During the QTO and CE activities information related to these features are mostly absent from models. Similar to the model enhancing process, relevant information for each feature is identified and extracted from design models.

Rules to infer new knowledge about these features are developed which will forecast the value of cost-driving feature parameters (e.g. number and type of reinforcement elements).

4.3 Problem Solving Methods and Knowledge Roles

Bases of the developed rules are PSMs and their knowledge roles. An example of application of a PSM and different knowledge roles is illustrated in Figure 3. In this example if the value “50’’ is extracted from a design model as an observable for “total building height”, it can be abstracted to “above max height” using another observable of “45’’ which is a user input for “max feasible column height”. This abstracted
observable will be followed by applying a solution abstraction of “divide to pieces below max length”, which will produce the solution of “column piece length”.

Note that to generate this solution, the PSM requires another set of inputs which are “panelization rules for columns”. These rules are the end results of the process of knowledge acquisition from domain experts and knowledge representation. These rules themselves comprise of a cluster of PSMs that define the actions and the rationale behind each action and their implementation might require additional inputs.

This figure also shows how different classes of inputs including data, information and knowledge which are acquired from different sources including parametric design models, users and domain experts are used to carry out a task and infer new knowledge. The nature of these inputs also cover a wide range from dynamic (e.g. design model data that is project-specific and often even changes throughout a project lifecycle) to relatively static (e.g. panelization rules based on architectural, structural and supply chain rules) that can almost be considered fixed until standards or production technology change at which point they need to be refined.

5 Conclusion

This paper presented a framework developed for a streamlined knowledge-based quantity takeoff and cost estimation of precast concrete products using 3D parametric models. The focus was on development of a knowledge base and rule libraries for extracting implicit knowledge used by experts in CE. Validation will be inferred when estimators accept and have confidence in the results of CE from BIM models The next step, after completion of the knowledge base and rule libraries, is to implement this framework for selected categories of precast concrete products on one of the major domain software platforms.

References


The Power of Technology in Bridge Construction Project Management

Fahimeh Zaeri\textsuperscript{a} and James Olabode Bamidele Rotimi\textsuperscript{b}

\textsuperscript{a,b}Construction Management Programme, Auckland University of Technology, New Zealand

fzaeri@aut.ac.nz; jrotimi@aut.ac.nz

Abstract-

Bridge construction projects are inherently complex and iterative, and these place great demands on project management to apply innovative approaches for more comprehensive analysis of performance data under uncertain conditions. Although new technological-based methods such as simulation have proven to be powerful techniques to cope with cyclic and uncertain project behaviours, implementation of simulation-based modeling is below par in the construction domain especially in bridge construction. This paper presents information on significant role of recent modeling methods in the construction domain. The study is an aspect of wider research that explores capabilities of simulation-based methods in scheduling and managing of construction projects in New Zealand by considering their repetitive, uncertain and complex features. The study design and data collection are briefly described to demonstrate the power of simulation technology in bridge construction management. It is hoped that the study will benefit both construction planners and managers.

Keywords

Simulation, Bridge construction, Resource Utilization, Uncertainty, Scheduling

1 Introduction

Construction projects are complex and are usually considered the most complex undertaking in any industry. The construction industry experiences great difficulty in coping with the increasing complexity of its major projects. For example complex interdependencies in a construction project’s components make its analyses complicated. In addition, multiple interacting feedback processes included in a complex system such as large scale construction projects means that mental models and traditional cost and scheduling tools such as CPM do not adequately account for feedback effects [1]. Project tools such as Gantt charts, PERT (Program Evaluation and Review Technique) and CPM (Critical Path Method) have been very helpful in the scheduling of activity sequences, but are unable to solve problems related to the dynamic nature of construction projects [1]. For example, CPM-based analysis only determines the effect of a change in the time required to complete an individual step which may affect the total completion time for a project. The analyst estimates the required time on the basis of historical data, past experience or judgment. The implicit assumption is that the time required to carry out all other steps is unaffected and consequently all other interactions are ignored [1].

Further, the construction industry and academia have realized that critical path methods and other time-based methods are not suitable for repetitive projects like bridge construction, high-rise buildings, housing projects, tunnels and highways which are resource-driven in nature [2].

Sripraset and Dawood [3] are of the opinion that project failures and low productivity are due to the ineffectiveness of traditional construction planning methodologies to support today’s project features. Thus, proper planning of complex project operations is important to ensure their timely and economical completion [4]. It is common knowledge that the planning phase is crucial, knowledge-intensive, ill-structured, and the most challenging phase in project development cycles. More so that construction processes have become more complicated, interactive and dynamic in nature [5].

The paper reviews extant literature on simulation in construction projects and especially in bridge construction from the point of view of project scheduling, and demonstrates the applicability of simulation on a case study project. The paper is an aspect of a larger research programme which explores the potential capabilities of simulation-based modeling methods in New Zealand...
construction projects. Contextual information on the larger research on which this paper is based is provided. The information provided includes the overarching aim of the research, its objectives, and a brief description about the research methods that is being employed to accomplish the given research objectives.

2 Research Motivations

Bridge construction works are associated with uncertainties that may become exacerbated by resource constraints, operational logistics and the work environment [6]. Inevitably progress on a bridge construction project is dictated by construction sequence, resource availability and structural adequacy [7]. Therefore according to Ailland et al. [8], bridge construction sites are commonly characterized by enormous pressures for time, cost and quality achievements within difficult logistic requirements.

Currently only a small number of researchers have applied simulation technology on bridge construction processes [9], while none have been applied in the New Zealand context. Although computer simulation has proven to be one of the powerful techniques for modeling uncertainties, its application in the construction domain is still limited [6]. This lack of uptake in the construction domain has generally been attributed to the difficulties associated with learning and applying simulation languages to its projects. Since the construction industry is characterized by individual products that vary in type, size, function, material and other attributes, modeling techniques would need to vary from one project to another one. Ailland et al. [8] also suggest that construction conditions change within proposed construction processes and shift as the processes progress. Based on these characteristics, such processes create opportunities for performing comprehensive analysis using simulation techniques. Therefore, developed simulation models would need to be unique yet adaptable to varying circumstances.

The use of simulation-based tools to investigate construction process/operation has the advantage of improved modeling of construction work sequence beyond those provided by MS Project or Primavera. Both MS Project and Primavera are merely planning and scheduling tools which are limited tools for seeking the needed improvements to production outputs within the construction industry.

3 Bridge Construction Projects

Continuous movements of workers and machines in a linear geographic layout are involved in many construction projects such as highways, tunnels and pipelines. In this class of projects, crews parade the site one after the other during the process. These are referred to as continuous repetitive projects as opposed to discrete repetitive projects, therefore work performance cannot be measured in discrete work units, but rather continuously with certain amount of distance and time interval kept between operations [10-12].

As [13] suggest, work processes in bridge construction projects are usually linearly repetitive involving the erection of a series of abutments, piers, spans, and other segments of the bridge’s superstructures Therefore, bridge construction planners need at first generate a hierarchical work breakdown structure that starts from a high-level activity, such as “Construct Bridge” and follow by lower level activities such as “Construct Abutment”, “Construct Pier”, and “Construct Superstructure”. For each level, the planner selects and applies certain construction methods for each of those activities [14]. However, construction sequence and resource availability underpin planning progress in bridge construction projects [7]. Where, unavailability of resources has been realized as a main factor for causing uncertainties in these projects [6]. Moreover, as there is no uniform repetition of a module network in continuous repetitive projects, then traditional scheduling techniques, such as CPM (critical path method) and Gantt chart, have been criticized for their inadequacy to accurately model continuous repetitive projects [15, 16](as cited in [17]).

4 Scheduling of Bridge Construction Projects

Wu et al. [18] introduced scheduling of construction projects as an allocation of resources of workers, machines and materials in a time-efficient way. Moreover, baseline schedule is proposed as an important step that the contractor has to realistically estimate the duration of a project. The importance of this step is such that project owners can evaluate the feasibility of contractors based on that and also successful completion of the project is attributed to this step [19].

The differences between design, productivity of resources, availability of resources, scheduling techniques are the effective factors to make repetitive activities unique in the point of view of durations. In other words, their durations are rarely identical in each unit since those factors contribute to activity and resource schedules, and definition of the repetitive activities’ characteristics. Consequently, characteristics of repetitive activities are what creates the need for sophisticated scheduling techniques and tools to schedule projects under precedence and resource constraints [2]. Further, construction project scheduling is almost experience-based. In other words, human knowledge plays an important role in project scheduling. In this point of view, previous researchers tried to capture human knowledge to create a powerful system to deal
with scheduling issues. Which they only have been able to represent the expertise in the form of a set of data and rules on the computer [14].

Most of the previous works associated with scheduling of repetitive construction projects focus on deterministic scheduling. In this way, uncertainty, which is an inherent feature of construction projects is ignored [20].

Further, those approaches only deal with a certain degree of complication as they were limited to deterministic problems. There are few probabilistic scheduling methods that capture the stochastic nature of construction projects which employ simulation techniques. Although they attempt to take stochastic nature into account, none of them guarantees continuous resource utilization [2].

Comparison between Critical Path Method (CPM) and Discrete Event Simulation (DES) shows that simulation can provide an unbiased distribution of construction project duration. Therefore, DES methods can predict construction project duration more realistically than CPM [21, 22].

Although, simulation has gained acceptance as a particular scheduling tool in other fields like operations research and in the manufacturing industry, but it has not been yet popular in the construction industry. Absence of a reliable method in selecting probability distributions for construction activity duration has been realized as the fundamental reasons for the limited use of simulation as a scheduling tool in construction domain [19].

Multitude of requirements such as a technological dependencies and resource capacities should be taken into account together with principal guidelines of project duration and available funds. However, the scheduling of construction projects become more complex as Wu et al. [18] (2010) found constraint-based simulation could be a more proper technique. Although they have been successful to overcome intricacies associated with scheduling construction projects by generating effective schedules virtually automatically but creating the necessary input data manually has been the major drawback of these approaches [18].

5 Resource Allocation in Bridge Construction Projects

Traditional scheduling methods such as CPM and graphical methods like Line-Of-Balance (LOB) cannot consider dynamic and resource-driven features of construction activities [2]. It seems that mathematical methods have been successful to model behaviours of shared resources through dynamic programming but in comparison to resource-driven simulation, they cannot be so effective. In addition, some of resource-driven simulations such as STROBOSCOPE use a conditioning node (e.g. Fork node) within shared resources allocation processes which offer a great advantage in the modeling of construction activities by considering their dynamic and resource-driven feature. The way of assigning shared resources between construction activities has an important role in successfully management of continuous repetitive projects [2].

CPM, LOB and Repetitive Scheduling Method (RSM) methods assume that activities require only one resource each and resource availability constraints are modeled by using precedence constraints. In these methods, a resource serves only one activity which is called a “dedicated resource” while in reality, activities may share the same resources. Where, resources and activities are called shared resources and resource-sharing activities respectively. However, precedence constraints approaches have failed to present resource availability constraints for shared resources [2].

Resources constraints are classified into 1/ resource availability constraints, and 2/ resource continuity constraints. The first category controls the output of activities for which their resources are available during a particular period while the second one includes the resources which need to work continuously and without interruption from the time they first arrive to the job site until they leave. As Srisuwanrat [2] proposed, considering this repetitive feature is indispensable to resource allocation within scheduling a construction project.

6 Simulation in Bridge Construction Projects

Kim [19] described simulation as a building and investigation process for a computerised model of a system which captures various time measures such as real time, expanded and compressed time to improve the behaviour of a process or system. Simulation is able to model any system with any set of conditions in a more practical way since it runs the computerised model of a system rather than finding analytical solution. This potential of simulation makes it more advantageous than traditional scheduling methods like CPM and PERT. In other words, the considered system does not need to be analytically managed. Moreover, fewer assumptions are required when simulation is used to schedule construction projects.

In the simulation approach, individual activities, interdependencies among them and resource availability are taken into account. This capability makes simulation suitable for detailed investigation of construction schedules [18].

Although simulations have been successful their implementations have not drawn as much attention in bridge construction processes.

Few examples of studies that have applied simulation within the bridge construction domain include works done by Ailland et al. [8], AbouRizk and Dozzi [23], [24], [25] and [26]. In their work, AbouRizk and Dozzi [23] used CYCLONE to facilitate dispute resolution in
bridge jacking operations. Huang and Halpin [24] simulated the construction operations in a cable-stayed bridge in Washington by using DISCO simulation software. Chan et al. [25] used SDESA to simulate field processes for a pre-cast bridge, resulting in optimal solutions to the pre-cast segment inventory problems.

Others like Marzouk [6] utilized simulation model like STROBOSCOPE as a simulation engine which was coded by Visual basic 6.0 to develop a special purpose simulation model to assist in the planning of bridge deck construction. This simulation engine considers uncertainties and the interaction amongst resources used for the construction works. Marzouk et al [6] had modelled the 15th May Bridge located in Cairo, Egypt which was constructed using an incremental launching technique. Marzouk et al [6] examined the results of the developed model and illustrate its capabilities in modeling two construction methods; single form, and multiple form. A sensitivity analysis was performed in their study to evaluate the performance of the system under different combination of resources. The study eventually enabled planners to estimate duration and production rate in each combination within those different methods of bridge construction and also provided them more understandable results to study the impact of assigned resources when estimating project duration.

Another research study undertaken on bridge construction by Said et al [27], reflects how simulation can facilitate construction process planning. Said et al had employed a developed STROBOSCOPE simulation engine, called ‘Bridge-Sim’ in a case study of the El-Warrak Bridge in Cairo - Egypt, to estimate the total duration of deck execution and the associated total costs. Said et al [27] suggests that Bridge-Sim also enables planners and contractors to evaluate different scenarios of construction plant utilization which represents various combinations of construction methods, crew formations, and construction sequencing. For example they compared cast-in-place on falsework method and cantilever carriage construction methods for the El-Warrak bridge. Simulating of the two construction process methods demonstrated the potent capabilities of simulation method in the creation of comprehensive documentation systems that helps planners in analysing construction alternatives where the project involves many repetitive activities, complex interdependencies between construction resources and uncertainties.

Table 1 presents information on typical construction projects where, how and why simulation was implemented on a variety of projects.

Simulation modeling methodologies vary depending on the nature of the projects to be modeled, but it could be observed from the table, that simulation is applicable to a wide spectrum of construction operations [28].

Simulation can be applied in general or for special purposes. General purpose simulation (GPS) refers to the formulation of a simulation model for a system under-investigation, running the simulation and analysing the results to decide whether the system is acceptable or not. If it is not acceptable, the process would need to be reiterated and an alternative system considered. Various GPS software systems have been developed for a wide range of industries e.g. AweSim [29] and GPSS/H [30], and for construction: Micro-CYCLONE [5] and STROBOSCOPE [31]. Special purpose simulation (SPS), on the other hand, is developed for a specific domain of application through the creation of a definite platform or template [32, 33]. The steps for simulation are the same in both GPS and SPS except the first step (involving the construction of the simulation model) where the platform includes the characteristics and behaviour of the considered system in SPS. Also, the modification is limited to the input parameter(s) of a pre-defined system and not to the characteristics and behaviour of the system [6].

It is conclusive from the foregoing that simulation is an appropriate tool to support construction planning, reducing the risks associated with time, budget, and quality and also increases productivity and operational efficiency [34].

7 Design of the Research Study

The section describes aspects of the larger study programme (doctoral research project) on which the current study is based. The study is at a relatively early stage in New Zealand. The main objective of the doctoral research is to determine how simulation methods could improve construction projects performance and productivity. The study is expected to provide some general and specific benefits for the NZ’s construction sector as outlined below:

(i) General

1. The development of an object-oriented framework for modeling construction operations toward improving projects productivity. This will contribute to current industry initiatives for productivity improvement by 20% in 2020.

2. Providing a general understanding of the features of simulation modeling at functional/operational levels of construction management as a complementary tool to planning and analysing complex construction projects.

(ii) Specific

1. Complement existing project planning methodologies currently being used at the operational level for the case study project with simulation modeling.
<table>
<thead>
<tr>
<th>Simulation engine/software</th>
<th>Simulation strategy/method</th>
<th>Benefits</th>
<th>Researcher (Year)</th>
<th>Construction method</th>
</tr>
</thead>
</table>
| Fuzzy Logic Simulation    | Stochastic method         | 1- Estimate process duration  
2- Using sensitivity analysis to validate simulation strategy | Ailland et al. (2010) | Using mobile scaffold, double webbed t-beam cross section made of pre-stressed concrete |
| CYCLONE                  | General Purpose Simulation | Facilitate dispute resolution | AbouRizk et al (1993) | |
| DISCO (dynamic interface for simulation of construction operations) | Graphical construction simulation system | 1- Provide a graphical environment in which modeling and simulation of construction operations can be conducted in an interactive fashion.  
2- Analyze system productivity as well as resource utilization by examining graphical report. | Huang, R. et al. (1994) | Segmental balanced cantilever technique to build on both sides of the pylon at the same time |
| SDESA(Simplified Discrete Event Simulation) | General Purpose Simulation | Simulate field processes to reach optimal solutions to the pre-cast segment inventory problems | Chan et al. (2005) | |
| STROBOSCOPE/Bridge-Sim Software | General Purpose Simulation | 1- Facilitate construction process planning  
2- Estimate the total duration of deck execution and associated total cost  
3- Enable planners and contractor to evaluate different scenarios of construction plant utilization | Marzouk et al. (2008,2009) | Cast-in Place on Falsework and Cantilever Carriage |
| Pattern-Based Simulation | Discrete Event based environment and Constraint-Based strategy/Preparator | 1- It consists of a sequence of work packages for realizing standard bridge parts which allows user to specify the available resources  
2- Facilitate schedule generation and cost analysis in bridge construction projects  
3- Encapsulate the repetitive construction processes  
4- Guarantee a high flexibility in modelling processes | Wu, IC et al. (2009) | |
| SimoFit | Discrete Event based environment and Constraint-Based Strategy | 1- Deliver important data for construction simulation based on existing huge data sources (data source is the base of high level simulation solutions.)  
2- The construction tasks and their constraints for production such as technological dependencies, availability and capacity can be specified and valid execution schedules can be generated.  
3- Developed simulation models to investigate the impact of different influencing parameters on the performance of construction activities. | Melzner et al. (2011) | |
2. Enhancing decision making on construction process strategy because when the model is developed, it could be used for planning and analyses of a wide range of construction processes.

3. The case study organization could adopt the simulation tool as a best practice approach to analysing and planning for the execution of their complex projects. The workshops planned at every stage of the model development could serve as capacity building workshops for operations personnel.

Information will be gathered from a case study investigation and will be complemented with semi-structured interviews and focus group studies. The case study envisaged by the research is a bridge construction project which is at an early stage of development in Auckland - New Zealand. The project involves the erection of four ramps using Self-Lunching Twin Truss Gantry construction technique.

8 Conclusion

Dealing with repetitive, uncertain, and dynamic features of construction projects within scheduling methods have been sources of concern in the construction domain. To achieve accuracy and successful management of construction projects, then planners need to implement better modeling methods. Regarding the ineffectiveness of traditional planning methods, project planners would need to seek integrated approaches using new technologies in construction management.

Considering that there have been fewer studies related to implementation of simulation in this domain the current study thus finds its usefulness. The study represents an innovative holistic study within the NZ construction sector that would deal with intricacies involved in construction projects management and help improve projects’ delivery.

References

19. Kim, J., An investigation of activity duration input modeling by duration variance ratio for simulation-
A BIM Based Construction Site Layout Planning Framework Considering Actual Travel Paths

J.C.P Cheng* and S.S. Kumar*

*Department of Civil and Environmental Engineering,
The Hong Kong University of Science and Technology, Hong Kong
E-mail: cejcheng@ust.hk, ssk@ust.hk

Abstract -
Construction site layout planning (CSLP) is recognized as a crucial component of construction management. The objective of CSLP is to determine the best arrangement of temporary facilities on a construction site, which would minimize the transportation distance of site personnel and equipment. It could be achieved by creating dynamic layout models that allow layout planners to cater the changing requirements of the site. However, these models are project specific and require large amount of manual data input by the layout planner. Besides being tedious, this approach is not practical either, since any changes to the design or construction plans would have to be manually updated into the models, resulting in unnecessary work by the layout planner.

In this study, we propose an automated framework to create dynamic site layout models leveraging the BIM technology. Using the information in BIM models and construction schedules, a dynamic layout model can be automatically created for facility layout optimization. Furthermore, the actual travel distances among facilities instead of Euclidean distances are considered in our framework when performing the facility layout optimization. A* algorithm and genetic algorithm heuristic method are used. The proposed approach and framework could reflect the actual site situation and facilitate the facility layout planning on construction sites. A case example is presented in this paper to demonstrate the framework and compare its results with those using Euclidean distances.

Keywords —
Construction site layout planning (CSLP); Building information modelling (BIM); Genetic algorithm (GA); Non-Euclidean distance

1 Introduction

The objective of construction site layout planning (CSLP) is to determine the optimal layout of temporary facilities (such as storage areas, fabrication shops, machines, residence facilities and equipment) within the boundaries of a construction site in order to enable the safe and efficient movement of materials, equipment and labour [1]. The CSLP problem can be subdivided into three parts – (1) determining the required size of facilities, (2) identifying at which stage of construction each facility is required, and (3) allocating facilities to different site locations. Site layout models fall under two broad categories – static models and dynamic models. Static models assume a fixed layout for temporary facilities through the entire duration of a project, whereas dynamic models attempt to capture the changing facility requirements during different stages of construction [2]. Dynamic models are better than static models in addressing the requirements of a construction site and recent research mainly focuses on dynamic layout planning [3]. However, modelling the dynamic facility requirements of a construction site is a complicated task that takes a significant amount of time and effort by the layout planner. Current site layout planning tools require a large number of project specific variables to be inputted manually by the layout planner. This method is cumbersome, especially since any changes to the original construction plan would require all of the updated project variables to be inputted again.

In this study we demonstrate how construction site layout planning can be automated using building information modeling (BIM) technology. Although BIM has been used in the construction industry for over a decade, its use in construction planning is still limited to clash detection and 4D simulation [4]. However, BIM models are rich information sources and hence can be used for site layout planning as well. Information stored in BIM models can be used as a basis for estimating the size, type and number of temporary facilities required by a construction project over time. This paper will demonstrate how BIM can be used in conjunction with project schedules to facilitate the creation of dynamic layout models for CSLP. Furthermore, since this methodology pivots on BIM, design and construction changes will automatically be reflected in the layout models, significantly reducing unnecessary work by the
This paper presents a CSLP system framework based on BIM that we developed with the consideration of non-Euclidean distance among facilities on construction sites.

An important step of the CSLP process is optimizing the layout of facilities by allocating spaces to facilities so as to ensure efficient site operations. This problem is considered to be ‘NP-hard’ and several research studies have attempted to arrive at solutions using heuristics [5] and mathematical optimization techniques [6]. In most layout optimization approaches, the goal is to minimize the total inter-facility transportation cost, which can be expressed as:

\[ \text{Min} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{ij}f_{ij}R_{ij} \]  

where \( n \) is the total number of facilities, \( f_{ij} \) is the frequency of transportation between facilities \( i \) and \( j \), \( R_{ij} \) is the unit cost of transportation between facilities \( i \) and \( j \), and \( d_{ij} \) represents the distance travelled by site personnel between facilities \( i \) and \( j \). Most of the previous studies take the parameter \( d_{ij} \) to be the Euclidean or straight line distance. However, on a construction site, due to the presence of obstacles and other restrictions, it is impossible for site personnel to always follow straight line paths. Instead, they travel from one point to the other avoiding any obstacles that may come in their way. Sanad et al. [7] claimed that using linear distances, such as the Euclidean distances, leads to an under-estimation of the site layout problem and hence introduced the concept of using actual travel paths instead. Yahya and Saka [8] introduced the concept of obstruction distance which was added to the straight line distance to account for site obstacles. However, neither of these studies laid much emphasis on accurately modelling the paths followed by site personnel. In this study, we used the A* algorithm in order to realistically model the travel patterns of personnel on a construction site. By doing so we could develop a more accurate estimate of their travel distances thereby improving the reliability of the optimization model. The superiority of the actual travel path driven optimization is demonstrated on an example project.

2 The BIM Based Construction Site Layout Planning Framework

In CSLP, construction facilities can be categorized as either fixed facilities or temporary facilities. Fixed facilities, such as buildings under construction, are assumed to be immovable and their locations are fixed prior to construction. Temporary facilities, on the other hand, refer to site objects that do not have fixed locations prior to construction and need to be allotted space on the construction site through careful planning. Temporary facilities may refer to laydown areas, fabrication shops, site offices, batch plants, etc. Since the site layout represents the work conditions of the labour crew for the entire duration of construction, thorough consideration must be put into deciding the site layout. Proper layout planning ensures good working conditions and hence improves the morale and hence productivity of the entire labour force. The CSLP process can be subdivided into three major tasks: (1) estimating size requirements of facilities, (2) identifying at which stage of construction each facility is required, and (3) allocating each facility a position on the construction site. The following subsections will describe these three tasks and how BIM can be leveraged to perform these tasks.

2.1 Facility Sizing

Facilities should be sized so as to facilitate (1) best practices for material storage, (2) safe work conditions for labourers, and (3) efficient functioning of the facility. Material storage facilities should ensure that there is adequate space for storage throughout the duration of construction. Facilities such as fabrication shops, carpentry yards and site offices should provide an unhindered working environment. Production facilities such as batching plants must be sized keeping in mind their production capacities. Therefore, the size of temporary facilities depend upon variables such as the estimated quantity of work, rate of consumption of resources, number of workers and site area. Using information gleaned from prior projects, layout planners generally develop thumb rules or parametric

![Figure 1. Calculating facility sizes](image)
relationships to estimate facility size requirements. Based on interviews with practitioners, we developed a set of relationships between facility sizes and their dependent variables, as shown in Table 1. In general, the facility sizes depend on the peak consumption rates of resources.

Table 1. Parametric relationships for facility sizes

<table>
<thead>
<tr>
<th>Facility</th>
<th>Dependent Variable</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebar Storage</td>
<td>Peak consumption / day</td>
<td>10 m²/ton</td>
</tr>
<tr>
<td>Rebar Storage</td>
<td>Peak consumption/day</td>
<td>10 m²/ton</td>
</tr>
<tr>
<td>Bending</td>
<td>Peak consumption/day</td>
<td>5 m²/box</td>
</tr>
<tr>
<td>Storage Tiles</td>
<td>Peak consumption/day</td>
<td>2 m²/bag</td>
</tr>
<tr>
<td>Storage Cement</td>
<td>Peak consumption/day</td>
<td></td>
</tr>
<tr>
<td>Engineer’s Caravan</td>
<td>Total engineering crew</td>
<td>15 m²/person</td>
</tr>
</tbody>
</table>

The process of facility sizing could be automated by adopting the following steps, as shown in Figure 1:

- **Step 1**: Calculating the total amount of resources required by each activity. By performing quantity take-offs on the BIM model, we were able to obtain detailed information pertaining to the quantity of resources consumed by each activity. Thus, for example, the total volume of rebar and concrete needed for an activity such as “Level 1 Columns” could be derived directly from the BIM model.

- **Step 2**: Calculating the durations of activities. This information could be read from the planned construction schedule. For example, the duration of activity “Level 1 Columns” would refer to the number of days assigned to this activity in the project schedule.

- **Step 3**: Determining the peak rate of consumption. For each activity, the rate of consumption was calculated by dividing the quantity of resources consumed by the duration of the activity. The peak consumption rate could then be obtained as the maximum rate of consumption among all activities.

\[
\text{Peak rate of consumption} = \max \left( \frac{\text{quantity consumed}}{\text{duration of consumption}} \right)
\]  

- **Step 4**: Applying the parametric relationship to yield the required facility size. For example, if the peak rate of rebar consumption is 7.85 ton/day, the required size of the rebar storage yard could be calculated as 78.5 m² by multiplying the peak rate of consumption by the parametric relationships. The size of the engineer’s caravan could be computed assuming six engineers. In this work, Autodesk Revit was used as the BIM tool and Microsoft Project was used for scheduling. The planned construction schedule as well as the resource information from the BIM model were exported into Excel spreadsheets. We then created a program to read these files, perform the necessary calculations and hence obtain the required facility sizes (see Figure 1). Changes to the design and schedule could be updated into the Excel spreadsheets, enabling them to be reflected in the final calculations. Therefore, our framework allows a quick and easy method for facility sizing and does not require the users to manually input project specific information. This framework also achieves a significant reduction in effort when coping with changes to design and construction plans.

### 2.2 Dynamic Layout

Not all facilities are required throughout the whole project from start to finish. Many facilities are required on site only for limited durations after which they are dismantled. The space which they occupied then becomes available for setting up other facilities. On sites with limited available space, multiple facilities might occupy the same position on the site during different stages of construction. In order to model these changing space requirements, the CSLP process must be treated as a dynamic layout optimization problem. Dynamic layout refers to the sequence of layouts spread over distinct time intervals, which when taken together span the entire duration of construction [9]. In other words, the construction project is split into multiple phases and layouts for each phase have to be determined individually. Tommelein et al [9] developed a dynamic layout tool called MovePlan which took activity relationships as input in order to determine the dynamic facility requirements of a site. In this study we have developed a methodology of determining the dynamic layout requirements from the planned construction schedule. This was done by mapping each activity on the schedule to its corresponding facility requirements, as depicted in Figure 2.

For instance, in order to determine during which time interval of construction a rebar storage yard would be needed, we would first need to identify (1) the activities which require a rebar storage yard, and (2) their start and end dates. Using the logic that a facility should be present on site only as long as there is an activity which requires it, we could correlate the requirement of the rebar storage yard to the durations of related activities.

Most commercial software allows for schedules to be exported as csv files, enabling us to create a program to read activities and assign them to phases. As a result, the process of dynamic modelling could be automated.
2.3 Objective Function of the Facility Layout Optimization Problem

After identifying the size and number of facilities required in each phase of construction, we need to formulate the CSLP problem into an optimization problem. On a construction site, personnel travel from one facility to the other in order to perform activities or transport materials. Every such movement is assumed to incur a cost which is directly proportional to the distance travelled. The objective of layout planning is to minimize the total inter-facility transportation cost of materials and labour while adhering to a specified level of safety. This is aligned with the principles of lean construction, which recommend minimizing unnecessary material handling and transportation costs. The objective function was mathematically defined as:

$$\text{Min } \sum_{p=1}^{N} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} e_{ijp}P_{ijp}$$

s.t.

$$\sum_{p=1}^{N} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} e_{ijp}P_{ijp} \leq S_{ijp}$$

$$e_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

where $N$ refers to the number of phases, $n$ refers to the total number of facilities, $d_{ijp}$ refers to the actual travel distance between facilities $i$ and $j$ during phase $p$ considering the presence of site obstructions, which will be explained in the next section. $R_{ijp}$ and $f_{ijp}$ refer to the cost and frequency of transportation between facilities $i$ and $j$ during phase $p$. $P_{ijp}$ is a parameter which denotes the presence of facilities $i$ and $j$ during phase $p$. Eq. (3) represents the total distance travelled by site personnel during construction. Having facilities too close to one another creates a safety problem, especially when heavy materials are being lifted by cranes. This is expressed in Eq. (4), where $S_{ijp}$ denotes the trade-off between safety and closeness between two facilities.

The layout corresponding to the minimum total travel distance is assumed to represent the best possible layout.

2.4 Actual Travel Distance

A drawback of most studies on this topic is that they used linear distances in facility layout optimization. Park et al. [10] demonstrated the superiority of using actual travel paths over linear distances in solving the floor-level material layout problem for an indoor environment. In this study we have used the actual travel distances of site personnel instead of unrealistic linear distances. It is a known fact that on a construction site, personnel always seek the shortest path from source to destination in a bid to minimize their effort in transportation. Hence, the actual travel path is assumed to be the shortest path from one facility to another, considering the presence of site obstacles.

To determine the actual travel path, we converted the site into a grid, mapped the locations of obstacles on it, and then used the A* algorithm to compute the shortest path between two points, as illustrated in Figure 3. In order to ease the computation effort we limited the path to move in eight independent directions. This leads to a slight deviation in the actual travel path of personnel and the path generated by the A* algorithm. However, this does not take away from the fact that the distance calculated by the A* algorithm is more accurate than the Euclidean distance.

Figure 3. Linear and Actual Travel Paths
Another characteristic of construction sites is the width of the path required by personnel and equipment. Lighter materials may be carried by labourers whereas heavier materials require wheelbarrows or forklifts for transportation. The width of the path required is different in each case and is related to the size of the equipment. In order to ensure that different locations are accessible by both labour and equipment, it is essential to select layouts which provide safe and navigable paths. In our model we assumed the path width as 1m for labourers, 2m for wheelbarrows and 5m for larger equipment. This allows a more accurate representation of site activities.

2.5 Constraints of the Facility Layout Optimization Problem

2.5.1 Available Site Space

BIM encompasses all of the functionalities of CAD, hence by extracting the site layout plans of the construction project, we were able to automatically generate available workspaces. This was done by extracting data contained in the layout plans and formulating them into a set of mathematical inequalities. The closed area bounded by the inequalities represented the total site area available for construction. The fixed facilities (buildings) were then identified in a similar manner and subtracted from the total area encompassed by the site. The remaining area represents the total space allocated for setting up of temporary facilities.

2.5.2 Overlapping Constraint

We used a mathematical inequality to ensure that facilities do not overlap with one another or the building to be constructed. Representing a facility by the coordinates of one corner \((x, y)\) the constraint to avoid overlapping between facilities \(i\) and \(j\) can be expressed as:

\[
\max [(x_j - x_i - l_i)(y_j - y_i) - w_i(y_j - y_i + w_i)] \geq 0
\]

where \(l\) and \(w\) represent the length and width of the facility respectively.

2.5.3 Tower Crane Constraint

Some facilities such as material storage yards need to be located within the reachable radius of the tower crane. This is mathematically represented as:

\[
(x_t - x_f)^2 + (y_t - y_f)^2 \leq R_t^2
\]

where \((x_t, y_t)\) represents the axis of rotation of the tower crane, \((x_f, y_f)\) represents the corner of the facility which is farthest away from the tower crane and \(R_t\) represents the tower crane’s pickup radius.

2.5.4 Site Accessibility Constraints

In order to facilitate smooth site operations, facilities should not obstruct the paths of material delivery trucks or other machines used on site. For the safety of the site personnel it is also necessary to ensure that work areas are not too close to these paths. These requirements could be incorporated into the model by defining the commonly used paths by trucks and machines and maintaining a clear distance between facilities and these paths.

2.5.5 Miscellaneous Constraints

Based on the layout engineer’s discretion and environmental concerns, a range of project specific constraints may be added. Sometimes the site may share one of its boundaries with a school or hospital. In such situations it must be ensured that facilities which generate a lot of noise should not be positioned near this boundary. In certain situations from prior experience, the layout engineer might require a certain relationship in the locations of two facilities. Such conditions can be formulated into mathematical relationships and incorporated into the optimization model.

2.6 Optimization using Genetic Algorithm

The optimization problem was solved using Genetic Algorithms (GA), which is a popular heuristic for solving the site layout problem. First a randomly generated large set of initial layouts is populated. The size, shape, orientation and coordinates of each facility were then encoded into genes and measured against a fitness function. The fitness function was taken as the inverse of the objective function. Thus, a layout with a large total travel distance will have a lower fitness value than a layout with a smaller total travel distance. We used the technique of Best First Search (BFS) in order to reduce the convergence time of the GA. The next step involves mimicking the process of natural selection by selecting the fittest genes, allowing them to create offspring and eliminating weaker genes. Each offspring undergoes a crossover, mutation and inversion operation with a certain probability, to prevent them from getting stuck at a local minimum value. The process is repeated over and over again until the genes in successive generations show no significant improvement. Finally, the gene with the best fitness is assumed to represent the solution of the optimization problem.
3 Demonstrative Example

The developed BIM based CSLP framework was tested and illustrated using an example project as follows. The project involves construction of a multi-storey reinforced concrete commercial building and an adjacent site office on a construction site of area 3600m². Autodesk Revit was used to create the BIM model with all of the material and structural information, as shown in Figure 4. The schedule was created using Microsoft Project allocating 300 days for the completion of construction activities.

By exporting the site layout plan from Revit, we could identify the available workspace as well as the boundary conditions of the site. The detected space was discretized into a square grid of 60 units x 60 units, to facilitate the use of the A* algorithm. Using the quantity take-off option in Revit 2014, we were able to export material information from the BIM model into csv files. Using the previously defined methodology, we were then able to determine the size of facilities such as the (1) scaffolding stockpile, (2) cement warehouse, (3) rebar storage yard, (4) rebar bending yard, (5) storage for mechanical fixtures, (6) storage for tiles, and (7) storage for panels. The size of the engineer’s caravan was selected to be 80m² based on the assumption that totally 12 engineers would be needed for the project. It was assumed that labourers would make multiple daily trips from temporary facilities to the required work locations inside the buildings. For activities involving rebar, we assumed that labour personnel would travel from the rebar storage yard to the rebar bending yard, and from the rebar bending yard to their respective work locations. For the sake of simplicity we assumed the centroids of the facilities to be the starting and ending points of each path. The construction schedule on Microsoft Project was exported as a csv file into our program which subsequently generated the facility requirements pertaining to three phases of construction. The required facility sizes and the phase of construction in which they appear are shown in Table 2.

Table 2. Temporary Facilities, their size and phase of requirement

<table>
<thead>
<tr>
<th>No.</th>
<th>Temporary Facility</th>
<th>Size (m²)</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Engineer’s caravan</td>
<td>80</td>
<td>P1,P2,P3</td>
</tr>
<tr>
<td>F2</td>
<td>Scaffolding stockpile</td>
<td>65</td>
<td>P1,P2</td>
</tr>
<tr>
<td>F3</td>
<td>Cement warehouse</td>
<td>75</td>
<td>P1,P2</td>
</tr>
<tr>
<td>F4</td>
<td>Rebar storage yard</td>
<td>80</td>
<td>P1</td>
</tr>
<tr>
<td>F5</td>
<td>Rebar bending yard</td>
<td>80</td>
<td>P1</td>
</tr>
<tr>
<td>F6</td>
<td>MEP storage</td>
<td>75</td>
<td>P2,P3</td>
</tr>
<tr>
<td>F7</td>
<td>Tiles storage</td>
<td>80</td>
<td>P3</td>
</tr>
<tr>
<td>F8</td>
<td>Storage for panels</td>
<td>85</td>
<td>P3</td>
</tr>
<tr>
<td>T.C.</td>
<td>Tower Crane</td>
<td>25</td>
<td>P1,P2,P3</td>
</tr>
</tbody>
</table>

3.1 Results

Table 3 compares the layouts generated from two cases: (1) Euclidean distance (the conventional method) and (2) actual travel distance (the proposed method). The layout generated from the conventional method yielded a total Euclidean distance of 2,700km, whereas the actual travel distance of this layout was calculated to be 3,750km. This shows that the presence of obstacles on the site accounted for the actual travel distance to be significantly greater than the centre to centre distance. This finding strongly highlights the under-estimation of the site layout problem by using linear distances. Figure 5 shows the layouts generated by the two approaches.

The actual travel distance driven optimization, yielded in a total travel distance of 3,130km, a reduction of 16.5% over the linear method. This shows that linear distance based optimization could lead to the generation of sub-optimal layouts. Furthermore, with an increase in the number of obstacles on the site, the superiority of the actual travel distance driven optimization becomes
even more significant. These results confirm that, optimizing the actual travel distance between facilities, generates layouts which significantly reduce unnecessary site-level transportation.

Table 3. Comparison of layouts using linear distances and actual travel paths

<table>
<thead>
<tr>
<th>Layout obtained considering linear Euclidean paths</th>
<th>Total Euclidean distance (km)</th>
<th>Total actual travel distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layout obtained considering actual travel paths</td>
<td>2,900</td>
<td>3,130</td>
</tr>
</tbody>
</table>

4 Conclusion

This paper presents a methodology framework which enables automating the CSLP process by leveraging the BIM technology. Based on this methodology framework, we created a tool to optimize the dynamic layouts of facilities on a construction site. This tool is flexible to changes in the design or construction plans and would significantly decrease the time spent in layout planning. Furthermore, by considering the actual travel distances in the CSLP optimization problem, a more accurate representation of construction activities is achieved. In the demonstrative example, our method generated layouts which reduced the total travel distance by 16.5%, a significant improvement over conventional methods. This improvement would be even more pronounced when dealing with larger sites with more obstacles. Future efforts will aim to integrate this model with 4D construction simulation to further aid the decision making of layout planners.

References


Figure 5. Comparison of layouts from linear distance and actual travel distance driven optimization


Contrivances to Assist Forest Machine Operator on Forest Road with Steep Slope

Katsutoshi Saibara a, Shigeomi Nishigaki b, Fujio Matsuda c and Shinichi Kubota d

a Kick, Co., Ltd., Japan
b Mazaran, Co., Ltd., and Kick, Co., Ltd., Japan
c Hitachi Construction Machinery Co., Ltd., Japan
d Mizobuchi Forestry, Co., Ltd., Japan
E-mail: saibara@c2mp.com, sleepingbear@c2mp.com, f.matsuda.kf@hitachi-kenki.com, mizobuchi-ringyo1967@deluxe.ocn.ne.jp

Abstract -
As for forest harvesting in Japan, machine operators have to handle their own forest machines on forest road with steep slope in the deep mountains. One common hazard that they are always facing is machine stability under different ground conditions. This paper presents the contrivances to provide the people concerned with low-cost Best Management Practices (BMPs), which would be firsthand knowledge of the harvest area being considered, and could support the machine operators who work for harvest operations, and logging road construction and maintenance. First, this paper presents visualizations of terrain attributes. Secondly, are described methods to reduce time and effort on collecting data required by construction, maintenance and restoration of forest road. Thirdly, this paper shows advance alert system pertaining to posture of machine body during the operations, and methods to provide the people concerned in advance with information on the situations of the forest road and the surroundings. Finally, remarks are described.

Keywords -
Forest road; Machine stability; Contrivance; BMPs; Open source software; R; Advance alert; Advance proximity warning

1 Introduction

A large quantity of man-made forest of Japanese cedars and Japanese cypresses have been planted in the postwar Japan. The harvest operations have been and are executed by small-scale firms of forestry contractors and landowners. Forest managers, machine operators, loggers, self-employed landowners, and other workers have been and are working for the harvest operations.

Figure 1. Deep forest mountains being analyzed
In this study, it aims to develop contrivances to provide them with low-cost Best Management Practices (BMPs), which would be firsthand knowledge of the harvest area being considered and could support them who work for harvest operations. The contrivances would be developed by utilizing open source software based on GNU General Public License [3].

Of particular concern in this study are the contrivances that aim to:
(1) Visualize terrain attributes
(2) Reduce time and effort on collecting data required by construction, maintenance and restoration of forest road,
(3) Provide operators with advance alerts pertaining to machine stability during the operations, and
(4) Provide machine operators, forest managers and landowners in advance with information on the situations of the forest road and the surroundings, where harvest operations would be planned and scheduled.

2 Terrain Attributes

The deep forest mountains are outlined and visualized below. The visualizations and the statistics presented here could be done by the R, which is a language and environment for statistical computing and graphics [4]. Besides, the other statics, for example, plan curvature, profile curvature, roughness, etc. could be also computed by the R [5]. For want of space, these statistics are left out in this paper.

Figure 2 shows the ground surface contour of the deep forest mountains and the existing logging road.

![Figure 2. Ground surface contour and existing logging road](image)

Figure 3 shows histograms regarding the statistics of the slope, aspect, TPI, and TRI of the deep forest mountain. The term of TPI is an abbreviation of the Topographic Position Index. The TPI is the basis of the classification system and is simply the difference between a cell elevation value and the average elevation of the neighborhood around that cell. Positive values mean the cell is higher than its surroundings while negative values mean it is lower [6]. TRI (Terrain Ruggedness Index) is the mean of the absolute differences between the value of a cell and the value of its 8 surrounding cells [7].

The slope distribution appears bimodal, two peaks (local maxima), and moderately skewed left. It can be seen from the histogram that all over the mountains has steep slopes more than 10 degree. The aspect distribution is bimodal, that is, easting or westing. The TPI distribution is symmetric and peaky, and has some outliers. The outliers are probably canyons or ridgelines. The TRI distribution is somewhat like bimodal, moderately skewed left, heavy-tailed at left, and has some outliers at right. It shows that some places in the mountains are very rugged.

![Figure 3. Histograms regarding the statistics](image)

There are two evaluation criteria sets with respect to terrain attributes as shown below.
(1) Criteria set based on TPI and slope [8]:
1) Canyon Bottom: TPI < -8,
2) Gentle Slope: -8 < TPI < 8 and Slope < 6°,
3) Steep Slope: -8 < TPI < 8 and 6° < Slope , and
4) Ridgeline: 8 < TPI ;
(2) Landslide susceptibility [9]:
1) Low: slope $\leq 11^\circ$,
2) Moderate: $11^\circ < $ slope $< 16^\circ$, and
3) High: $16^\circ < $ slope.
Figure 4 shows that there are many spots with steep slopes more than 20 degree. Although logging road should be constructed with avoidance of the spots with these steep slopes as far as possible, the logging road has been helplessly constructed for commercial thinning.

Figure 4. Visualization of slope

It can be said from Figure 3, Figure 4, and the above criteria sets that this terrain has:
(1) Some canyon and ridgeline;
(2) Many spots with steep slopes, and
(3) High landslide risk.

Figure 5 shows the TRI (Terrain Ruggedness Index). Looking Figure 5 at a glance, we can easily see that the existing road runs through the comparatively small difference in elevation of the deep mountains.

The logging road has been constructed for use all year long for harvesting and transport, and has unpaved surfaces of dirt and gravels with ancillary structures, such as:
(1) Skid trails
(2) Landings,
(3) Turntables,
(4) Turnouts
(5) Bridges,
(6) Fords,
(7) Grade breaks,
(8) Ditches,
(9) Culverts,
(10) Cross pipe,
(11) Riprap at ends of pipe or culvert, and
(12) Diversion swales, and forth

Figure 5. Visualization of TRI

The grade break is a gradual break in grade on a logging road to limit water flow by decreasing the concentration and the velocity. The diversion swale is a vegetated trough that carries or diverts water to reduce the erosion and to avoid washing over the top of the bank.

3 Reduction of Time and Effort on Collecting Data

It is time-consuming to gather data in order to identify dangerous spots and excessive inclined spots, and any damage and deterioration of the ancillary structures. The dangerous spots and excessive inclined spots might be latent in logging road, and could cause skidding or rolling over of machine and make operator slip, trips, and fall out of the machine. The damage and deterioration of the ancillary structures could cause falling rocks, landslide, water flood, and so forth. The advance alert system could be applied to reduce time and efforts on collecting data pertaining to dangerous spots latent in and slope of the logging road, and any damage and deterioration of the ancillary structures.

3.1 Finding Dangerous Spots Latent in Logging Road

Forest environments in the deep mountains are always changing dynamically. The changes are often unpredictable. Machine operators have to handle their own forest machines based on their own realistic
sensation given on their seat of pants in such severe environment.

The dangerous spots might be caused by changes in operating conditions such as wash-outs, gulleys, wheel ruts, puddling, frost heaving, soil disturbance, erosion, disruption due to water flood, sediment, and so on. Then the dangerous spots might be latent in the logging roads.

An operator could drive a light motor vehicle with the advance alert system on the logging road. Severe phenomena could be detected by identifying change points based on existing trends of consecutive increasing or decreasing values in time series of tri-axial acceleration responses. Finding seven consecutive increasing or decreasing values is almost the same as the probability to have a car accident in one year in Japan. Then it becomes possible to find severe phenomena related to the dangerous spots of those, for example, front and back skiddy spots, sideway skiddy spots, and severe bumpy spots.

Figure 6 shows the time points of advance alert “caution” that was triggered based on tri-axial acceleration responses obtained by driving a light motor vehicle on the logging roads.

Figure 6. Time points when advanced alert “cautions” being triggered

In order to avoid any accident of skidding or roll-over, the advance alert system would give operators advance proximity warning during driving their forest machine when entering into the proximate area of 100m range from the dangerous spots like those. The operators could watch the advance proximity warning on the in-cabin monitor and would slow down the speed.

Front and back skiddy spots, sideway skiddy spots, and severe bumpy spots have almost same position on the logging road. For want of space, the sideway skiddy spots are shown in Figure 7.

Figure 7. Sideway skiddy spots

3.2 Measuring Slope of Logging Road

To do a walk through assessment of work environments, the light motor vehicle with the advance alert system in Figure 8 is repeatedly stopped during one minute and driven 100 m distance on the logging road.

Figure 8. Light motor vehicle with the advance alert system

The advance alert system could calculate and record cross and longitudinal slopes automatically based on the static acceleration of gravity when stopping the light
motor vehicle. The cross and the longitudinal slopes of the logging road are shown in Figure 9 and Figure 10, respectively.

It is said to avoid constructing logging road with grades in excess of 10 % [6], [10]. From these Figures, we could easily find dangerous spots with slope in excess of approximately 6 degree, that is, grades of 10%.

Drive Downhill and slash disposal work on the above steep logging road are shown in Figure 11 and Figure 12, respectively.

Figure 11. Drive Downhill

Figure 12. Slash disposal work

In the machine operator’s opinion, it becomes hard to operate forest machine at the spot in excess of 12 degree. By the way, logging road should be less than 18%, that is, 10.3 degree in compliance with the logging road construction standard by Kochi prefecture [11].

3.3 Periodically or Rapidly Mobile Inspections

Generally, inspection activities are periodically conducted. Then inspection results are documented and pictures are taken as evidence of current conditions or damages or deterioration with a measuring scale or some other reference object. Furthermore, rapid inspection activities would be done for the purpose to determine if further attention to and reconstruction of the ancillary structures are necessary, typically following a typhoon, earthquake, and so on.

When stopping the light motor vehicle, operator will inspect and take pictures of current conditions of the ancillary structures and the surroundings. The each field
data of those is automatically tagged with time, GPS locations and recorded in digital camera, mobile phone or tablet devices, and also manually uploaded to the database management system.

Automatically or manually, acquiring and recording data as described above could reduce time consumed to gather field data in ordinal situation, and to narrow down and choose certain points to be watched.

Operators could also review previously the field data before their timber-harvest operations. Forest manager could get significant information on planning and activities necessary for the forest harvesting management. Landowners could confirm the current conditions of the logging road and the ancillary structure.

4 Advance Alert System of Machine Posture

When operating forest machine at dangerous spots as noted before, it is significant and required to compensate operator’s realistic sensation. The advance alert system consists of an on-board tri-axial accelerometer as shown in Figure 13 and an in-cabin monitor system as shown in Figure 14.

In this study, it is assumed that acceleration responses measured along cross dimension might represent sideway skid, and ones along longitudinal dimension front-and-back skid. The advance alerts are derived from the existence of trends such as consecutive increasing or decreasing values in time series of acceleration responses, which are sent via wireless communication from on-board tri-axial accelerometer fixed on the carriage body of machine to the in-cabin monitor system.

As mentioned before, seven consecutive increasing or decreasing values equals to the probability to have a traffic accident. Furthermore the thirteen equals to the probability to be killed in a traffic accident in Japan. In the advance alert system, the former is defined as “Caution,” the latter “Danger,” and otherwise “Safety.”

The advance alert system enables the operators to watch advance alert pertaining to sideway and front-and-back skids, and bumping of their own machine. The monitor displays numeric and graphic indicators pertaining to warnings of the colored classes such as:

1. Green referring to safety,
2. Yellow referring to caution, and
3. Red referring to danger.

Operators could watch the advanced alerts in real time, which would compensate their own realistic sensations given on the seat of their pants during operation. Then, the risk latent in harvest activities could be reduced.

5 Remarks

People concerned with forest harvesting have to consider far-flung and different coverage problems related to the work plan and operations, the conditions of the work site and the logging roads, the silviculture plan and maintenance, and so forth.

Operators have to examine many different factors in site level decision-making regarding the following matters [10], [11], [12]:

1. Timber-harvest operations with suitable risk controls as considering location of possible dangerous spots,
2. Preferred location of log extraction, tracks and log landings,
3. Machine servicing,
4. Minimization of impacts on forest resources
5. Potential water quality,
6. Seasonal ponds,
7. Soil productivity,
8. Work environment information,
9. Latitude, longitude, and altitude,
10. Ground surface contours,
11. Marginal plot of across each coordinates of forest mountains,
12. Fundamental statistics of slope and aspect, and
Forest manager needs to investigate and deliberate problems in decision-making regarding:
(1) Administration of commercial timber-harvest,
(2) Number, size and design of forest access roads,
(3) Resource allocation,
(4) Maintenance of forest health and productivity, and
(5) Cost effectiveness on harvesting and transportation, and also to work on the best settings for forest harvesting.

Landowners, especially absentee ones and persons of advanced age, want to know current conditions of their own forest mountain at their own home for:
(1) Commercial thinning, which would be carried out at the right stands at the right time under appropriate stand conditions [13],
(2) Precommercial thinning to enhance woodland productivity,
(3) Redistribution of growth potential to fewer trees past the sapling stage,
(4) Leaving a stand with a desired structure and composition,
(5) Improving the quality of the stand by removing dead, disease and deformed trees, and
(6) Long-term silviculture and harvest plan

Figure 15 shows scheme of contrivances to provide them with low-cost Best Management Practices (BMPs), which would be firsthand knowledge as described above. Since the contrivances are inexpensive and easy-to-use system, its up-front cost might be a less burden to small-scale firms of forest contractors and landowners.

The contrivances consist of the front-end devices and the back-end platform. The former is on-board devices that function real-time monitoring system for the advance alerts pertaining to dangerous spots and machine stability, and for measurement slope of logging road. The former enable operators to confirm the situations of logging road, the ancillary structures and the surroundings in advance, where the forest harvesting is scheduled. Besides, the operators could watch the advanced alerts in real time, which would compensate their own realistic sensations during their operations. The operators armed in advance with information on the work site could make a suitable plan of harvest operations as considering environmentally sensitive maintenance and restoration. In addition, this deviceBrowse work environment information by on-demand processing of requests.

The latter functions exploratory data analysis to generate statistics being required and to visualize them, and manages data gathered and front-end devices, and so on. Machine operators in the cabins and a forest manager on site will be able to import findings on site, which are related to the matters as described before, into this platform from their front-end devices. The people concerned, for example, landowners, forest managers, can obtain information relevant to their harvest operations by accessing the back-end platform.

Consequently, the risk involved in harvest activities would be reduced.

References
[8] Dickson, B. and Beier P. 2006. Quantifying the influence of topographic position on cougar (Puma

Figure 15. Scheme of contrivance


Importance of Planning for the Transport Stage in Procurement of Construction Materials

A. Ahmadian F.F.*, A. Akbarnezhad*, T. H. Rashidi*, and S.T. Waller*

*School of Civil and Environmental Engineering, University of New South Wales, Australia
E-mail: ahmadian@unsw.edu.au, a.akbarnezhad@unsw.edu.au, rashidi@unsw.edu.au, s.waller@unsw.edu.au

Abstract -

The use of logistics milestones, as predecessors for construction works, is gaining popularity to improve the control over the material procurement stage and minimize the risks of delay in the project completion resulting from late material deliveries. Transportation is an important phase in the procurement process which can account for up to 20% of the total project expenditure in some industrial construction projects. However, despite its importance, transportation is usually overlooked when planning the material procurement stage and evaluating the potential risks of delay. This study evaluates the current practices in managing transportation stage of the construction material procurement process and the perception of the industry practitioners about the importance of considering material transportation in project planning using the results of a limited systematic survey. In addition, the importance of transportation in the procurement of construction materials is studied by considering the actual data on shipment of construction materials/components in two industrial projects. Results indicate that material handling procedures adopted in industry are highly disorganized and transportation variables are ineffectively articulated. The results highlight the need for methods to plan, monitor and control the transportation stage as an independent activity in the material’s life cycle. Besides the travelling distance, weight, dimension, mode of transportation, and terms of delivery were identified as the main factors affecting the transportation of the construction materials.

Keywords -
Transportation; Construction Material; Supply Chain Management; Planning

1 Introduction

Construction contractors are increasingly engaged in supply of material from diverse sources around the world and statistics have declared that more than 65% of a construction project budget is spent on procurement of materials [1]. This process starts with design and engineering pursued at manufacturers’ workshop, and ends with a series of journeys from factories to the construction site. Planning such a long supply chain requires scrutinizing all the stages involved in terms of time and costs [2].

On the other hand, delay in construction projects has always been an issue of concern in both academia and industry. Late delivery of construction materials and components has been identified as one of the main causes of delay in major industrial construction projects. Therefore, timely delivery of materials is essential to ensure meeting completion date of construction activities [3].

However, on-time delivery of materials is a complicated job and requires planning, monitoring and control of different stages of the supply chain including the transportation stage [4]. Off-site transportation of materials has been estimated to account for 10 to 20% of the total project expenditure in typical industrial construction projects [5]. However, despite its importance, little has been done to investigate the efficiency of the current transportation management practice and potential strategies to improve the latter. This study aims to investigate the current practice in planning of off-site transportation in the industrial construction projects and identify the main factors influencing the efficiency of the transportation stage. The available literature on planning, monitoring and control of the transportation stage is reviewed. The current practice in management of the transportation stage of the construction material procurement process and perception of industry practitioners about the importance of considering material transportation in project planning is evaluated using the results of a limited systematic survey. In addition, the importance of transportation in the procurement of construction materials is studied by considering the actual data on shipment of construction materials/components in two industrial projects.
2 Literature Review

2.1 Importance of Transportation in the Supply Chain

The current practice in the construction management field for developing the supply chain is to optimize the overall process by considering the costs trade-offs between three main components of the chain; i.e., transportation, inventory, and production [6]. As a result, this has shaped the dominant mindset in construction industry researches to concentrate on devising cost efficient approaches for logistics and shipment strategies, as one of the three elements. An illuminating example is a part of the research that has been conducted by Irizarry et al. [7] in which Building Information Modelling (BIM) and Geographic Information Systems (GIS) are deployed to manage logistics perspectives of construction material supply chain and then reduce the relevant costs through the trade-off between inbound and outbound transportation costs. Their developed model is also capable of finding the optimum way of transporting material from a supplier to the construction site, in terms of number and time of orders, order quantity, transportation methods, and the associated cost of transportation. These efforts are supposed to be used before the formation of contract between a general contractor and suppliers.

Although the main theme in analysing transportation is founded on cost, there is an interdependent relationship between cost and time. As a result, calculation of transportation time can lead to deciding about optimum transportation options, even in term of cost [8]. On the other hand, measuring time in real situations is simpler and more practical than cost for logistics management. Therefore, it is a reasonable outlook to focus on duration in order to come up with the best option for shipment of goods. By using time as the target for managing movements of purchased materials, it can simply help evaluating how well the companies have planned their shipment and how well they have managed to carry out the plans. In addition, a key to successful improvement of any process is accessibility of adequate information about that process as well as effective management of information [9]. Whereas gathering information about time is relatively easier rather than obtaining and collecting highly sensitive data on cost, using time measures for providing feedback and improving the performance of project material delivery process would have a greater chance to be realized. Therefore, in this research it has been decided to just concentrate on time data of construction material supply chain to assess performance.

2.2 Diversity of Material and Their Lead Time

To address problems associated with the delivery of construction materials, it is vital to comprehend different types of materials and their corresponding supply chain. There are four groups of materials which are named as engineered-to-order (ETO), made-to-order (MTO), assembled-to-order (ATO), and made-to-stock (MTS) products [10]. ETO products are special items that are made from raw material in a manufacturer workshop based on detailed design information and drawings received from engineering companies. Therefore, ETO items are expected to have a relatively long supply chain. MTOs are defined as components that are fabricated (or prefabricated) upon receiving orders from customers without intensive needs to engineering data from other parties. They differ from ATO products as ATOs are assembled from standard available elements. Finally, MTS products are resources only controlled by the in-stock availability of the manufacturer. These four categories of products are usually prioritized based on longevity of their supply chain as ETO, MTO, ATO, and MTS, respectively.

Another common classification of material in the construction industry is based on the lead time. Two main categories of material identified in this classification are critical items and noncritical items. Critical items that are also prominent as long lead items are “those components of a system or piece of equipment for which the times to design and fabricate are the longest” [11]. As expected, noncritical items should be those with shorter supply chains. By comparing these two material categories, it may be reasonable to assume that ETOs and part of MTOs will constitute the majority of the critical items; whereas the rest of MTOs, ATOs, and MTS products may be classified as noncritical items.

Although supply chain of any type of material has its own problems, transportation concerns mainly emerge from delivery of prefabricated components in terms of packing and dimension [12]. The prefabricated components can be put into either MTO or ATO categories of material. There is a lack of rich literature on the transportation phase of supply chain and its importance on the efficiency of the procurement process. This mainly originates from the existing perception in the industry that impact of the transportation stage on plans is insignificant and any problem associated with the transportation of construction materials is in fact a consequence of another problem in the previous steps of the supply chain [13].
2.3 Role of Communication Technology in Managing Delivery of Construction Material

A number of previous studies have focused on logistics to improve transportation of construction items. The dominant methodology to evolve current practices for timely delivery of material appeals to application of communication and information technologies [14]. Monitoring and tracking of deliveries has been determined to be an effective way to ensure timely and efficient handling of delivered material at the jobsite and prevent assembly interruption. For this purpose, using advantages of the Internet was examined in a study conducted by Williamson et al. [15]. To make use of the current achievements in visual monitoring in delivery of construction materials, a system combining BIM and GIS has been established to provide project management team with early warnings on any late arrival of construction components [7].

Lu et al. [16] made a communication network, by integrating Global Positioning System (GPS) with vehicle navigation technology, Bluetooth, and commercial mobile that are connected to a central PC, to position concrete mixer trucks in highly dense urban areas. It was shown that they had better utilization of those vehicles at concreting sites. Radio Frequency Identification technology (RFID) also has been used as a tool to automate custom clearance process resulting in reduction of stoppages and saving time that is normally being lost in import cargos in international air cargo terminals [17].

Although a great deal of effort has been made to improve the efficiency of the supply chain of construction materials through the use of communication technologies; little has been done to investigate the importance of the transportation stage and developing methods to improve the efficiency of the transportation stage.

2.4 Delivery Strategy

In international trades, transport of goods and material between seller and buyer requires a number of decisions to be made. Two important decisions that may influence the cost and the speed of the shipment are mode of transport and terms of delivery. While mode of transport is usually chosen after a trade-off between cost and duration of delivery, finalizing terms of delivery depends on risk-taking attitude and power of the parties involved [18]. In other words, assuming availability of all transportation options, the decision to use train, truck, vessel, and airplane is generally conformed to the limitation of time and cost. In statement of trade terms, the location where ownership is transferred, as well as arrangements for payment of any cost associated with delivery of items are clarified and it is a likely approach that each party would try to transfer all the risks and responsibility of paying all costs to the other party. INCOTERMS that are defined and published by the International Chamber of Commerce are put into action to minimize the probability of misunderstandings in such deals. INCOTERMS are a set of three-letter standard trade terms that are classified into four groups and distinguished by increasing obligations and risk assumed by the seller. In the E terms, seller is just responsible to make goods available at its own facility while in F terms his obligation is extended to delivery of the goods to the carrier appointed by the buyer. In group of C and D, however, seller will accept an increasing responsibility toward delivery of goods to the place of destination [19].

3 Current Practices in Planning Delivery of Construction Material, Outcomes of a Qualitative Survey

Twelve construction industry professionals with characteristics summarized in Table 1 were surveyed about their experience in planning logistics and transportation of components between suppliers’ workshop and construction site. To increase reliability of the results, interviewees with different backgrounds in terms of organization, industry, and functional role were selected.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Owner</th>
<th>Contractor</th>
<th>Consultant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>12-20</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>20-25</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Above 25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.Sc.</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>M.Sc.</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Industry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oil/Energy</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Petrochemical</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mineral</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Functional Role</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Management</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Procurement</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Planning</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

They were from energy, petrochemical, or mineral processing sectors in which supply of material is a significant part of executing a project. Each of them possessed responsibilities for managing the delivery of...
material, either in owner, contractor, or consultant companies. Survey participants were asked to judge mainly by considering only the items which require international trips as a part of their delivery process. The outcomes of interviews are discussed in the following.

3.1 Significance of Transportation in Supply of Material

The results of the survey performed as a part of this study showed that transportation plays a significant role in the procurement process under following circumstances:

- **Reducing Cost of Inventory:** The cost of inventory usually accounts for a considerable portion of the material handling costs. Therefore, reducing the costs of inventory through using more sophisticated material management approaches such as the just-in-time inventory approach is commonly applied in some industries to reduce the overall material handling costs. Successful implementation of such methods to reduce the inventory duration and costs requires precise planning of transportation as the last step before storing material.

- **Space Limitation:** Space limitations on the job site is another common problem faced by project managers in many construction projects. When facing difficulties in placing commodities of project on the job-site, project managers try to postpone the delivery of materials until they are actually needed. In such cases, keeping track of the delivery of goods to the site and thus management of transportation gain considerable importance. This situation may be observed in offshore projects, projects implemented in urban areas, or in the middle of some congestive construction projects.

- **Propagation of Delay:** Transportation will be brought under special considerations, when the previous stages in the supply chain of material are delayed and the float time has elapsed. In such cases, alternative modes of transport are usually evaluated to accelerate the delivery timing.

- **Mandatory Completion Date:** When there is a tight schedule for the whole project, like projects with a fixed finish date, all project tasks including transportation activities should be scheduled carefully.

3.2 Planning the Delivery

The planning of delivery is conducted to develop delivery schedules, terms of delivery and the mode of transport. The results of the survey conducted to investigate the current practice in planning of deliveries are summarized in the following:

- **Scheduling:** To detail material delivery milestones, pro-rata allocation of shipment time based on the overall supply chain duration is usually performed; i.e. a longer manufacturing time, more time is needed to transport it!

- **Term of Delivery:** Trade terms for delivery of materials between supplier and contractors are set while trying to transfer risks to the suppliers. This is because there is a common perception among contractors that suppliers are generally more experienced in material movement and handling.

- **Mode of Delivery:** In cases where contractors are kept responsible for transportation, the major deciding factor for selecting the mode of delivery is the cost at the time of planning. Decision making about the mode of delivery is usually performed based on personal judgments of logistics team about the costs and a detailed analysis is seldom performed. Unforeseen conditions sometimes force the management to change the delivery mode.

3.3 Special Issues in Managing Delivery of Components

- **Short-Term Nature of Dealing with Suppliers:** One-off and temporary relationships between construction firms and suppliers were identified by the survey participants as the root cause of poor management of the delivery process.

- **Traffic Management:** Obtaining the required permits from road and transportation authorities is in some cases a time-consuming process. It takes endeavours to precisely study the current road limitations and detail allowable load and dimensions of the consignments in conjunction with the considerations for the existing traffic.

The results of the survey show that this is one of the common difficulties faced when projects require the shipment of heavy and bulky items fabricated offsite. The survey highlights the necessity of in-advance definition of strategies for ordering large and heavy equipment and components.

4 Case Studies

The transportation stage of construction materials in two actual industrial projects were studied to investigate the current practice in planning of transportation in procurement of construction materials. The projects were selected intentionally from oil-refinery construction projects which require usually the procurement of a considerable amount of materials and components fabricated offsite. The scope of the first project was revamping and expansion of an existing oil refinery, while the scope of the second project was construction of a new gas refinery. Both projects are
located in the Middle East and are currently in the construction phase by different contractors. Components required in both projects are procured from a series of international suppliers.

Table 2. Number of purchase orders with various characteristics arrived at the construction sites

<table>
<thead>
<tr>
<th>Characteristics of Orders</th>
<th>Structural</th>
<th>Mechanical</th>
<th>Piping</th>
<th>Instrument</th>
<th>Electrical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (ton)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light (&lt;1)</td>
<td>7</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Medium (&gt;1, &lt; 10)</td>
<td>6</td>
<td>5</td>
<td>22</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Heavy (&gt;10, &lt;100)</td>
<td>9</td>
<td>14</td>
<td>15</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Super heavy (&gt;100)</td>
<td>21</td>
<td>12</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dimension (m³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small (&lt;1)</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Medium (&gt;1, &lt; 10)</td>
<td>6</td>
<td>5</td>
<td>25</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Large (&gt;10, &lt;100)</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Super large (&gt;100)</td>
<td>16</td>
<td>11</td>
<td>10</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Term of Delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>33</td>
<td>21</td>
<td>29</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

These projects require the procurement of more than 3500 diverse items which can be grouped into structural, mechanical, piping, instrumentation, and electrical groups. The total number of purchase orders made and delivered to the construction sites by the time of this study is 149. Table 2 summarizes the characteristics of the orders in terms of weight, dimension, and trade condition of delivery.

In compliance with practitioners’ point of view, investigation of initial time schedules revealed that the planned transportation duration, from vendor’s shop to the final destination, has been estimated initially as a proportion of the overall supply chain length. The ratio of planned transportation duration to the total procurement duration was found to vary from 10 to 20% for the majority of items. In this study, the overall procurement duration is considered as the time period between placing a purchase order and its arrival at the construction site.

Table 2 shows that 66% and 11% of all items are categorized respectively under C and D terms of delivery. This originates from the expected risk-aversion tendency of contractors persuading them to transfer as much risks as possible to sellers.

The data available on material handling durations demonstrates considerable deviations from the planned durations for transportation activities. Figure 1 examines the relationship between the duration of transportation and the total supply chain duration for various items delivered. In this Figure, items have been divided into ETO, MTO, ATO, and MTS categories described earlier. As shown, the ratio between duration of transportation and total supply chain duration varies from 2 to 70%. This clearly contradicts the planning strategy adopted in these projects which assumes a directly proportional relationship between the duration of transportation stage and the total supply chain duration. The results presented in Figure 1 show that this commonly used rule of thumb for estimating the duration of transportation is invalid for about 60% of the items investigated in this study and therefore highlight the need for a more reliable method for planning the transportation stage. It may be concluded that the period of shipping should be estimated by identifying the parameters affecting the transportation stage and regardless of the duration of other stages of supply chain. Therefore, investigating the factors affecting the duration of transportation and developing methods for quantifying the impact of each factor are highly appealing.

Figure 1. Actual duration of transportation vs. total supply chain length for different categories of construction material

Mean values and standard deviations (STDEV) for delivery period and average delay of items in different material categories are compared in Table 3. As shown, material categories with shorter lead time show generally higher deviations from the planned transportation durations than items with longer lead times. On the other hand, as shown in Figure 1, for materials with relatively shorter lead times, transportation may be considered as the most dominant
phase of the supply chain in terms of duration, accounting for more than 40% of the total supply chain duration in some cases.

Table 3. Delivery duration for different categories of material and its deviation from planned value

<table>
<thead>
<tr>
<th>Material Category</th>
<th>Actual duration of Transportation (Days)</th>
<th>Delay (Days)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (STDEV)</td>
<td>Mean (STDEV)</td>
</tr>
<tr>
<td>ETO</td>
<td>26.49 (1.28)</td>
<td>1.20 (0.85)</td>
</tr>
<tr>
<td>MTO</td>
<td>29.18 (5.71)</td>
<td>5.84 (6.67)</td>
</tr>
<tr>
<td>ATO</td>
<td>25.33 (9.63)</td>
<td>10.31 (6.02)</td>
</tr>
<tr>
<td>MTS</td>
<td>22.31 (7.84)</td>
<td>12.7 (4.91)</td>
</tr>
</tbody>
</table>

In other words, whereas travel time for different types of components from a specific destination will relatively remain the same, the effect of probable transportation delay on the overall supply duration seems to increase with a decrease in the lead time of items. Therefore, the results suggest that despite the common perception, the effects of material transportation on material procurement delays are more significant for off-the-shelf or MTS components and therefore, a greater deal of attention should be paid to planning the transportation of the items in this material category.

![Figure 2. Transportation duration versus overall procurement duration](image)

This fact is highlighted in Figure 2 which indicates a schematic view of timescale for supplying different categories of material (from a specific origin) and draw attention to importance of precise planning for delivery of items with shorter lead time. Among different categories of construction material, any delay in ETO and MTS types may directly push back completion date of the project. This is because ETOs are needed to perform critical activities and MTSs are general bulk materials that are consumed in both critical and non-critical paths of the project. Hence, this fact will intensify significance of taking proper strategies for planning delivery of MTS components.

![Figure 3. Effect of shipment variables on probability of delay in transportation period](image)

The shipment data of the case projects were also used to investigate the effect of weight, dimension, mode of transport, and term of delivery on probability of delay. Figure 3 shows the variation in the likelihood of time overrun in transportation of materials to the construction site with variation in the aforementioned variables. As it can be seen, the results indicate that despite the common perception in the industry, an increase in weight and/or dimension of consignments is not necessarily accompanied with an increase on the likelihood of delay in the transportation phase. Variations in weight and dimension of items seem to have relatively similar effects on the likelihood of delay although the amount of occurred delay is absolutely different for each group (Table 4).

![Figure 3. Effect of shipment variables on probability of delay in transportation period](image)
the delivery of items which could be resulted in boosting risk of delay, because that party will accept responsibility of shipment even in the region of the counterpart. In general, the results of this case study suggest that realistic transportation plans cannot be achieved by considering the travelling distance as the only variable. Besides distance, the dimensions and weight of item, mode of transport and terms of delivery should be taken into account when planning the transportation activities.

Table 4. Range of delay in shipment of orders classified based on combination of weight and dimension

<table>
<thead>
<tr>
<th>Density Index</th>
<th>Weight</th>
<th>Dimension</th>
<th>No. Of Orders</th>
<th>Average Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Light</td>
<td>Small</td>
<td>14</td>
<td>5.41</td>
</tr>
<tr>
<td>2</td>
<td>Light medium</td>
<td>Medium</td>
<td>6</td>
<td>6.4</td>
</tr>
<tr>
<td>3</td>
<td>Light heavy</td>
<td>Large</td>
<td>2</td>
<td>6.75</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
<td>Medium</td>
<td>21</td>
<td>6.17</td>
</tr>
<tr>
<td>5</td>
<td>Heavy</td>
<td>Medium</td>
<td>27</td>
<td>10.03</td>
</tr>
<tr>
<td>6</td>
<td>Super heavy medium</td>
<td>Super large</td>
<td>28</td>
<td>9.35</td>
</tr>
<tr>
<td>7</td>
<td>Heavy super heavy</td>
<td>Super large</td>
<td>32</td>
<td>4.95</td>
</tr>
<tr>
<td>8</td>
<td>Super heavy</td>
<td>Super large</td>
<td>19</td>
<td>3.78</td>
</tr>
</tbody>
</table>

As shown in Table 4, a density index (based on the range of weight and dimensions of items) was defined and introduced to evaluate the effects of various combinations of item’s weight and size on delay. It shows that in the case projects presented here, the items with density index of 5 experienced the maximum delay. Items with density index of 5 normally include packages of large numbers of different small items such as packages of thousands of piping accessories and fittings. One probable reason for the higher delay experienced by such items is the typically long and labour intensive custom clearance process required. The second highest delay corresponds to items with density index of 6. According to the project database, majority of items in this category have irregular shapes and centre of gravity which are likely to experience delays during loading and unloading. Table 4 moreover shows that items with density index of 7 and 8 showed minimum delays. Items in this category were mostly huge equipment in cubic or cylindrical packs which require usually less inspection in customs and experience less difficulties in loading and unloading.

5 Conclusion

Construction contractors are permanently concerned with improving supply chain of material. In this way, advancement in logistics management has been proven to be an important ingredient for success. Although reducing costs of logistics can benefit involved parties, it should not direct them to the pitfall of overlooking time management in movement of material. Moreover, time is a readily available measure than can be used as an indication of performance in cost management. Consequently, it is essential to spend efforts on developing cautious time schedule for transportation of construction material and carry out this plan.

Current practices in delivery of material to the construction sites are mainly in connection with movement of special items on local roads and squeezing schedule of transportation to reduce effects of delays during prior stages. While such practices are accredited, still modification of the mind-set in estimating duration of shipping is required. In addition, dimension and weight of orders as well as mode of transportation are effective factors in calculation of delivery length. Parties also should be aware that transferring risks through trade terms may create a new risk causing late arrival of their consignments.

Transportation is an issue that calls for approaches which are not normally influenced by duration of their engineering phase or fabrication at factories. Therefore, considering factors like traffic management, loading and unloading problems, custom clearance issues, and cooperative approaches between buyer and seller is vital to realistic estimation of delivery time and would help the project team to successfully manage this milestone. Nonetheless, it may have significant contribution in time management for supply of material that have short lead time.

6 References

[1] Scholman H.S.A., Subcontracting by main contractors, Economic Institute for the Building,
Amsterdam, 1997.


A Formwork Layout Model based on Genetic Algorithm

D. Lee*, H. Lim*, T. Kim*
*H. Cho* and K.I. Kang*

*School of Civil, Environmental and Architectural Engineering Korea University
Seoul, Republic of Korea

Research Institute of Engineering and Technology, College of Engineering, Korea University
Seoul, Republic of Korea

* Corresponding author (hhcho@Korea.ac.kr)

Abstract -
The recent increase in the number of irregularly shaped tall building has caused a continuous increase in the formwork costs. To reduce the cost of forms, modular forms are widely used since they are cheap and reusable. However, they cannot be applied around columns which are irregularly arranged. This study introduces a formwork layout model using a genetic algorithm to reduce formwork material costs. The proposed model seeks an optimum formwork layout to maximize the proportion of modular forms by arranging 12 types of modular forms with the columns, for which the positions and sizes of are changeable. From a case-study, the model showed a decrease in the ratio of non-modular forms covering the area by about 11.9%, and the rental cost of forms can be reduced by 10.4% compared to the previous heuristic method. The model can be used in the structural planning stage by a structural engineer to reduce the cost of formwork, and increase the efficiency of the column layout.

Keywords - formwork layout; optimization; genetic algorithm; tall building

1 Introduction

The recent increase in the number of irregularly shaped tall building has caused a continuous increase in the formwork costs (Kim et al 2012). The material cost of forms has especially increased, while the labour cost has been reduced by the application of systemized formwork methods, since the number of non-modular form covered areas was highly increased because of irregularities. As the material cost of forms depends on the number and type of forms, the layout of formwork has become a key factor determining the formwork cost (Cha 2012).

However, the present formwork layout planning method has several problems. First, planning is carried out in highly limited conditions because the structure of the building has already been fixed. For fixed structure layout conditions, only a few inefficient formwork layouts can be used. Second, the current formwork layout planning is conducted by someone experienced in heuristic approaches. Since there are an infinite number of formwork layout alternatives in a floor, it is impossible to make a quantitative decision about which layout solution is better within a short time, especially in large-sized and freeform shaped panels.

This study suggests a model for formwork layout planning to reduce formwork cost using genetic algorithm. The proposed model finds an optimal formwork layout alternative which maximizes the ratio of modular form-covered area by adjusting the size and position of the columns. To demonstrate the benefits, a case-study is implemented and then the result is compared with the layout alternative by an expert.

2 The layout planning of formwork

Existing studies related to formwork were overly weighted on selection of the formwork method. They focused only on the constructability and work productivity of the formwork, regarding the layout planning which must precede the formwork as unimportant (Lee et al 2009; Huang et al 2004; Tam et al 2005).

However, through in-depth interviews with practitioners from manufacturers, formwork engineers and suppliers, we found that one of the main issues with horizontal formwork is the process for finding optimal formwork layout solutions with the least cost.

When conducting formwork, the cost depends on the number and the type of the forms used. Therefore, the cost can be analysed quantitatively with different formwork layout alternatives (according to how many
or what kind of forms are used).

There are two types of aluminium forms: the modular form, which is standardized in shape, and the non-modular form, which is not. The non-modular form is not reusable, commonly making the cost about 1.5~2 times more expensive than modular forms. This means that the formwork cost can be minimized by arranging forms of standard sizes to cover as much of the given space as possible, i.e., by minimizing the use of special units.

The objective of previous formwork layout planning was to cover the area with as many modular forms as possible. An expert arranged the forms on the floor plan from a start point to the end, from the biggest to the smallest, based on experience. The modular forms (standard sizes of which are 300mm, 400mm, 450mm, and 600mm in width by 900mm, 1050mm, and 1200mm in length) are preferentially arranged, and non-modular forms are used only in the case when a modular form cannot be arranged.

In addition, when the formwork engineer designs aluminium formworks, beam units like end beam (EB) and main beam (MB) are always used to do the work more easily and safely. The beams attach aluminium forms more tightly, and prevent any kind of deflection even if all the forms are not similar enough (in size or position). Therefore, when making a layout planning model, the combination of beams must be considered. See Fig 1.

Figure 1. Main beams between the forms

According to the shape and the size of the floor plan, there are infinitely many kinds of formwork layout plans that can be made, so it is hard to find great layout alternatives within a short time. Especially, as the shape of buildings has become irregular and the layout of columns arranged unsystematically, the number of non-modular forms has increased highly around the columns. This directly influences the rental cost of forms.

In short, previous formwork layout planning methods have two problems. One is that the formwork layout planning is conducted by someone with individual experience and intuition, making the planning lower-quality and cost-inefficient. The other is related with the formwork planning time, in that planning is carried out after finishing the structure layout planning. Since the layout planning is decided on a fixed structure layout, including locations of columns and walls, if the structural layout is complex, it cannot be helped but to use many non-modular forms.

Therefore, a computerized and automated layout planning model for optimum formwork layout which arranges forms at the planning stage without any limitation of the columns layouts would be greatly beneficial.

3 Formwork layout model using genetic algorithm

Through interview of experts who majored in the structure of buildings, we realized that changing the column layout (length, width, position etc.) within an allowable nominal area should be flexible at the planning stage. It would be critical to change the structure for the formwork, but considering that such change is required by the constructor and employer to reduce cost, limited adjustment can still be valuable enough.

Formwork layout planning is conducted with lines since all of the forms are rectangular and are arranged in good order. This is similar to the two dimensional rectangle cutting stock problems. However, it is an NP(Non-Polynomial)-hard problem, because there are many variables. The position and size of the columns, forms and beams are changeable. There is no stiffness when planning the layout and the placement must be done one by one, which is why a heuristic algorithms was chosen instead of a mathematical method.

Like the previous formwork arrangement method, as many modular forms as possible were arranged to optimize the space, from the biggest form to the smallest one. Only when forms encountered columns, columns were deleted and the space was used as an optimization area. If there were columns in the same section in line, they were joined together and optimized at the same time.

We found that the beams that connect two different
groups of forms always adhere to the form in the wider direction (Fig 1.), so they were counted as a part of the forms. Since the width of the beams is always 150mm, it was added to the width of the forms. The adjusted types of the modular forms are shown in Table 1.

Table 1. The types of the modular forms (12 in total)

<table>
<thead>
<tr>
<th>Height(mm)</th>
<th>Width(mm)</th>
<th>Area(m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1050</td>
<td>0.315</td>
</tr>
<tr>
<td>300</td>
<td>1200</td>
<td>0.36</td>
</tr>
<tr>
<td>300</td>
<td>1350</td>
<td>0.405</td>
</tr>
<tr>
<td>400</td>
<td>1050</td>
<td>0.42</td>
</tr>
<tr>
<td>400</td>
<td>1200</td>
<td>0.48</td>
</tr>
<tr>
<td>400</td>
<td>1350</td>
<td>0.54</td>
</tr>
<tr>
<td>450</td>
<td>1050</td>
<td>0.4725</td>
</tr>
<tr>
<td>450</td>
<td>1200</td>
<td>0.54</td>
</tr>
<tr>
<td>450</td>
<td>1350</td>
<td>0.6075</td>
</tr>
<tr>
<td>600</td>
<td>1050</td>
<td>0.63</td>
</tr>
<tr>
<td>600</td>
<td>1200</td>
<td>0.72</td>
</tr>
<tr>
<td>600</td>
<td>1350</td>
<td>0.81</td>
</tr>
</tbody>
</table>

The formwork layout planning optimization process is as follows. First, input (1) the standard size of forms and their cost, (2) initial position of the columns as x, y coordinates, (3) starting point and direction of arrangement. Second, arrange as many modular forms (600×1350) to the floor plan as possible. If there are any forms which are interfered with by columns, delete them and set the area as an optimization area. Third, grids are set to every optimization region at 50mm intervals, considering the minimum unit of standard aluminium forms. After measuring the width and the height of the optimization area, the model will search for optimal solutions. The procedure shown in Fig 2., and the concept of the optimization is shown in Fig 3. respectively.

Figure 3. Schematic concept of movement of columns

A genetic algorithm is used in this process to obtain near optimal solutions. The arrangement of forms and positioning of the columns is determined according to the gene sequences. The x axis gene sequence variable is $a_i$, and the y axis is $b_j$. The column layout variables $x, y, s$ (the coordinates and the sizes of columns) are defined. For example, the normal width size of forms, 1350mm, 1200mm and 1050mm can be allotted to the x axis array, and the height size of the forms, 300mm, 400mm, 450mm, and 600mm can be allotted to the y axis array independently. In the column layout array, any kind of $(x, y)$ coordinate in the optimization area and the column’s size can be allotted within 50mm. The gene sequences are shown in Fig 4.

$x$ axis array: $a_1, a_2, ..., a_k, ..., a_n$
$y$ axis array: $b_1, b_2, ..., b_k, ..., b_n$
Column layout: $x_1, y_1, s_1, x_2, y_2, s_2, ..., x_k, y_k, s_k, x_{n-1}, y_{n-1}, s_{n-1}

Figure 4. Gene sequences for formwork layout

Once the three gene arrays are allotted with numbers, a layout alternative is made. Fig 5, 6, 7.

$F(a,b) = \begin{bmatrix}
F(a_1, b_1) & F(a_2, b_1) & \cdots & F(a_k, b_1) & \cdots & F(a_n, b_1) \\
F(a_1, b_2) & \cdots & \cdots & \cdots \\
\vdots & \vdots & \cdots & \vdots \\
F(a_1, b_n) & \cdots & \cdots & F(a_n, b_n)
\end{bmatrix}$

Figure 5. Formwork layout alternatives
In this model, the objective function is represented by cost, the total formwork cost. By analysing the formwork layout matrix and the layout of the columns, the total cost can easily be calculated. Therefore, the total cost is a function of $a$, $b$, $x$, $y$, and $s$, as follows:

$$Total\ cost = \sum f(a_i, b_i, x_i, y_i, s_i)$$

(1)

The model continuously changes these three sequences with a genetic algorithm, and compares the objective function until no better alternatives can be deduced.

All of formwork layout alternatives are deduced only using modular forms, and the forms are only regarded as non-modular when they are interfered with by obstacles. Arrangement of forms and columns is carried out by filling the gene array. In this process, the larger forms should be arranged first to reduce the total number of forms, and the summation of the total height of the forms should be perfectly fit to the height of the optimization area (same as the width).

In addition, the model user must deliberate the optimization boundary of columns and set the distance limitations, to secure safety and maintain the original design. Even if an alternative was verified as the most efficient result, it is possible that it may not be used because of safety or other factors such as design or law.

The cost of the forms are calculated with the type of the forms (modular or non-modular), and the covered area. If any forms encounter columns, the forms are estimated as non-modular forms.

4 Case study

4.1 Case Description

The model for formwork layout planning was applied to an actual project to verify its applicability and demonstrate the benefits. The project was a 50-story building, for which the floors are very irregular.

The computed layouts were compared with the layout obtained with the heuristic approach from a professional who had not only worked in formwork planning for 10 years, but had also participated in designing the aluminium forms. By comparing the modular form-covered area, the cost was quantitatively analysed. The detailed floor plan of the project is shown in Fig. 8

As can be seen, the shape is very irregular, with unsymmetrical column layouts.

4.2 Result and discussion

The layout solutions from the model and the heuristic approach are shown in Fig 9. As can be seen, the solution generated by the proposed model used more modular forms than the heuristic solution.

In this analysis, we assumed the rental cost of the aluminium form to be US $3 per meter square, and the cost of non-modular forms to be 2 times more expensive than the modular forms. The total area of the floor was originally $741.36\ m^2$. The 50th story was analysed. The
detailed differences are shown in Table 2.

An accurate rental cost cannot be estimated, since it is determined based on the number of reuses, which could continuously be changed according to the floor plan and construction time.

<table>
<thead>
<tr>
<th>Items</th>
<th>(a) by an expert</th>
<th>(b) by the model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-modular form covered area ($m^2$)</td>
<td>105.157</td>
<td>17.188</td>
</tr>
<tr>
<td>Rental cost of slab formwork (US$)</td>
<td>126,977.6</td>
<td>113,782.2</td>
</tr>
<tr>
<td>Total time for optimization</td>
<td>4 hour 30 min</td>
<td>30 min</td>
</tr>
</tbody>
</table>

From the results, about 88 $m^2$ (11.9%) of the originally planned non-modular forms, comprised of at least 105 units, could be changed to modular forms. This reduced the rental cost by almost $10,195 US. If the model is applied to bigger and higher buildings, the difference will increase. Also, the total time for optimization was reduced from 4 hours, by the expert, to only 30 min, by the model. As can be seen in Fig 9 (a), non-modular forms are especially required around the columns, near the walls, and in the transition zone of the floor. On the other hand, it can be seen in Fig 9 (b) that the black marked area, non-modular forms, was considerably reduced. In addition, the array of the forms looks far more simple and easy to manage, consequently enabling minimization of the formwork cost.

In this study, we suggested a formwork layout model which can be used in the structural design stage to reduce formwork cost. The model assumed that columns can be moved within limited boundaries; however, this may sometimes be a critical issue to discuss, since columns are one of the main structural members and design factors, and their placement is also related with safety. So, consideration of safety should be made, depending on the circumstances.

This study has a major role in the cost consideration of formwork layout planning. A structural engineer who is aware of potential formwork costs has to consider the advantages of structural changes, and function at reduced costs. This advanced model can bring many benefits to owners, engineers and contractors.

**Conclusion**

We have proposed a new formwork layout planning model which is suitable for the irregularly shaped buildings. The movement of columns within a limited region can be a great solution to reduce formwork cost without any losses. This model can be used in the structural planning stage by a structural engineer to reduce the cost of formwork, and increase the efficiency of the column layout. In addition, the computerized model can reduce the time required for layout planning, and enables flexibility in response even if there are any changes to the plan.

Through further studies, the consideration of changing the column design and the space distribution problem will be solved. In addition, another optimization algorithm will be studied to enhance the efficiency as well as to reduce the optimization time.
Acknowledgment

This research was supported by a grant (Code#09 R&D A01) from the High-Tech Urban Development Program, funded by the Ministry of Land, Infrastructure and Transport Affairs of Korea. Special thanks to Mr. Seung-Bum Lim, who is an expert engineer in formwork design.

Reference


The Case for BIM Uptake among Small Construction Contracting Businesses

P. Forsythe

Abstract - Advocates of BIM have the ambition of making it a central tool for information management and collaboration among stakeholders on construction projects. Even so, its implementation is often difficult for small construction contracting businesses, albeit that they make up such a large proportion of the construction industry. This paper uses data from a targeted workshop of 47 industry professionals concerning issues in transitioning from traditional to BIM based construction project management and the impact of BIM on professional training and development. The findings support the proposition that BIM uptake in construction contracting is relatively low and that small contracting businesses currently constitute the weakest link in BIM based supply chains. In order to realise the full potential of BIM, there is a greater need to make it accessible to small businesses but forced uptake is not recommended. Practical difficulties include software application costs, the need for multi-tasking, risk management issues, and confusion over the type of skilled people and training needed for operating BIM software. A “touch the BIM lightly” approach is advocated as is the development of stylised BIM applications targeted to better suit the capabilities of small construction contracting businesses, as distinct from design authoring businesses. Greater attention should be placed on low learning time, low cost, site-tasks and one-way information flows. All such recommendations are consistent with implementation using smartphone and tablet technology.

Keywords - BIM; small business, construction contracting

1 Introduction

BIM can be defined as the “modelling technology and associated processes to produce, communicate, and analyse building models” [1 P.16]. Nowadays, the literature spans broadly on this topic and examples of the following technical themes figure strongly in the literature: interoperability [2]; construction safety [3]; lean production [4]; conflict reduction [5]; sustainable design [6]; 4D construction scheduling [7]; conformance modelling [8]; 5D cost estimating[9], integrated process implementation [10] and project management [11].

Other efforts have focused more specifically on how people and organisations interact with BIM technology. Instances include BIM as a collaboration platform [12], as a procurement framework [13], as a tool in supply chains management [14] and in the multi-disciplinary expectations of participants involved in construction projects [15].

Despite the potential benefits of BIM, there is still concern about the slowness of its uptake in industry. For instance, authors such as Bew and Underwood [16] who have been instrumental in the UK implementation of BIM (see for example http://www.doi.org/doi_handbook/1_Introduction.html), identify that for an organisation to implement BIM it must be collaborative across the entire supply chain but it must concurrently make sense for businesses to be involved in such technology; it must be realistic relative to that business’s current capabilities. Here, the ideal situation is that a virtual information model will be handed from the design team through the supply chain and then to the client on completion [17]. Even so, such a scenario is only as strong as its weakest link. Unfortunately, in the Australian construction industry there seems to be relatively little BIM penetration beyond the design stage of a project and especially when moving down the supply chain where it is common to find smaller construction contracting (and subcontracting) businesses in the context of BIM uptake in Australia.

2 What Does Small business mean

Authors such as Clifford et al [18] identify that size,
structure and self governance are important aspects of describing small business. To this end, they propose that small business can be defined in terms of an employee range from 8 to 50 people but qualify this by stating that the lower end of the range maybe characterised by owner operators, the mid-range by owner managers and the upper range by owner directors. Despite the contextual relevance of the above, it would still seem that “number of employees” dominates the main key criteria used to define small businesses [19]. For instance in Australia, the Australian Bureau of Statistics defines small businesses as having 19 employees or less and defines micro businesses as having 4 employees or less [20].

With this in-mind, 47.7% of the Australian construction industry involve small businesses based on value added terms [20] and 63% of employment in the sector comes from small businesses [20]. Clearly, small businesses have a large impact on the construction industry.

In probing further, approximately two-thirds (67.0%) of these workers exist within construction services including the likes of bricklaying, plumbing, concreting, roofing services, structural steel erection, electrical services, air-conditioning services, fire and security alarm installation services, plastering, carpentry services and glazing [20-21].

In this area, small businesses represent a huge 97.8% of the construction services sector [20]. This is not a feature limited to Australia, as similar situations exist in the likes of the UK[22], the United States [23] and Asia as well [24].

In analysing the literature specific to the implementation of BIM in small businesses, most of it relates to the slightly broader category of small to medium enterprises (SMEs). Here, Gledson et al’s [25] study of middle to senior level managers confirms that significant differences exist between the large versus SME scale businesses. One aspect they identify is whether or not the proposed savings from BIM will be passed down the supply chain. Another concerns the degree of multitasking among SME staff relative to more dedicated functional roles in larger businesses – hence limiting the ability to provide dedicated BIM staff in SMEs. More generally they cite cultural, legal and commercial barriers as features impeding BIM uptake in SMEs relative to larger businesses – a view supported by other authors such as Yan and Damien [26] on cultural barriers, Sebastian [17] on commercial barriers, Greenwood et al [27] on legal barriers, and Takim et al [28] on software licensing costs. Oluwolé [29] goes further on this last issue by stating that software costs can account for as much as 55 percent of total BIM implementation costs.

Others delve deeper into the specifics of people oriented issues within construction related businesses. For instance, Arayici and Coates [30] and Arayici et al [31] found that BIM implementation was impeded in design practices by a lack of operational skills, staff training and the additional layer of complexity posed in understanding BIM specific protocols and standards. Anderson et al. [32] found that BIM is perceived as being difficult to access, may not be trustworthy and may not necessarily be neutral insofar as information being created and controlled by others. Harris et al’s [33] large scale study of SMEs found that incremental levels of innovation were possible where there was an emphasis on people or organic aspects of BIM and TQM, as distinct from more mechanistic approaches to implementation. Even so, there was still no clear link between such technologies and radical levels of innovation.

Given the above, it is not surprising that Porwal and Hewage [13] assert that organizational and people centred issues pose the greatest challenge for BIM implementation and that contractors must be integrated as early as possible in the design phase for BIM to have a strong impact on project outcomes.

What can be gleaned from the above is simply that the technology intensive nature of BIM appears to be particularly difficult in terms of people issues and this is exacerbated further when taken in the context of small businesses. As raised earlier in the paper, this research focuses upon gaining a better understanding of the extent to which such issues impact on small construction contracting businesses in Australia. This aims to underpin and direct discussion about the best way forward in assisting stronger engagement concerning BIM uptake (across the entire supply chain).

3 Research method

In addressing the above, a large industry workshop on the implications of BIM on future professional needs was undertaken as a mechanism for feedback and consensus concerning BIM uptake. Whilst the workshop covered a variety of issues, only those implicating the perspective of small construction contracting businesses (including subcontracting) have been reported in the findings.

In total, 47 middle to executive level managers participated in the workshop with representation predominately consisting of architectural firms, specialist design consultants, quantity surveyors, head contractors and subcontractors.

The structure of the workshop began with formal yet brief presentations from a small group of speakers covering cogent topics of relevance to industry uptake. This was used to set the tone for debate during the ensuing breakout sessions which focused on
transitioning from traditional to digitally based construction project management, and posing key questions such as the most important issues arising from BIM in practice, and the associated impacts of these changes on professional training and development.

In executing the breakout sessions, participants were directed into special interest subgroups covering different disciplinary roles such as construction technology, structures, services, procurement, design management, time management, cost management and urban design. Each group debated the abovementioned issues and then reported back their key findings and conclusions. This approach provided a direct account of the plight of construction contractors in BIM uptake, and also provided the way that others in the supply chain viewed the uptake of BIM by such contractors.

The discussion arising from the breakout sessions was transcribed and then a thematic analysis was undertaken to derive the main perceptions and ongoing issues of relevance. Here, thematic analysis serves to process raw qualitative data into categories of thematic meaning. It is similar to content analysis but the focus is on categorising theme frequency rather than word frequency. Promotion of the deeper level of meaning contained in themes (as opposed to words) has been acknowledged as far back as Cicourel [34] and its more modern implementation is elaborated upon in detail by qualitative researchers such as Boyatzis [35].

The findings are discussed in terms of the main themes identified from the analysis and are supported by indicative quotes from different participants made during the breakout sessions (as shown in italics).

4 Analysis and findings

The findings arising from the study extend and add further context to themes reported previously in the literature review. At an over-arching level, it was apparent that contracting businesses in general were perceived as having relatively low uptake of BIM on building construction projects. This especially appears to be the case where involving small subcontract businesses e.g. “A lot of them say, what is BIM, what does that mean?” and, “That’s our problem, as a constructor, there’s no subbies (sub-contractors) engaged in this sort of stuff. We need to get them engaged. Sure there’s a whole heap of architects here and probably some structural guys, but as a constructor we are finding that the subbies are very much behind the eight ball”.

These and related issues lead to the perception that construction contractors have a tendency to inflate their tender prices to undertake BIM based construction work e.g. “Now some of the contractors actually inflate their tender because of BIM. They’re saying, if you want a BIM model, we will charge you $100K more, or something like that”.

Consistent with the above, it was perceived that there was a gap in knowledge that effects subcontractors more than most e.g. “So collaborations with subcontractors and narrowing the gap of knowledge, or working out ways to help companies that might not be at the required level - how do we bring those people up to speed?”

In delving deeper into the day-to-day operations of construction contractors’ onsite, there was the view that spending large amounts of time and resources interrogating a BIM to extract construction specific information was not an option for them. They felt that they did not have the skills to interact with the BIM; dealing with the BIM was an extra cost to production; the BIM is often large and unwieldy in terms of the information that subcontractors need to extract from it e.g. “If you’re down the pecking order at a subcontractor level, you don’t necessarily want to see, or you want to very quickly drill down to the bits (of information) you need to know.”.

A similar view was apparent in terms of BIM usage in work flow scheduling. Here, the dominance of the traditional Gantt chart was not underestimated and reflects the level of technology that is commonly used by construction contractors, as captured below:

“Unfortunately, the Gantt chart has been run for half a century and we’re still using it.

We don’t want something high-tech and we don’t need something more complicated than the critical path method.

(With reference to BIM) Not because people don’t want to change but because people don’t want to change for complication. We want efficiency. We want something simple on the construction site. So, hopefully, BIM as the new technology could provide the solution.

Despite the above, some have had a degree of success by stylising interaction with the BIM (mainly the 3D view of the model) in a way that was simplified and well tailored to meet onsite needs using the likes of laptop technology e.g. “They want a laptop of the model, with a viewer onsite. They’ve done that and I think it’s been giving them huge benefits, in terms of onsite understanding.”

From the above, it would seem that in order to optimise BIM across the supply chain, it may be worth re-thinking specific aspects of implementation to suit organisational size and capability. Here attention must
be given to small construction subcontracting organisations and targeting a specific level of involvement for them that will be both viable and will allow the efficiency of the supply chain to improve overall.

4.1 BIM Advocacy and the impact on construction contractors

It was apparent from the breakout groups that BIM advocacy was stronger among design professionals than construction contractors. Clearly, the sophistication of BIM usage among design professionals (e.g., architects, structural engineers, mechanical engineers) was linked to their long term affinity with digitally based tools for communicating and documenting design information on projects. It was also evident that this advocacy would inevitably impact on construction contractors – as captured in the following quotations:

I think we're going to see a marked division in the industry of the people who have adopted BIM processes and the ones that have not.

I'd really like for it to be (mandatory) because architects have technical capabilities.

Giving people no choice (about the use of BIM) is a great way of getting them into this process.

I think it all comes back to collaborating again and unfortunately we're getting a complete lack of models from certain parts of our industry. We would also love to see the contractors getting involved much earlier.

A contractor may not be ready for construction sequencing using BIM or location-based analysis, even cost estimation. But everybody can see and understand clash detection and it just saves money very, very directly.

I think one of the important issues surrounding people issues is that very soon you are finding BIM model managers taking over the role as your traditional contract managers, and perhaps rightfully so, because they are the key holders to all the important information.

A key issue arising from the above (and consistent with previous discussion) is simply that small construction contractors seem to represent the weakest link in the BIM based supply chain on projects. A question arising from this concerns what is the appropriate level of BIM uptake for such businesses including, is it best to force change, or advocate for a more mediated approach based around incremental uptake of the technology.

4.2 Fast moving technology versus trying to keep up

Unsurprisingly and adding to the previous point, a key theme from the workshop concerned the rapid evolution of BIM technologies and the ability of the supply chain to keep up. The basic tenets of 3D technology have gradually added cost, time, sustainability, facilities management and geo-spatial locating variables. Work continues to progress on improved visualisation and simulation abilities. Common themes in the breakout sessions alluded to "Increased knowledge requirements" and "Increased levels of collaboration and communication." A specific problem for small construction contractors concerned the cost and resourcing of hardware and data storage requirements e.g. This is a major issue because if you're passing around a model that is, say, half a gigabyte, even with very high bandwidth communications, you can't pass that around very much and just archiving that sort of information is incredibly difficult.

It was therefore apparent that the higher the level of information technology, the higher the level of business resourcing needs (including hardware, software and technical expertise). In adding to this, it was evident that small construction contractors needed to learn and understand project BIM specific platforms and data compliance standards in order to collaborate and check the consistency of data, thus creating further training and resourcing needs.

Here, it is relevant to point out that many small construction contractors do not see their primary role as designing or facilitating information flows but rather they are typically receivers and adaptors of information for the purpose of constructing things; they are often towards the end of the information chain and subsequently information often has a one way directional flow. As such, they have less express interest in the big picture of BIM; a smaller perspective of what they will get from it; and a smaller budget to commit to it. In real terms, BIM may confront their preferred business model of maintaining low operating overheads to remain competitive. BIM tends to contest this premise as it is not easily scalable at present. Instead, it potentially creates a new layer of expertise within small contracting businesses. Further, BIM may only be used on a limited number of the business’s projects and so dedicated BIM staff may be underutilised at a broader organisational level. Of course there is also the likely problem that this layer of BIM expertise may reside within a single person which means that operations become pivotal around a new but relatively unknown area of expertise, thus creating
unwanted operational risks. As mentioned by some during the breakout sessions, it is also unclear who is the right person (people) to hire in terms of the expertise that will most benefit the business’s operations e.g.

The most common story I get is builders say, how are we going to do this? Answer, I know, we’ll go and find someone who’s an expert in [design authoring software] – apologies. They’re not the right people, they’re not builders. They don’t actually know how to do it(...) but they drive a solution that’s in the architectural world. They’re not driving a solution that’s in the construction management world.

4.3 Risks divested down the supply chain

A feature of discussion from the workshop concerned the accuracy of BIM information provided down the supply chain, to those involved in construction contracting. Comments suggest that designs often lack continuity and appropriate detail for construction. Many instances of this were put forward including the following:

Architects don’t currently have continuity between their sketch design, detail design, construction documentation. If architects can’t get continuity within their own profession and their own office, how can they get continuity anywhere else.

So if you have incomplete models, models with missing information and a missing level of detail, I think that causes problems. If a model is done correctly, clearly it helps mitigate risk. If it is the other way around, then you probably enhance the chances of encountering a problem.

Certainly I think people on the table are very critical and think that it’s an important issue about completeness of the model and who is accountable for if anything goes wrong and it’s implication on the various contractual issues that are on the table.

If you’re relying on models from architects, they’re probably not going to be done at a suitable standard and there needs to be a lot of work at that front end around integrating - which gets us onto integrated project delivery.

For the understanding or for the people that are working inside these processes, they must have a basic understanding inside the usage of BIM, how processes are constructed, how processes are developed on the construction site and even how progress is reported in the construction industry.

The main point from these quotations is that those undertaking the design process do not necessarily have a full understanding of the ramifications concerning how their information will be used further down the supply chain i.e. for trade level construction purposes. The previously mentioned point about lack of confidence in the accuracy of information may mean that contractors are ultimately not prepared to make serious usage of it.

It is also unclear who is responsible for inaccurate information. The lack of fully detailed design documentation is certainly not new but it does raise the obvious issue that a contractor is unlikely to place confidence in a BIM that is not formally linked to the contract and is suspected of having inconsistent or inaccurate information. It really only leaves contractors with the choice of doing their own exploration, investigation and checking of data. Even so, the previously discussed problems of resourcing, standards protocols and training, creates an obvious disincentive for this to occur in real terms. It also tends to mean that subcontractors simply place stronger faith in traditional 2-D documentation, which they are more used to searching for mistakes and emissions.

5 Conclusion

The findings from the study are consistent with other work in the area but add context and qualifying statements that assist understanding. For instance, whilst BIM has a degree of uptake among design professionals on Australian construction projects, there currently appears to be relatively low uptake among general construction contractors beyond design management and (some) program planning activities. The level of uptake appears even lower for small subcontracting businesses being the main target of interest arising from this paper. They seem to represent the weakest link in the supply chain concerning BIM uptake.

A general theme from the literature review – as captured by Bew and Underwood [16] – was that for BIM implementation to be successful, it must be collaborative across the entire supply chain but it must concurrently make sense for businesses to be involved in such technology; it must be realistic relative to that business’s current capabilities. The findings from this study suggest that BIM is not particularly realistic for small subcontract businesses in Australia at this point in time, thus preventing full realisation of the collaborative benefits that BIM potentially offers.

The question for small construction contractors is therefore at what rate should they get involved, and what is it worth to them? There is currently a basic problem between the economies of scale presented by BIM and the apparent disconnect for small subcontract
enterprises in fitting in with this scale of operation. For instance, BIM aims to offer a decreasing cost per unit of information output but this does not seem to necessarily convert to reduced costs for small subcontract businesses. Instead, BIM appears to require significant resourcing including big learning curves, big files, big hardware, big human resourcing commitments, big design management involvement and big ITC involvement. This level of “big” commitment is essentially counter to the basic premise of small enterprises whose modus operandi generally revolves around low overheads and low investment. In the short term, the inability of small businesses to absorb higher fixed costs within their operating overheads means that BIM will likely come at an additional cost to project specific budgets – a factor that is not necessarily seen as a value-add in terms of tangible outputs. A key point here is simply that unlike architects, consultants and project managers who to some extent make fees by creating and managing information flows, trade based sub-contractors ostensibly make income and add value via the tangible outputs they deliver.

A further conclusion from the research is the general push by some to mandate and force BIM usage along the entire supply chain. Some may not have full understanding of the capacity of the supply chain to actually undertake this in a way that is genuinely beneficial to the industry. For instance, correctly set regulation can help lead BIM progress but incorrectly set regulation may only serve to create dysfunctional market mechanisms. With this in mind, trying to force training and technical advancement is one way of attempting to progress small businesses but not the only way and not necessarily the easiest way. In the worst cases scenario it may only promote competency in the likes of large BIM authoring software which ultimately has little relevance to the needs of small trade based sub-contractors – they have greater interest in simply manipulating, detailing or extracting already authored design information.

Where BIM would benefit small subcontract businesses is in software applications that directly assist productivity onsite during the physical execution of the work. To name but a few instances, they require construction accurate information relating to assembly details, management of site quality control, checking of orders, making orders, sorting materials deliveries, marking where materials should go, recording completed work, obtaining site instructions and dealing dynamically with site identified safety issues.

It is therefore worth re-thinking or at least adapting selective BIM implementation to suit the specific needs of small contracting businesses. Here, greater attention should be placed on the scalability of BIM applications and consideration of one way information flows. For instance, rather than using high end, expensive and complicated software applications, the focus should move towards smaller, simpler and trade package specific applications. Consequently, a “touch the BIM lightly” approach should prevail. This should target low levels of BIM proficiency, based around extraction of information and relatively limited return information to the master BIM. In many ways, such an approach would be similar to many smartphone or tablet “apps” which typically exhibit:

- free or very low cost,
- quick, simple and intuitive learning
- an emphasis on user needs and how information will be used
- a focus on doing relatively small tasks well
- (Occasionally) providing output files that can be transferred to more comprehensive software.

Of note, this would go some way to addressing Oluwole’s point [29] that software costs can account for as much as 55 percent of total BIM implementation costs. Further, it would circumvent related resourcing requirements such as software training costs and the need for software specialists. Such an approach pushes the virtues of a simplified approach to BIM implementation for targeted users.

In a sense, this would support the democratisation of BIM (by making it available to a greater cross section of the supply chain) and would concurrently reduce the potential onset of learned helplessness among those who to date, have seen it as being too difficult to implement.

Despite the potential benefits of such an approach, it would also seem that those higher up in the supply chain must commit more fully to the accuracy of information passed down the chain in order to realistically provide improved efficiency and the mitigation of risk to small contracting businesses.

References


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

ENERGY AND ENVIRONMENT
Energy-Efficient Air-Cooled DX Air-Conditioning Systems with Liquid Pressure Amplification

V. Vakiloroaya and Q.P. Ha

School of Electrical, Mechanical and Mechatronic Systems, University of Technology, Sydney, Australia
E-mail: vahid.vakiloroaya@enginee.com, quang.ha@uts.edu.au

Abstract -
The objective of this study is to explore an optimal strategy on energy consumption for a direct expansion (DX) air-conditioning system by using a refrigerant pump in the liquid line to allow the system to operate at a lower condensing pressure. An existing DX rooftop package of a commercial building located in a hot and dry climate zone is used for data collection. The theoretical-empirical modelling approach is used to obtain system model, from which the proposed strategy is formulated. A numerical algorithm is developed to analyse the system transient performance, using an iterative loop. As a minimum pressure differential is required across the expansion device, liquid pressure amplification (LPA) devices can be used on DX systems that operate with fixed head pressure control. They can be fitted to new or existing systems. Results show that the LPA approach is more effective when the ambient temperature is falling, with electricity saving around 25.3% in average.

Keywords -
Energy Saving; Experimental Study; HVAC; Modelling; Liquid Pressure Amplification

1 Introduction

Nowadays, heating, ventilation and air conditioning (HVAC) systems typically account for around 40% of total electricity consumption of buildings [1], attributed mainly for creating peak electricity demand. In recent years, hotter weather has meant the demand for peak energy growing much faster than for the base power. As a result, finding new ways to reduce energy consumption in buildings without compromising comfort and indoor air quality is an ongoing research challenge. One proven way of achieving energy efficiency in the vapour compression refrigeration systems is to reduce the compressor discharge pressure [2,3], which decreases its compression ratio, and in turn, causing to less electricity consumption.

Many studies report that reduced compressor discharge pressures and the corresponding reduction of compression ratios in a refrigeration cycle are advantageous when it comes to HVAC energy saving and increasing its service time. There are, however, some solutions to significantly lower compressor discharge pressures whilst preventing the flash vapour occurrence. For example, pre-cooling the ambient air before it reaches the air-cooled condenser of the vapour compression cooling systems can reduce the condensing temperatures to drop the condensing pressure. Yu and Chen [4] investigated how the coefficient of performance (COP) of air-cooled chillers can be improved by using mist pre-cooling. They estimated around 18% decrease in the annual electricity usage could be achieved with mist pre-cooling of air entering the air-cooled condenser of chiller, used to serve a hotel in a sub-tropical climate. The application of LPA to air-cooled air conditioning systems can assist in achieving a considerable reduction in compressor discharge pressure. LPA can be realised with a hermetically-sealed, magnetically-driven liquid refrigerant pump installed in the liquid line between the condenser and expansion valve. The LPA pump increases the pressure of the liquid refrigerant before it enters the expansion valve. This method allows the condensing temperature to fluctuate with ambient temperature changes, and hence, reducing the condensing pressure and lowering overall energy consumption.

The compressor is the largest power consumer in a vapour compression system. As a result, many previous studies have investigated the influence of various technologies on the compressor performance to enhance the operating and energy efficiency of vapour compression refrigeration systems [5]. However, compared with a water-cooled air conditioning system, air-cooled cycles are less energy-efficient. Wang et al. [6] studied the impact of two performance improvement techniques applied to a compressor with different refrigerants. The first technique involved cooling the compressor motor via external means, while the second used isothermal compression achieved by transferring heat from the compressor chamber. Their results
showed that these approaches could reduce the power consumption of the compressor by up to 16% for external cooling and 14% for the isothermal technique. In [7], the system performance of an R410A inverter air-conditioner could be improved using an evaporative-cooled condenser, increasing the tested system’s COP by 18.32%. A comparative study of heat recovery and floating condensing temperature techniques was carried out by Arias and Lundqvist [8] to show that with a condensing temperature of 40°C, a system using the floating condensing temperature technique could achieve a 50% improvement in energy savings when compared to a conventional refrigeration system.

The objective of this study is to explore the influence of an optimal strategy on energy consumption of an existing direct expansion air conditioning system. For this purpose, an actual air-cooled rooftop package air conditioning system of a real-world commercial building located in a hot and dry climate zone in the northern hemisphere was used for experimentation and data collection. Field tests were conducted to quantify and determine the system variables. Mathematical models were obtained by using a theoretical-empirical approach, and the proposed strategy was formulated on the basis of these models. A transient simulation software package, TRNSYS 16 [9] was used to predict the HVAC energy usage. From the TRNSYS codes and real-world test data, a simulation module for the cooling plant was developed and embedded in the software. Performance predictions were then compared with actual performance measurements to verify the models and to compare the strategy performance. Findings reported in this paper show that energy saving of 42% is possible using the proposed method.

2 System modelling

Many models for DX rooftop package air conditioning system have been developed using various principles. In the following, a combined theoretical-empirical approach will be developed for component-wise modelling. A single stage vapour compression direct expansion air conditioning system consists of four major components, namely a rotary scroll compressor, an air-cooled condenser, an expansion valve and a DX evaporator. Figure 1 shows a schematic block diagram of the conventional and developed DX air conditioning system while their pressure-enthalpy diagram is depicted in Fig. 2. In a conventional system, the cycle starts with a mixture of liquid and vapour refrigerant entering the evaporator (point 1). In the following, notations used in the analysis are given in the nomenclature following Section 5. The DX evaporator used in the plant is of a rectangular finned tube type of heat exchanger, in which both the refrigerant and air are assumed to be counter-flow. When heat from warm air is absorbed by a DX evaporator coil, the refrigerant is changed from liquid to gas and becomes superheated at the evaporator exit. Superheat is required to prevent slugs of the liquid refrigerant from reaching the compressor to cause any serious damage. The supply air temperature can be estimated by:

$$T_{sup} = T_{eva,a,i} - \frac{\dot{m}_{r}}{\dot{m}_{eva,a}C_{p,a}}(h_2 - h_1), \quad (1)$$

in which the enthalpy of the refrigerant leaving the DX evaporator, $h_2$, is determined as:

$$h_2 = h_{eva,r,sat} + C_{p,r}(T_{eva,r,sh} - T_{eva,r,sat}). \quad (2)$$

where $T_{eva,r,sh}$ is the temperature of the superheated vapour refrigerant leaving the evaporator to be measured during the experimentation. The input power
of the evaporator fan, which is adjusted for part load operating conditions, is proposed as a function of the refrigerant mass flow rate, building cooling load and supply air temperature as:

\[
P_{\text{evax, fan}} = a_0 + a_1 m_r + a_2 m_r^2 + a_3 T_{\text{sup}} + a_4 T_{\text{sup}}^2 + a_5 Q_b + a_6 Q_b^2 + a_7 m_r T_{\text{sup}} + a_8 m_r Q_b + a_9 T_{\text{sup}} Q_b,
\]  

(3)

where coefficients \(a_0\) to \(a_9\) are constant to be determined by curve-fitting of experimental data.

The superheat vapour then enters the compressor, at point 2, where an increasing pressure will in turn increase the temperature. In this study, a steady-state polytropic compression is considered, assuming the compressor speed reaches its specified speed instantaneously. The compressor mass flow rate is given by:

\[
m_r = \frac{V_f h_f}{v_{\text{sup}}},
\]  

(4)

where \(\eta_c\) is the volumetric efficiency of the compressor and can be obtained by:

\[
\eta_c = 0.92 - 0.08 \times \left[ \frac{1}{\left( \frac{p_{\text{dis}}}{p_{\text{sup}}} \right)^k} - 1 \right].
\]  

(5)

The work of compression is obtained by:

\[
W_{\text{in}} = \frac{m_r}{\eta_{\text{comp}}} P_{\text{sup}} v_{\text{sup}} \left( \frac{k}{k-1} \right) \left( \frac{p_{\text{dis}}}{p_{\text{sup}}} \right)^{\frac{k-1}{k}} - 1.
\]  

(6)

The empirical expression to determine the compressor power consumption \(P_{\text{comp}}\) is proposed as:

\[
P_{\text{comp}} = b_0 + b_1 T_{\text{sup}} + b_2 T_{\text{sup}}^2 + b_3 T_{\text{sup}} T_{\text{dis}} + b_4 T_{\text{sup}}^2 T_{\text{dis}} + b_5 T_{\text{dis}}^2 + b_6 T_{\text{sup}} T_{\text{dis}} + b_7 T_{\text{sup}}^2 T_{\text{dis}} + b_8 T_{\text{sup}} T_{\text{dis}}^2 + b_9 T_{\text{dis}}^3,
\]  

(7)

where coefficients \(b_0\) to \(b_9\) are constant, obtained by curve-fitting of the collected data.

The refrigerant enthalpy leaving the compressor is calculated as:

\[
h_3 = \frac{m_r h_2 + W_{\text{in}}}{m_r}.
\]  

(8)

The required heat rejection capacity of the condenser is obtained from:

\[
Q_{\text{con}} = AU \Delta T_m,
\]  

(9)

where \(\Delta T_m\) is the mean temperature difference given by:

\[
\Delta T_m = \frac{T_{\text{con, a, o}} - T_{\text{amb}}}{\ln \left( \frac{T_{\text{cond}} - T_{\text{amb}}}{T_{\text{cond}} - T_{\text{con, a, o}}} \right)},
\]  

(10)

and

\[
\frac{1}{AU} = \frac{1}{\eta_{\alpha_d} A_0} + \frac{1}{2} \ln \left( \frac{d_0}{d_f} \right) \frac{d_0}{A_0 k_t} + \frac{1}{\alpha_f A_0},
\]  

(11)

in which \(A_0\) is the total heat transfer area, i.e.:

\[
A_0 = A_f + A_b,
\]  

(12)

and \(A_f\) and \(A_b\) are areas respectively of the fin and bare tube:

\[
A_f = \left( \frac{D - t_f}{D \times VTS} \right) \frac{\pi d_f}{4} N_{\text{row}},
\]  

(13)

\[
A_f = \left( \frac{2}{D} \left( HTS - \frac{\pi d_f^2}{4 \times VTS} \right) \right) \frac{\pi d_f}{4} N_{\text{row}}.
\]  

(14)

The surface effectiveness is given by:

\[
\eta_o = 1 - \frac{A_f}{A_0} \left( 1 - \eta_f \right),
\]  

(15)

in which the fin efficiency \(\eta_f\) can be determined by approximation as described in Schmidt [10]:

\[
\eta_f = \frac{\tanh(mr \phi)}{mr \phi},
\]  

(16)

where

\[
m = \sqrt{\frac{2 \alpha_a}{k_f t_f}},
\]  

(17)

\[
\phi = \frac{R_{\text{eq}}}{r} \left[ 1 + 0.35 \ln \left( \frac{R_{\text{eq}}}{r} \right) \right].
\]  

(18)
\[
Re_q \frac{r}{r} = 0.635 \left( \frac{VTS}{r} \right) - 0.3 \right)^{0.5}, \quad (19)
\]

The refrigerant side heat transfer coefficient can be calculated in two cases:

- For single-phase section [11]:

for \( Re_d < 3500 \),

\[
Nu = 1.11 (Re_d)^{0.21} \Pr^{1/3} = \frac{\alpha_c d_i}{k_r}
\]

for \( 3500 \leq Re_d \leq 6000 \),

\[
Nu = 3.52 \times 10^{-7} (Re_d)^{2.03} \Pr^{1/3} = \frac{\alpha_c d_i}{k_r}
\]

for \( Re_d > 6000 \),

\[
Nu = 0.22 (Re_d)^{0.615} \Pr^{1/3} = \frac{\alpha_c d_i}{k_r}
\]

- For two-phase section [12]:

\[
\alpha_{fp} = \alpha_l \left[ (1 - x)^{0.8} + 3.8x^{0.76} (1 - x)^{0.04} \right] \Pr^{-0.38}, \quad (21)
\]

where \( \alpha_l \) is the heat transfer coefficient of liquid refrigerant calculated from the Dittus and Boelter correlation [13]:

\[
Nu = 0.023 Re_d^{0.8} \Pr^{0.3} \quad (22)
\]

The heat transfer coefficient on the air side is expressed as:

\[
\alpha_a = \frac{j G_{a, max} C_{p,a}}{Pr^{2/3}} \quad (23)
\]

where the Colburn factor \( j \) can be determined from the correlation of McQuiston and Parker [14].

The pressure drop on the air side of the condenser coil is given as [15]:

\[
\Delta p_{con,a} = \frac{G_a^2}{2 \rho_a} \left[ f_a \left( \frac{A_0}{A_{min}} \right) \rho_l + \frac{1 + \sigma^2}{\rho_l} \left( \frac{\rho_l}{\rho_o} - 1 \right) \right], \quad (24)
\]

in which the air side friction factor is calculated as:

\[
f_a = 0.0267 Re_c F_1 \left( \frac{VST}{HTS} \right)^{F_2} \left( \frac{P_l}{d_c} \right)^{F_3}, \quad (25)
\]

where

\[
F_1 = -0.764 + 0.739 \left( \frac{VST}{HTS} \right) + 0.177 \left( \frac{P_l}{d_c} \right) - \frac{0.00758}{N_{row}}, \quad (26)
\]

\[
F_2 = -15.689 + 64.012 \ln(Re_c) \quad (27)
\]

\[
F_3 = 1.696 - 15.695 \ln(Re_c) \quad (28)
\]

The power consumption of the condenser fan then can be expressed as:

\[
P_{con, fan} = \frac{\Delta p_{con,a} V_{con,a}}{\eta_{con, fan}} \quad (29)
\]

In a conventional vapour compression refrigeration system, the condensing pressure is designed to allow refrigerant condensation at high ambient temperature. In this way, the energy may be wasted in partial loads when the ambient temperature is low and a high condensing temperature is not required. In this study a refrigerant pump is incorporated after the condenser’s receiver in the liquid line to preserve the high pressure differential across the expansion valve. Consequently, the condenser pressure can be adjusted with variation of the ambient temperature to result in a lower discharge pressure during the partial load period, and in turn, to reduce the compressor power consumption. Therefore, the compressor head pressure is reduced from point (3) to point (b) at steady state conditions, as shown in Fig. 2. In addition, a small amount of the liquid sub-cooled refrigerant is injected into the compressor discharge line using a secondary pump. Therefore, the temperature of the discharge refrigerant gas entering the condenser is decreased to reduce the difference between the ambient and condensing temperature, and thus, to increase the efficiency of the condenser. This is shown in Fig. 2 by moving point (b) to point (c). The superheated refrigerant then enters the air-cooled condenser, where a final reduction in the refrigerant temperature takes place and causes it to de-superheat, and thus, the refrigerant liquid is sub-cooled as it enters the expansion valve. As a result, refrigeration effect of the evaporator is increased, enabling the system to deal with a higher load demand. As illustrated in Fig. 2, this is occurred by replacing point (1) to point (a) in the refrigeration cycle.
In this study, 5\% of the total refrigerant mass flow rate is injected to the discharge line of the compressor at each iteration loop of the simulation. Therefore, the temperature of the refrigerant entering and leaving the condenser in the LPA system can be determined by:

\[
h_c = \frac{h_h + 0.05 h_{inj}}{1.05}, \tag{30}
\]
\[
h_d = h_c - \frac{m_{con,a} C_{p,a}}{m_r} (T_{cond} - T_{amb}) \tag{31}
\]

The input power of the condenser fan is proposed as a function of the refrigerant mass flow rate and ambient dry-bulb temperature as:

\[
P_{con,\text{fan}} = c_0 + c_1 m_r + c_2 m_r^2 + c_3 T_{amb} + c_4 T_{amb}^2 + c_5 m_r T_{amb}, \tag{32}
\]

where coefficients $c_0$ to $c_5$ are constant to be determined by curve-fitting of the experimental data.

The high pressure sub-cooled refrigerant flows through the expansion valve, at point (d), which serves to reduce its pressure. The expansion valve is a refrigerant flow control device that adjusts the quantity of the liquid refrigerant entering the evaporator, and thus regulates the refrigerant superheat temperature leaving the evaporator. Furthermore, by this method the refrigeration effect of the evaporator is increased, enabling the system to deal with a higher load demand. However, due to a minimum pressure difference across the expansion valve required for proper operation of the thermostatic expansion valve, the amount of the discharge pressure to be reduced is limited. Therefore, the energy saving potential here is a function of the minimum acceptable pressure differential across the expansion valve and the ambient temperature. For an isenthalpic process in the expansion valve, we have $h_d = h_1$ in conventional systems, or $h_d = h_u$ in this proposed LPA system.

3 Case Study

The proposed strategy is modelled on the existing DX rooftop package unit serving an office building as our experimental set-up. The floor area of the air-conditioned room was 325 square meters with a height of 3 meters. The building is open from 8.00 am to 10.00 pm. The walls, windows, floor and roof are modelled according to ASHRAE [16]. More details about the building materials and internal loads can be found in [17]. The weather data that drive project simulations are based on a typical meteorological year. The data file distributed with TRNSYS 16 is generated using “Meteonorm” including Synthetic hourly weather data, including information about the month of the year, hour of the month, direct normal solar radiation, global solar radiation, dry-bulb temperature, humidity ratio and wind speed.

The DX rooftop package comprises an electrostatic expansion valve, a DX evaporator, an air-cooled condenser and a scroll compressor with nominal capacity of 52.5 kW. The working fluid is refrigerant R22. The evaporative temperature at designed conditions is set to be 4.4°C. Supply air flow rate is 8748 m$^3$/h at designed conditions. The designed electric power input of the evaporator variable speed fan at maximum air flow rate is 1.3 kW. Heat rejection capacity of the air-cooled condenser is designed to control the condensing temperature at 50°C. The nominal air flow rate of the condenser variable air volume fan is 12400 m$^3$/h and its rated power input is 1.1 kW.

3.1 Experimental Set-Up

Figure 3 shows the experimental set-up. High precision sensors/transducers were used for measuring all operating variables. The temperature sensor for the supply and return air is of platinum resistance type with accuracy of ±0.1°C. The refrigerant mass flow rate passing through the compressor is measured by a Coriolis mass flow meter with accuracy ±0.1%. Four platinum resistance thermometers of PT100 type with a calibrated accuracy of ±0.5°C for a temperature range from -50°C to +260°C are used to measure the refrigerant temperature before and after of each component. These sensors are in direct contact with the refrigerant to achieve high accuracy. Two pressure transmitters with a calibrated accuracy of 1% are respectively installed in the suction line and discharge line.

Figure 3. Experimented DX rooftop package
The pressure measuring range for the suction line is from 0 to 2000 kPa and for the discharge line from 0 to 3500 kPa. The ambient temperature is monitored by a digital thermometer of accuracy ±0.8°C. Electric component powers are measured by a digital power clamp of accuracy ±3.5%. All measurement signals are acquired with a 20-minute sampling time. Therefore a total 3234 points of the system’s power consumption and other variables were collected. Data were logged at the test system using history sheets available in the system measuring tools.

3.2 Model Validation

The mathematical model of the system components is coded into the fully integrated visual interface known as TRNSYS simulation studio using FORTRAN. Under real conditions there are many factors that can influence the HVAC system performance. In this study, all dimensions of the system components and dimension of connected pipes are included in the program code. The experimentally obtained data were entered through a dedicated visual interface. The model included a subroutine to evaluate thermodynamic properties of refrigerant R22. This simulation code was then used to obtain parameters of main interest such as the evaporative and condensing temperature, condensing pressure, temperature of refrigerant leaving the compressor, temperature of refrigerant entering and leaving the condenser and liquid line temperature as well as power consumption of the compressor, evaporator variable air volume (VAV) fan, and condenser VAV fan.

According to the data provided by field tests, the corresponding coefficients of the models are obtained by regression techniques using MINITAB [18]. The R-squared value of each model indicates a good fit. Regression process yields coefficients according to variations of the overall energy consumption. These coefficients for the rooftop package used in this study are described in [17], where the integrated simulation tool was validated by comparing the predicted and measured total power consumption of the system for a week in mid-summer.

4 Results and Discussion

Having described the specification of plant and developed strategy, the simulation is run with a time interval of 20 minutes, i.e. equal to the monitoring time step in the real test process. The cooling load is calculated prior to the simulation performed with TRNSYS. The references for indoor temperature and relative humidity during the cooling load calculation were set respectively at 23°C and 50%. The peak cooling load is estimated at 36.7 kW. As mentioned before, the LPA technique allows the compressor to be run at a lower delivery pressure, with the pump providing the required stable pressure for the expansion valve and thus the condensing pressure can float with ambient temperatures.

4.1 Performance Prediction

The variation of condensing temperature and condensing pressure, discharge temperature, evaporative temperature and temperature of refrigerant in liquid line are discussed in this section. The condensing temperature profile is shown in Fig. 4, where it can be seen that the condensing temperature in a conventional system is maintained relatively constant at an average value of 43°C while in the system floating with LPA, it varies in a larger range. The reason is that the system floating with LPA can operate at a lower condensing temperature as a result of decreasing the compressor head pressure. Figure 5 shows the variation of compressor discharge pressure and discharge temperature for both the conventional and LPA system. The results show that the average refrigerant discharge pressure drops from 1750 kPa to 1170 kPa after using an LPA pump whilst the refrigerant discharge temperature is reduced from 81.5°C to 62.3°C.

Nevertheless, results show that changes in the condensing temperature do not have a considerable effect on the evaporative temperature, as shown in Fig. 6. This is because the variation of the evaporative pressure is dependent on the heat transfer rate and fairly insensitive to both the condensing temperature and pressure.

The temperature of refrigerant in the liquid line in the proposed LPA system is notably less than that for the conventional system, as shown in Fig. 7. The reason is that the sub-cooled refrigerant injected into the discharge line quickly flashes and reduces the temperature of the superheated refrigerant vapour. As indicated in Fig. 7, the average temperature of the refrigerant at the liquid line after floating the condensing pressure by LPA pump is dropped from 31°C to 17°C. The drop of the refrigerant temperature entering the condenser leads to a reduction in refrigerant temperature leaving the condenser. This means a more substantial sub-cooling at the condenser outlet and therefore a reduction of the refrigerant enthalpy entering the evaporator. This enhances the system refrigeration effect and thereby its COP. Simulation results show that the average COP of the system floating with LPA, at various evaporative temperatures, is greater than that under the commonly-used designs, as shown in Fig. 8.
The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

Figure 4. Condensing temperature

Figure 5. Discharge pressure and temperature

Figure 6. Evaporative temperature

Figure 7. Liquid refrigerant temperature

The comparison study using simulation results indicated significant energy saving potential in steady-state conditions. The enthalpy of the refrigerant entering the condenser with and without the proposed configuration is found to reduce from 440.16 kJ/kg to 432.78 kJ/kg.

Furthermore, temperature of the refrigerant entering the condenser is decreased from 63.05°C to 48.71°C, leading to a temperature reduction of the refrigerant leaving the condenser from 30°C to 17°C. This temperature reduction yields to lessen the displacement volume of the compressor from 0.0127 m³/s to 0.0116 m³/s to drop the compressor work. Moreover, the condensing temperature in this condition decreases from 35°C to 25°C, causing a reduction of condensing pressure from 1355 kPa to 1044 kPa. Therefore, due to a discharge pressure drop, the compressor pressure ratio becomes lower to result in an additional diminution of the compressor work. In the steady state, compressor power consumption for a conventional cycle and the LPA system are respectively 13.9 kW and 9.04 kW, indicating 34.9% energy savings. Furthermore, enthalpy of the refrigerant entering the evaporator by using the LPA system is nearly 6.85% less than of a conventional system, i.e. more refrigeration effect for the LPA system evaporator. However, simulation results in the designed steady state show a less energy saving potential due to a high ambient temperature in that condition. It means that the potential of energy savings using the LPA technique is exceedingly dependent on the ambient temperature and would increase with dropping the ambient temperature, as shown in Fig. 9. Moreover, according to results the amount of energy saving for the designed steady-state condition is 8%, almost due to more sub-cool temperature created by liquid refrigerant injection to the discharge line rather than floating the
condensing pressure. As a result, the energy saving potential of the LPA system depends more on floating the condensing temperature at a low ambient temperature, and more on increasing the sub-cool temperature at a high ambient temperature.

4.2 Energy Analysis

To estimate the energy saving potential of the LPA system over the summer, its energy consumption is compared with that of a conventional rooftop package. The hourly power usage of both systems is shown in Fig. 10, as obtained from TRANSYS, indicating nearly 25.3% less energy consumption than in the conventional system. It is noted that the compressor and DX evaporator VAV fan power consumptions for the LPA system are less than those for the conventional system while the air-cooled condenser VAV fan power usage is higher than that for the conventional system. The average energy savings potential of the proposed design for the compressor and evaporator fan are respectively 26.5% and 3.6% while condenser fan power consumption increases by 4.8%. The reason is that the temperature reduction of the refrigerant entering the condenser yields in a temperature decrease of the refrigerant leaving the condenser, i.e. refrigerant sub-cool temperature at the condenser outlet is increased. Therefore, enthalpy of the refrigerant entering the evaporator is reduced to increase refrigeration effects, and in turn, the system COP. This causes more heat transfer in the evaporator to decrease the supply air temperature, thus providing building cooling demand to slightly reduce the supply air flow rate, hence to result in less power consumption of the supply fan. Increasing the refrigeration effect also leads to more heat exchange between air and refrigerant in the evaporator. Additional heat absorbed by the refrigerant, however, results in a slight increase in the heat rejected by the condenser to the ambient air, which tends to increase slightly power consumption of the condenser fan. Overall, the total energy usage for each summer month can be obtained by summation of the whole system’s energy consumption in each working hour. Figure 11 shows simulation results for the average energy consumption of the conventional and LPA systems and compares it in each summer month. It can be seen that the power consumption of the LPA system is significantly less than that of the conventional plant.

5 Conclusion

We have presented a promising technique for energy efficiency improvement in DX air conditioning system by using liquid pressure amplification.

The simulation-empirical method has been applied for system modelling. The obtained models are validated and used for performance prediction. Simulation results show that by allowing temperature and pressure to fluctuate within the condenser, the reconfigured DX air-conditioning systems with LPA consume less electricity. The method can be used on DX systems that operate with fixed head pressure control. LPA devices can be fitted to new or existing systems. Results showed that the LPA approach is more effective when ambient temperature is falling. The proposed system can save electricity around 25.3% in average.

Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>area (m²)</td>
</tr>
<tr>
<td>AU</td>
<td>overall heat transfer coefficient (kW/°C)</td>
</tr>
<tr>
<td>Cₚ</td>
<td>heat capacity (kJ/(kg°C))</td>
</tr>
<tr>
<td>D</td>
<td>centre-to-centre spacing between fins (m)</td>
</tr>
<tr>
<td>d</td>
<td>diameter (m)</td>
</tr>
</tbody>
</table>
Subscripts

- $a$: air
- $Al$: aluminium
- $AR$: aspect ratio
- $avg$: average
- $amb$: ambient
- $b$: bare tube
- $comp$: compressor
- $cond$: condenser
- $Cu$: Copper
- $dis$: discharge
- $eva$: evaporator
- $f$: fin
- $fr$: frontal
- $hr$: heat rejection
- $i$: inlet, inner
- $l$: liquid
- $max$: maximum
- $min$: minimum
- $o$: outlet, outer
- $r$: refrigerant
- $sat$: saturated
- $sh$: superheated
- $suc$: suction
- $sup$: supply
- $tp$: two-phase
- $w$: wall

References


A Closed-loop System of Construction and Demolition Waste Recycling

J. Brennan\textsuperscript{a}, G. Ding\textsuperscript{b}, C.-R. Wonschik\textsuperscript{c} and K. Vessalas\textsuperscript{d}

\textsuperscript{a}School of Civil Engineering, University of Technology Sydney, Australia
\textsuperscript{b}School of the Built Environment, University of Technology Sydney, Australia
\textsuperscript{c}Faculty of Automation and Computer Sciences, University of Applied Sciences, Germany

Email: Jane.Brennan@uts.edu.au, Grace.Ding@uts.edu.au, rwonschik@hs-harz.de, Kirk.Vessalas@uts.edu.au

Abstract -
This study discusses the construction and demolition waste recycling stream both in Australia and in Germany. Differences and commonalities in commercial practices between the two countries are outlined, and open research questions are introduced. Following McDonough and Braungart’s cradle-to-cradle theorem, and ideal closed-loop system within the building life cycle is proposed. Deficiencies and efficiencies in the closed-loop system are reported and assessed and related parameters promoting or hindering the closed-loop system are evaluated. Results of the study demonstrate that reusable and non-reusable materials generated from construction and demolition operations, which are destined for landfill, are categorized differently between the recycling systems used in Australia and Germany.

Keywords -
Construction and Demolition Waste, Closed-loop system

1 Introduction
Worldwide building materials, which constitute approximately 50% of all used materials and of the solid waste generated, have considerable environmental impacts throughout the entire construction process [1]. In order to move towards more sustainable construction, it is necessary to create a material flow that returns as much as possible into the building life cycle. This paper proposes an ideal closed-loop system and investigates potential drivers for a more sustainable building life cycle. Since waste is a large component of the material flow, the following discussion will focus on construction and demolition (C&D) waste which is often defined as “the solid component of the waste stream arising from the construction, demolition or refurbishment of buildings or infrastructure” ([2], p.3). It contains no\textsuperscript{1} foreign material in the case of inert C&D waste or less than 25% foreign material in the case of mixed C&D waste. C&D waste includes “bricks, concrete, tiles and ceramics, steel and inert soils.” ([2], P.3) Permissible foreign materials exclude “Municipal Solid Waste, Liquid, Listed, Hazardous or Radioactive Waste.” ([2], p.3) Based on a case study on C&D waste recycling in Australia and Germany, the ideal closed-loop system was developed and potential drivers that could be used to move towards such an ideal system were investigated using figures from both countries.

The next section discusses challenges and drivers of C&D waste before introducing the closed-loop system. A general discussion on commercial practices in C&D waste recycling in Australia and Germany follows. There are two case studies, one for an Australian mixed C&D waste recycling facility and one for a German (inert) C&D waste recycling facility. In both cases the pricing structures are outlined and then compared in relation to general waste statistics to establish a rapport on the potential causes of different recycling rates.

2 Challenges and Drivers of C&D waste recycling
Waste is produced in different types and quantities throughout the life cycle of a building with the bulk of the waste being produced during the construction and demolition phases. The environmental problem posed by waste generated during the C&D phases is not only from its increases in volume, but also from its method of treatment [3]. This process generates waste and this waste has caused a serious environmental problem. The figures generated for construction-related waste amount

\textsuperscript{1} Less than 5%
to approximately 30% in the US, 35% in Canada and 50% in the UK [4, 5, 6]. In Australia C&D waste accounts for about 38% of the total waste generated in 2006/7 and about 43% of C&D waste went into landfill [7, 8]. The waste is often generated due to an inefficient management onsite or workshop. Most of this waste can be recycled and reduced if properly planned and managed.

C&D waste minimisation is increasingly realised as an ecologically sustainable strategy in the construction industry as a way to ease the need for virgin materials, minimise the space and cost for the disposal of waste materials, and subsequently enhance efficient use of the materials [9].

Traditionally, waste produced during construction or demolition is considered as a homogenous by-product of the process. However this approach of handling C&D waste represents a loss of opportunity as most of these wastes are not waste, but valuable resources if they are salvaged and separated properly [10]. Therefore recovering waste and keeping it separated offers significant economic benefits in the construction industry. It minimises the transportation and disposal costs for landfilling. In addition the salvaged materials generate economic benefits either by selling them for recycling, or by incorporation into future projects. However according to Marie and Quarsawi [9] the current use of recycled C&D wastes are largely down-valued as they are mostly used as hardcore or backfilling materials. Therefore proper management will need to be considered at the outset if the value of C&D waste is to be fully recognized.

Barriers and challenges have been well examined and addressed in research studies. The negative perceptions and unawareness for the usefulness of recycled materials in the construction industry has been the major hurdle for C&D waste management. Osmani et al. [11] conducted a questionnaire survey to the top 100 architectural and contracting companies in the UK and realised that C&D waste management is not fully recognised and considered in the design process of a project. Yuan and Shen [12] suggest that concern about extra labour and equipment cost in collecting and sorting have been the key barrier to the recycling practices in construction.

Quality of recycled materials is another concern as it may be hampered by cross-contamination with other materials in the salvaging, collecting and storing process [13]. Asbestos contamination is a well-documented problem and the most common form of quality control is at the recycling yard for visual inspection. The lack of an industry-wide standard and material testing program to validate compliance to the prescribed standard are major problems for the C&D waste.

As the range of products and materials increases there will be a vital need for training and education to provide greater awareness for the importance of waste materials, and knowledge of how to use them successfully in projects as alternatives to virgin materials. The perceptions will only be changed if more data for field trials show how to use these materials to their optimum performance. The lack of technology and equipment to sufficiently clean or segregate materials has caused significant volumes of materials being sent to landfill. There is a clear need for investment in research and development to overcome technological barriers.

Over the years, research and development in the minimisation of waste has gone underway and a waste management hierarchy has been well established. The waste management hierarchy builds on the principles of 3Rs and 5Rs. The principle of the C&D waste management hierarchy is to reduce consumption of raw materials for environmental conservation along with the continued recycling and reusing of C&D waste. This principle allows and directs the construction activities to an environmentally-friendly process, therefore reducing the eventual material within landfill. The 3Rs principle in C&D waste management refers to reduce, reuse and recycle. It has significant impact on reducing the need to extract raw materials, reducing the amount of materials going to landfill sites and thus reducing the life-cycle costs of buildings and building materials. The 3Rs principle has further been developed into a 5Rs Rule where refuse and repair have been added into the hierarchy [9, 14].

Reuse of building materials deals with a serious resource issue. The reuse of materials involves the consideration of the material and joining techniques so as to enable the reuse and replacement of components, either in parts or as a whole. When the reuse of a component is not possible, it may still be possible to recycle it in whole or in parts. Reduce relates to the reduction in the use of resources, space or elements. It is not necessary to undermine a good design solution such as reducing the amount of mechanical services to suit. It involves the adaptation of existing buildings instead of demolition and reuse of salvaged materials to minimise raw material consumption. Refuse refers to setting guidelines on what are and are not acceptable materials in buildings. For example, the European Commission released the integrated product policy to identify products within the construction sector for products that have the greatest lifetime environmental impact potential. Repair is a strategy that aims to reduce a wasteful lifestyle by overhauling and refurbishing to
extend the useful life.

3 A closed-loop model for the C&D waste stream

Traditionally the use of construction materials is characterised by a linear process of extraction, manufacture, construction, maintenance and refurbishment, demolition, and disposal. However, in the last two decades there has been a shift in the construction industry from linear production and consumption of materials and products to some cyclic manufacturing activities [15]. Therefore, products and components used in construction should be designed in such a way to enable materials to maintain their status as resources. As such at their end of life cycle, materials are re-entered further into another life cycle in substitution of virgin materials [16].

This shift considers a closed-loop approach in C&D waste management as an alternative to the traditional linear process. The closed-loop approach allows materials and components to be reclaimed, reused and recycled multiple times during the life-cycle. The closed-loop approach in C&D waste management plays a significant role in achieving the goal of sustainable construction as it aims at closing material life-cycle loops where waste from one process will be the resource for another. Figure 1 presents the model of linear and closed-loop system in C&D waste management.

The traditional linear value chain allows valuable C&D waste to end up in landfill and impact on the environment. However the closed-loop approach utilises the 5Rs principle by closing the material life-cycle loop.

The concept of cradle-to-cradle and zero waste is the heart of the closed-loop waste management system. It is related to product design and builds on the idea of eco-efficiency of materials with the aim of reducing the consumption of virgin resources and eliminating waste and pollution [15]. In principle materials are extracted from buildings and re-integrated directly or reprocessed and then re-integrated into buildings or put to useful purpose in other sectors without creating any waste. According to Sassi [17] the principle of the closed-loop approach would include the ability to be reprocessed infinitely through industrial or natural recovering without significant loss of material quality and mass, and without uncontrolled or significant pollution emissions. The closed-loop approach of C&D waste represents an ideal system which may be difficult to achieve, but it provides a goal for the construction industry to improve the current practices of construction related activities.

4 Commercial Practices in Australia and Germany

Governments all over the world have developed regulations and legislations to positively impact on the re-use of construction materials and other products derived from the C&D waste stream [1]. While there is the common goal of reducing the “energy and emission associated with building demolition and transport to landfills of materials that would not currently be recycled and reduced” [18, p. 204], the way this is realised in varying degrees can look very different across different countries or jurisdictions.

The general practice regarding C&D waste management is to remove the waste from the construction site for disposal or recycling elsewhere; there is usually not a lot of re-use onsite. What happens to the waste after removal of course, takes different forms. This can encompass a simple disposing in landfill, removing it to a waste separation company, or pre-sorting and then transporting the C&D waste to a specialised recycling facility. These specialised operators accept pre-sorted waste such as brick and concrete waste only and focus their operations on recycling. There are also options involving mobile recycling units which can be hired out and allow recycling directly onsite. The recycled material is then transported to other places. Sometimes for example in road construction; it might be used directly on the site again. This kind of onsite reuse is however the exception. Generally, recycled materials and products are offered on the market by the recycling facilities to be used elsewhere.

The main difference between the Australian and German C&D waste management processes appears to be in the common utilisation of separate salvaging operators in Germany. These companies collect the C&D waste from the construction sites or get them delivered by the builders, and then pass on the useable components of it to the C&D recycling facilities. In Germany, there
appears to also be a far greater rate of pre-sorting of C&D waste on the construction sites than in Australia. In Australia on the other hand, mixed waste facilities fill the role of the German salvagers by also accepting unsorted C&D waste, and sorting it into its recyclable and non-recyclable components. The mixed-waste facilities might also recycle the sorted C&D waste on their own sites, contrary to German salvagers who will pass it on to specialised recycling facilities.

In a series of interviews conducted, we found that the Australian, in particular New South Wales (NSW), construction industry is much more willing to use recycled C&D waste materials, partially due to a lack of easily accessible virgin materials, while there is a great reluctance on the German side; which seems to be slightly contradictory to the fact that there is a higher recycling rate in Germany than Australia (and NSW).

5 Case Study of Australian and German C&D waste facilities

This section will introduce a case study for an Australian and a German C&D waste recycling facility. It is important to note that the actual technology used for sorting and recycling is the same. This might include sieves, crushers, water bath and magnets. However, as previously mentioned, the process in the chain appears to be different in terms of which stakeholder performs which task. Hence the C&D waste management process will be considered as a supply chain along which the waste is generated either as supply for new materials or products that are based on it, or is disposed as landfill. According to the Australian C&D Waste Status Report [8], countries like the Netherlands produce virtually no landfill from C&D waste; everything is re-used or recycled in some future construction process. This is quite remarkable. Both Germany and Australia are still quite some way from this. In Australia, approximately 57% of C&D waste was recycled nationally in 2006-07 [8]. There is a high variability in the recycling rates across the country. Since the Australian case study looks at a NSW facility, it is worth mentioning that for the same period 67% of C&D waste was recycled in NSW. These rates have since increased; however, 2006-07 data will be used since these are the most comprehensive data sets available [8]. For a similar period, the nationwide recovery rate of C&D waste in Germany was more than 86% [8].

5.1 Australia

The mixed C&D waste facility Brandown in Sydney’s south-west is a privately owned and operated Resource Recovery Centre, General Solid Waste (Non Putrescible) Landfill and Quarry. They have a C&D waste recycling facility onsite, and ensure that anything that is recyclable in the C&D waste delivered is separated and made available as a resource². Brandown recycling has two streams of incoming C&D waste; a mixed stream which

---

includes “anything” except asbestos and a recyclable C&D waste stream. The former is sorted to separate reusable construction waste going to the recycling facility from residual waste, which is being deposited to landfill. Due to the considerably high NSW landfill levy, the prices for the landfill component are relatively high. It would be interesting to see if the recycling rates of C&D waste have risen in relation to the rising levy. In this paper, the focus will however be on current charges and how they might impact on recycling rates.

Table 1. Brandown Waste Disposal Charges

<table>
<thead>
<tr>
<th>Waste Type</th>
<th>$ per Tonne</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>13</td>
</tr>
<tr>
<td>C2</td>
<td>22</td>
</tr>
<tr>
<td>C3</td>
<td>50</td>
</tr>
<tr>
<td>MM1</td>
<td>22</td>
</tr>
<tr>
<td>MM2</td>
<td>196</td>
</tr>
<tr>
<td>MW</td>
<td>196</td>
</tr>
</tbody>
</table>

Brandown’s charges as of 1st July 2013 are examined in more detail. Selected items were chosen, which can be related to similar items in the German case study for the purpose of comparison. Prices are given for the following waste or recyclable materials: C1 (small pieces with no inclusion), C2 (medium pieces with a small amount of inclusion), C3 (large pieces with a small amount of inclusion), MM1 (mixed masonry: brick and concrete, small with no inclusion), MM2 (mixed masonry: brick and concrete, with inclusions), MW (Mixed/general waste, recyclable or non-recyclable). Table 1 summarises Brandown’s prices for these categories of waste.

5.2 Germany

As mentioned before, in Germany C&D waste is usually pre-separated in different bins on site. Salvaging companies might accept mixed waste, but since it is very highly priced, sorting takes preference on most construction sites. A pricelist including mixed waste prices will be provided in this section, however, the case study focused on a C&D waste recycling facility, since mixed waste C&D waste recycling facilities are very uncommon in Germany.

The Recycling and Sanierung Thale GmbH (RST) is a C&D recycling facility in north-eastern Germany, in the Harz Mountains. They have three streams of C&D waste treatment: high-pressure soil washing, deposit for dangerous and non-dangerous waste, and a C&D recycling facility. RST accepts sorted waste only. The waste is then separated further if necessary by washing, sieving and so on. It is recycled by the usual methods such as crushing and sieving. The technology used for these operations is almost identical to the recycling machinery used at Brandown, often it has been bought from the same European suppliers.

Table 2. RST and other local Waste Disposal Charges

<table>
<thead>
<tr>
<th>Waste Type</th>
<th>EUR per Tonne</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>6</td>
</tr>
<tr>
<td>C2</td>
<td>8</td>
</tr>
<tr>
<td>C3</td>
<td>11</td>
</tr>
<tr>
<td>MM1</td>
<td>16</td>
</tr>
<tr>
<td>MM2</td>
<td>24</td>
</tr>
<tr>
<td>MW</td>
<td>135</td>
</tr>
</tbody>
</table>

RST’s prices do not include mixed waste, since they do not accept this type of waste. Prices for C&D waste was directly taken from RST’s website, mixed waste prices were retrieved from a web database containing waste disposal prices for the area of Thale. The categories are similar to the categories defined for the Brandown case study.

5.3 Comparison

There are two aspects along which one can compare the C&D waste recycling operations in Australia and Germany. Firstly, a comparison can be done by the variations in the life-cycle model between the two countries. This was already done in the previous section on C&D waste recycling practices. Secondly, the different pricing structures and what effect they might have on the actual recycling rates can be examined more closely.

As previously mentioned, high recycling rates are strongly linked to legislative incentives. However, could the higher recycling rates in Germany in part also be explained by the pricing structure as shown in each case study? It is assumed that the structure examined in the following, can be seen as indicative to the Australian and German situations. Can it be assumed that the larger the difference between the prices for mixed/general waste and recycled waste is, the higher is also the recycling rate? Is there a direct correlation between

---

3 Available on their website (footnote 1)

5 www.abfallscout.de (queried on 20.1.2014)
6 In particular NSW
recycling and pricing rates? The rates were either taken or derived from statistics in the aforementioned 2011 C&D waste status report [8]. The price difference was calculated in percentage between general waste and each type of recyclable waste, thus showing the savings attainable for each case study (representative for the countries the case studies were situated in) when recycling C&D waste instead of directly depositing it into landfill. Table 3 summarises these statistics and Figure 4 compares the Australian and German savings.

Table 3. Savings made when recycling different types of waste compared to mixed waste prices

<table>
<thead>
<tr>
<th>Type of Waste</th>
<th>Australian Savings to MW in %</th>
<th>German Savings to MW in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>93.37</td>
<td>95.6</td>
</tr>
<tr>
<td>C2</td>
<td>88.78</td>
<td>94.08</td>
</tr>
<tr>
<td>C3</td>
<td>74.49</td>
<td>91.86</td>
</tr>
<tr>
<td>MM1</td>
<td>88.78</td>
<td>88.09</td>
</tr>
<tr>
<td>MM2</td>
<td>0</td>
<td>82.23</td>
</tr>
<tr>
<td>MW</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

As can be seen in table and figure, mixed masonry with inclusions does not give any economic benefit as compared to general and mixed waste in Australia. The savings are at least the same or higher in Germany.

The percentage of savings is also reflected in recycling rates, which are 57% nationwide in Australia, 67% in NSW (as derived from data in [7]) and 86% in Germany. If compared to the average cost savings to mixed waste of 57% in Australia and 75% in Germany, there is a clear correlation between pricing structures and recycling rates. This is shown in Figure 5.

6 Conclusions

There are great opportunities for C&D waste to be recycled and reused. The paper discusses issues and challenges of C&D waste in the construction industry and presents a closed-loop model for C&D waste as an approach to achieve the goal of zero waste. The closed-loop model may be ideal but it is possible and advantages in maintaining the sustainability of the natural resources on the environment.

The paper also presents results of a comparative study of C&D waste management processes between Australia and Germany. The results reveal that the German practice has a higher waste recovery rate than the Australian practice. The pre-sorting and separating facilities on construction or demolition site in Germany place an important role in achieving high recovery rate of waste for recycling and reuse. This aligns with the literature review that high recycling rates for materials are achieved when materials are captured closer to the source before mixing with other wastes. Whilst there was insufficient pre-sorting and separating facilities in Australia there is therefore a higher risk of cross-contamination that may hamper the recovery of useful materials in the C&D waste.

The closed-loop model may be viewed as an ideal system but it represents a goal for the construction industry to raise awareness among professional to achieve better practices. Ultimately the concept can encourage the construction industry to rethink the design and construction by taking into consideration for dismantling and reuse at the end of a building’s life at an outset and the use of processes and materials on the construction and demolition project and to innovate to meet the challenge of sustainable development.
References


Development and First Testing of a Framework for Predictive Energy Control of Underground Stations

A. Giretti\textsuperscript{a}, R. Ansuini\textsuperscript{a} and A. Carbonari\textsuperscript{b}

\textsuperscript{a}Università Politecnica delle Marche, DICEA Department, via Brecce Bianche, 60131 Ancona, Italy
E-mail: a.giretti@univpm.it, r.ansuini@univpm.it, alessandro.carbonari@univpm.it

Abstract -
Building Energy Management Systems consist of hardware and software components. The hardware set-up of BEMS is typically made up of a set of computers in charge of building control and sensor-actuator networks. The software side of BEMS is usually made up of a number of functional layers that implement standard management functionalities. This paper will present an application of Model based Predictive Control (MPC) targeted to energy management of the “Passeig de Gracia” metro station in Barcelona. This approach uses the predictions of future building status, obtained by means of a set of Bayesian Networks, in order to determine the optimal control policies. First the predictive Bayesian Networks were developed through the following steps: structural learning based on a simulated dataset; improvement of the network’s topology through enhanced datasets derived from the previous one; final refinement and validation based on experimental data collected through a pervasive wireless monitoring network. Then those networks were integrated within a control framework, including control algorithms, a Dymola\textsuperscript{TM} based virtual model of the station to simulate its evolvement and, on top of them, a user graphic interface to manage the system. The results about energy savings estimation determined by the application of model based predictive control to the station’s mechanical ventilation showed that as much as 35\% can be saved on average.

Keywords -
BEMS; predictive control; real-time control

1 Introduction

The development of an innovative adaptive control of HVAC based on the use of predictive models is part of an ongoing research project, funded by the EU Commission and called “Seam4us” (http://seam4us.eu/). The pilot of such a project is the “Passeig de Gracia” station in Barcelona (Spain). This approach will overcome the traditional homeostatic short-term feedback mechanisms which are applied singularly to each equipment type. This paper concerns the design and the development of a new type of intelligent building energy management system (which is usually referred to as BEMS), that is able to optimise the operation of the mechanical air supply systems of the Passeig De Gracia metro station in Barcelona. To the purpose of this application, predictive models were developed to support the optimal control of indoor environmental conditions in the station, which was necessary due to the many interacting variables of the domain. BEMSs usually consist of hardware and software components. The hardware set-up of a BEMS is typically made up of sensor-actuator networks that accurately monitors the indoor-outdoor environment and the building plants state and drive the systems accordingly. The software side of a BEMS consists of a number of functional layers that implement standard management functionalities like plant status monitoring, alarm management, demand driven plant management, reporting, etc.. [1]. Still plant and building set-points follow prescribed schedules and are rarely optimized in response to changing dynamic conditions, including weather, internal loads, occupancy patterns, etc. Nonetheless, there are significant opportunities for optimizing control set points and modes of operation in response to dynamic forcing functions and utility rate incentives. A number of studies [2] have shown potential savings for optimized controls in the range of 10\% to 40\% of the overall cooling cost.

Model Predictive Control (MPC) may be used to enhance BEMSs so that they can improve their control performances getting close to optimal behaviour. MPC is an advanced control technique [3] that uses the predictions of future building status, obtained by means of a model of the building’s dynamics, in order to solve the problem of determining the optimal control policies in advance and anticipate its reaction to external forces. But this requires the development of integrated models capable of predicting the near future behaviour of the controlled environment under specific conditions, so that the optimal solution can be sought through scenario analysis. Furthermore, MPC models must interoperate
with real sensor/actuator networks that usually, for cost reasons, cannot be larger than few tenths of devices and whose deployment is constrained by a number of external factors. Nevertheless, the model accuracy must be granted despite the reduced representation of the physical model and the suboptimal selection of the parameter set. The fulfilment of such competing requirements compels the definition of a modelling framework that, by guiding the MPC modeller through a set of methodological steps, will contribute to design accurate and robust models, which are sufficiently light to be embedded in real control systems. Thus far the model part was usually left to statistical models and it was usually targeted to quite simple domains. In this paper, a new probabilistic approach was tested, suggesting that Bayesian Networks can provide the means to manage very complex domains. In particular, they are shown to be able to make correct inferences in the case of a metro station, whose behavior is affected by a number of variable and interacting physical phenomena. Hence they supported the development of a MPC scheme.

2 The case study: underground station PdG

The PdG metro station in Barcelona is a 3-line connection station between metro lines no. 2, 3 and 4 (Fig. 1). Line 3 (L3) is located in the northern hub of the station, which includes spaces devoted to different activities: commercial, transportation, people movement, public and technical services, staff reserved rooms. A spatial survey in the station led to the identification of the following types of spaces: entrances (E), halls (H), corridors (C), platform (P) and rooms (R), including technical rooms, restrooms, vestibules and other areas whose access is restricted to the staff. Internal comfort is managed by means of several systems. The public access area is mechanically ventilated. The whole station is lit by means of regular, auxiliary and emergency light fittings controlled by several power circuits. People movement is favored by upward escalators. Other systems (e.g. split units, communication) are installed in commercial, technical and staff only rooms.

In this paper we will show how MPC can optimally regulate comfort by means of a dynamic control strategy, instead of by a set of predefined design constraints. To that purpose, the station must be capable of dynamically accommodating the user needs, by driving the fans located in the station’s technical room. The main ventilation ducts leave from here to convey outdoor fresh air into the platform (PL3). Air intakes are located above both platform’s sides and they supply air changes. Two CONAU V1080 injection fans (that are the main object of our control) are located in the station’s technical room, and other two fans are extracting air through ventilation shafts in the middle of the tunnels adjoining PdG-L3 (which are not controlled instead). The current daily summer ventilation schedule keeps injector fans on during the day (from 5 am to 10 pm) at their highest rate. They are switched off in the night. Similarly is valid in winter, but the fans’ input frequency is halved, and their air flow rate is reduced at about one third, as a consequence. Also, outdoor ventilation is conveyed through its five entrances and corridors leading to the platform.

![Fig. 1 – Spatial layout of the PdG underground station (a) and pictures (b).](image)

3 MPC control

As mentioned in the Introduction, the Bayesian networks developed in this chapter were used to provide forecasts about the future state of the PdG-L3 in Barcelona, given the knowledge about their current state. Any control in buildings is targeted to minimize power consumption while keeping required comfort level and guaranteeing robustness of the solution. To this purpose, the control system must be optimal and adaptive, which is "a special type of nonlinear control system which can alter its parameters to adapt to a changing environment. The changes in environment can represent variations in process dynamics or changes in the characteristics of the disturbances [...]" [4]. Reliability is also required, and the predictive feature is another opportunity for achieving high energy efficiencies: prediction gives the capability of taking soft control actions in advance instead of suddenly reacting to unexpected deviations from the required state, thus saving energy. MPC takes into account the (measured) current state of the system, future weather conditions and other disturbances (e.g. internal gains), in order to control actuators (e.g. HVAC, lighting and blind systems), so that energy and money usage are minimized. At the current point in time, a heating/cooling plan is formulated for the next several hours to days, based on predictions of the upcoming weather conditions. The control action is designed by running the model of the process over a given prediction horizon and evaluating the control sequence that gives the minimum value of the cost function [5].

One remarkable survey about the effectiveness of...
MPC was carried out by means of simulations and applied to office buildings [6]. First, the authors considered and compared a list of potential adaptive approaches, among which we cite reduction of the thermal comfort when the building is not used, widening of the room temperature comfort range, use of Indoor Air Quality controlled ventilation. Those preliminary simulations showed that the highest energy savings were determined by predictive control [7].

In the case of large underground buildings, like PdG-L3 metro station in Barcelona, interaction with the outdoors is very complex and occupancy figures result someway difficult to predict. Hence, the dynamics of the station cannot be solved – and predicted – though a simplified thermal model. Bayesian Networks will be shown to work well when it is necessary to reduce a complex building model into a more manageable one. In fact, they gave back a lumped representation of a complex system, involving thousands of variables.

The overall MPC control framework applied to the station is represented in Fig. 2. Inputs u to the system are the variables that can be driven by the controller (e.g. frequency that drives injector fans). The outputs y are the power consumption and indicators for comfort and health that must be controlled in order to reach certain desired reference level r. The relation between inputs and outputs is also significantly affected by a set of disturbances d, such as weather, train arrival, passenger flows and fans external to the station: they cannot be manipulated but only “accounted for” by using direct measures. At each control step, the prediction model receives candidate input sequences ū picked out by the controller; disturbance predictions come from disturbances models ḅ. Measured outputs m from PdG-L3 and the prediction model estimates the future output sequence ŷ. The optimal control sequence ũ is that one which minimizes a given cost function while complying with given constraints. Once the optimization problem has been solved, the first step u of the optimal sequence is applied as the best control action. The overall procedure is repeated at each step, thus closing the control loop. The implementation of those systems asks for the development of devices and services:

- monitoring systems and intelligent algorithms to interpret occupant’s behaviour, as deeply explained by the authors in [8];
- high-level control systems capable of solving optimization problems in real-time;
- accurate and fast dynamic models of buildings’ behaviour and their systems (which is object of this paper);
- accurate modelling of disturbances.

4 Development of predictive models

Indeed, Bayesian Networks may be thought as a directed acyclic graphs that encodes assertions of conditional independence [9]. In fact, it orders the variables in a domain U. They are suitable to reduce complex domains into computationally manageable models, which is a key feature when computations must be performed in real-time. Also, they are capable of managing incomplete (e.g. one or a few data are not available because the corresponding sensors are broken) and uncertain information (e.g. if we include uncertainty in sensor measurements or if inputs are relative to forecasts of disturbance actions).

They implement inference algorithms, thanks to the conditional probability relationships defined among the variables of the domain under analysis [10, 11]. In other words any node can be conditioned upon new evidences. This feature is particularly important in case a control system must work in real-time, because in that case evidences acquired about a state variable (i.e. from sensor measurements) must be propagated to update the state of the rest of the domain. When it is run in the MPC framework, the controller will make queries to a set of nodes belonging to the networks, whose probability distributions are computed from the state of other nodes, upon which observations (or evidences) are already available. In the case of PdG-L3 presented in this chapter, the Bayesian Networks were built in the Hugin™ software environment. The conditional probability tables were learned from datasets put together through numerical simulations, by means of the “EM learning” algorithm [9].

In order to validate their performances, different kinds of indices were developed. The difference between the predicted value ŷ and the actual value X is defined as error E = ŷ − X. The absolute error is |E| and its squared error is S = E^2. Percentage error will be: PE = 100 · E/X. In order to have a global performance index to be evaluated over the whole validation dataset made up of K samples, these
As indices of the predicted variables are related to different physical quantities with different units, they should be normalized with respect to their typical range of variation, by means of:

$$\text{MAE} \equiv \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

and the root mean square error is:

$$\text{RMSE} \equiv \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}.$$  \hspace{1cm} (2)

As indices of the predicted variables are related to different physical quantities with different units, they should be normalized with respect to their typical range of variation, by means of:

$$\text{NMAE} \equiv \frac{1}{|x_{max}-x_{min}|} \sum_{i=1}^{n} |x_i - \hat{x}_i|$$

and:

$$\text{NRMSE} \equiv \frac{1}{|x_{max}-x_{min}|} \sqrt{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2}.$$  \hspace{1cm} (4)

ASHRAE Guideline 14-2002 [12] establishes that for calibrated simulations, the CVRMSE and NMBE of energy models shall be determined for each calibration parameter by comparing simulation-predicted data to the utility data used for calibration. The proposed indices are the coefficients of variation of the root mean square error (CVRMSE) and normalized mean bias error (NMBE). Following this guideline, the RMSE has been selected as the main performance index for evaluating the accuracy of a BN. The range of the considered variable has been taken as a normalization factor and the NRMSE has been selected as final index for the design process of the BN because it includes information about both bias and variance of the error.

4.1 Predictive models

Basically, the development process of both Bayesian Networks consists of three main phases:

1. definition of the network topology;
2. preparation of the training set and learning of the conditional probability tables;
3. final assessment of the network.

In the case of this chapter, the behaviour of the metro station Passeig de Gracia was first simulated through whole building analyses, which provided datasets encompassing all the possible environmental conditions, such a knowledge was transferred into Bayesian Networks then. Three datasets were generated:

- the first one was made up of randomly generated data, which means that the inputs (e.g. weather, heat gains, occupancy figures etc.) were allowed to vary without additional constraints in their range;
- the second sample, called “likely” dataset, was generated through simulations whose inputs were allowed to vary within their same ranges cited above, their differential variations being constrained, so that the difference between the value of each variable at the present time step and the value of the same variable at the previous time step was limited by a threshold;
- the third “typical” sample was built through simulations, whose inputs were taken from real measurements, such as real weather conditions, number of people etc ...

The use of the random dataset was targeted to provide to the networks information about any kind of combination of events possible, including the less likely ones. Then, more information about the more likely scenarios included in the second and third datasets was added. These two last datasets were constrained by setting input variables within those values which were measured in past years (e.g. weather conditions, occupancy, driving frequencies of ventilation systems). The whole building model used for running simulations was developed as a lumped parameter model in the Dymola™ simulation environment, that is based on the Modelica language [13]. Starting from a validated library for building simulation developed by the Lawrence Berkeley National Laboratory [14], a specific library for underground stations was developed. However, such a model cannot be run in real-time when the controller needs to determine the best candidate control strategies, so it was reduced into the less computationally demanding form of Bayesian Networks. The PdG-L3 predictive model was split into two Bayesian Networks:

1. temperature prediction dynamic Network (TP-DBN), which is in the form of a DBN, because it forecasts expected temperature in the station given inputs about current and past time steps;
2. air flow prediction Bayesian network (AF-BN), which is in the form of a regular BN, because it estimates variables relative to air flow in the station and energy consumption of the fans, given its current status.

Once available, the two networks were run according to the scheme outlined in Fig. 3. At every iteration the controller will opportune query the two networks to get future estimations about the variables relevant to select the most opportune control policy to be adopted at each running step. To this aim, the networks need to be instantiated first: the current temperature in the station’s platform (PL3) and weather conditions will be provided by the permanent monitoring network installed in the station, along with candidate fan frequencies. Given these inputs, the controller is allowed to query the AF-BN in order to estimate fans consumption and air changes in the station at each time step. Such a prediction step takes a few seconds and is performed by the software Hugin™ through algorithms for belief propagation. Then, the TP-DBN will take these variables as inputs, along with other state variables (e.g. current PL3 temperature, temperature difference between inside and outside and forecasted weather, people, train arrival etc.) in order to
predict PL3 temperature at the next time step. Again, belief propagation is performed with this second network. Then, the same loop will be repeated at each iteration. Both the networks were built following the same methodology:
- first structural learning: it was determined first by the a-priori knowledge from the researchers and cluster analyses;
- improvement of the network’s structure: this was carried out through analysis of its performance indices, after learning conditional probability tables from the random dataset;
- final refinement using the two additional datasets: adding more datasets allowed the developers to quantify even probabilistic relationships among the variables;
- final evaluation of the networks.

Fig. 3: The basic loop of the predictive cycle adopted for PdG-L3 which involves both Bayesian Networks.

The first step started from the analysis of 81 variables included in the Dymola™ dataset. Iterative cluster analyses [15] were useful to group those variables into clusters and determine those which were redundant [16]. Finally, the number of variables was cut down to 25. This final set’s variables were naturally grouped into two sub-clusters of variables: one of them including those related to air flow processes, and the other one including those related to the temperature dynamics.

The second step helped in several tasks: meaning and dependencies between nodes have been reviewed according to the relationships suggested by physical laws; the number of intervals for discretizing the state space pertaining to every node, with the final purpose of minimizing the errors of the output variables given by the performance indices; a few links have been rearranged.

Fig. 4-a depicts the final structure of the dynamic Bayesian network (i.e. TP-DBN), which was used to predict PL3’s temperature in the next step (node TemPL3_p01), starting from inputs such as: forecasted number of people in the station at the next step (NPepSta_p01), forecasted internal gains supplied by trains at the next step (GaiTr_p01), current PL3’s temperature (TemPL3), forecasted outdoor temperature (TouMet_p01) and deviation of temperature from the past time step (DtePL3). The network’s intermediate variables are useful to perform computations and simplify conditional probabilistic relationships among variables. Similarly holds with the AF-BN network (Fig. 4-b), whose inputs are: forecasted frequencies of fans in the station and tunnels at the next time step (DfreTF1_p01, DfreTF2_p01, DfreSF1_p01), forecasted internal gains by trains (GaiTr1_p01), forecasted wind direction and speed (WiDMet_p01, WiSMet_p01), outdoor temperature (TouMet_p01) and current temperature (TemPL3). The main outputs are the power consumption of fans – in the station (PelSF1) and in the tunnels (PelTF1, PelTF2) – and air flow rates expected across the corridors leading to PL3: AfICNI_p01 (corridor CN1), AfICNop_p01 (sum of corridors Cno and CNp), AfICNq_p01 (corridor CNq) and AfISlb_p01 (station link). These estimated airflows are then summed up coherently to the Air Mass Balance for computing the overall air change in PL3 (ACOPL3), needed as input from the TP-DBN.

The third step was aimed at performing further refinement using the “typical and “likely” datasets. Technically, that means that the EM learning algorithm was implemented by adding the information included in these two datasets to the information already derived from the “random” dataset. This process allowed to include information about those scenarios which are likely to occur more often. The refinement was mainly...
performed in terms of tuning the subdivisions into intervals of all the nodes and in terms of converting discrete variable into continuous variables. The steps no. 2 and 3 required many iterations of learning, refining and validating. On the whole, 140 cycles were made with the TP-DBN (Tab. 1), which was useful to reduce the error from 4.98 to 0.72 °C (in the RMSE case) and from 17% to 4% (in the NRMSE case). The trend during the refinement process led to a continuous increment of performances, as shown in Tab. 1. In addition, 82 cycles were needed to optimize the AF-BN: for the control variable, Station Fan Power (PelSF1) RMSE fell down from 1858 W to 377 W, whereas NRMSE fell from 10.3% down to 2.3%. In the TP-DBN all the nodes were represented by discrete variables. In the AF-BN all the variables were continuous except the following ones: frequencies of fans (DfreTF1_p01, DfreTF2_p01, DfreSF1_p01) and wind direction (WiMet_p01).

\[
J = \sum_{k=1}^{N} \alpha_{PEITF1(k)} \left( \frac{|PEITF1(k) + PEITF2(k)|}{2P} \right)
+ \alpha_{PEISF1(k)} \left( \frac{|PEISF1(k) + PEISF2(k)|}{2P} \right)
+ \alpha_{DT} \left( \frac{T \cdot \Delta WS(k) - TempPI3(k)}{\Delta T} \right)^2
+ \alpha_f \left( \frac{TempPI3 - TempPI3(k)}{T} \right)^2
+ \alpha_{AC} \left( \frac{ACOP3 - ACOP3(k)}{AC} \right)^2
+ \alpha_{DF} \left( \frac{FreSF1(k) - FreSF1(k - 1)}{DF} \right)^2
\]

The inclusion of temperature in PL3 (TemPL3) and air changes per hour (ACOP3) were used to control comfort conditions. The respective coefficients of eq. (5) can be tuned to weigh the importance of the several concurrent factors.

### 4.3 Validation of the predictive models

Finally the performances of the two networks were verified also through simulations. Fig. 5 shows the good agreement between the real temperature simulated by Dymola™ in PL3 and the forecasted plot of PL3 as predicted by TP-DBN. The simulations performed by the Bayesian Networks in this case were carried out according to what already described. The input values at the first time step were instantiated as evidences taken by the Dymola™ model. Then, the outputs from the networks were used as inputs for the next time step in the networks and the simulations were iterated in the same way all over the period shown in the diagram. It’s clear that the predictive and dynamic Bayesian networks (BN) are able to accurately model the temperature plot sin PL3 and to give the right inputs to the controller, in order to evaluate the best control policy.

**Fig. 5: Qualitative comparison between the real temperature plot computed by Dymola™ and forecasts by the Bayesian Network TP-DBN.**

### 5 The Simulator

The Bayesian predictor and the MPC logics have been embedded in a simulation environment that accurately reproduces the thermal and air-flow
The Simulink (Mathworks©) architecture of the SEAM4US simulator is shown in Fig. 6. The simulator is made of four main components: the PdG environmental model, the passenger flow simulator, the lighting control simulator and the environmental MPC. In this paper we are showing the potentials of MPC applied just to environmental control. The PdG Model is that one developed in Dymola™. At compile time the PdG environmental model results in a matrix with tenths of thousands of unknowns. The PdG Model is interfaced with a weather file of Barcelona that provides the hourly external weather parameters, including wind speed and directions. The PdG environmental model receives as inputs passenger occupancy levels, lighting level of the appliances in each space, and fan control frequencies. It then outputs all the environmental parameters (e.g. air temperature and humidity, pollutant levels, energy consumption). These parameters are then fed back to the control logics as the basis for the next control step. In the SEAM4US simulator the large PdG Environmental model acts as the real station. The Bayesian models reported in sub-Section 4.1 support the controller by means of predictions on the future status of the station. The size of this predictor is small enough and its computational time short enough to suit the model embedding requirements.

The control logics implemented in the SEAM4US simulator is based on a particle filtering mechanism. The controller randomly generates a number of different control options that are sent to the predictor. The predictor updates the model with the control parameters and by means of Bayesian inference calculates the environment and energy consumption parameters. Then the controller ranks the predictor outcomes according to the cost function in eq. (5). The best performer is then selected and used in the next control step. Fig. 7 shows an example of a simulation results of three days of operation, which is relative to the environmental control. The simulation time is represented along the x axis, while the y axis represents the fan frequencies in Fig. 7-a Negative frequencies means that the fan direction is inverted (extracting air instead of supplying). Three curves are reported. The dashed curve (i.e. baseline) depicts the current policy used for fan control. The fan is driven at maximum speed for all the station opening time and it is turned off during the closure time. The second dash-dot curve represents MPC constrained to only two driving frequencies, while the third (continuous) curve is related to a continuous frequency driving. In addition to the fact that MPC control provides an energy saving rate that can rise up to 35%, it is noteworthy to realize why this happens. Comparing the baseline curve with the MPC controlled, it appears that in many cases the driving frequencies and the baseline have opposite signs. This means that in the standard baseline driving the station fans very often are opposed to the air flow induced by the external sources, and therefore contribute negatively to the air exchange. This is reflected by the temperature curves that are slightly lower – i.e. more comfortable - for the MPC controlled environment despite the huge energy saving (Fig. 7-b). Summarizing, these results show how the effectiveness of the MPC control of complex environment relies on the power and on the flexibility of the Bayesian predictor and of the Bayesian Inference paradigm.
6 Conclusions

Predictive control of buildings is one of the most effective ones currently being developed by researchers. However, it cannot be applied without a reliable predictor of the expected state of the controlled domain. Computationally demanding software programs cannot be used to produce predictions at run time, but they can be run to generate datasets and these datasets may be used to transfer knowledge into Bayesian Networks. In fact, inputs by the controller are instantiated in Bayesian Networks in the form of a set of evidences; then, inference algorithms are propagated and expected future values describing the energy and thermal state of the domain might be estimated. This procedure can be repeated thousands of times at each control step and it makes the implementation of MPC feasible.

When implemented in a real case, the results from inferences were shown to be very accurate with low deviations from the values estimated by means of more complex numerical models. In addition, our testing of the use of predictive Bayesian Networks embedded in a wider MPC framework to support the ranking of concurrent control policies was successful, too. So Bayesian Networks proved to be able to solve the problem of reducing complex models into more manageable tools for performing cumbersome inferences through limited computational efforts, while getting highly accurate results.

References


Task Specific Trajectory Profile Selection for Energy Efficient Servo Drive Movements

Christian Hansen, Kai Eggers, Jens Kotlarski, and Tobias Ortmaier

Institute of Mechatronic Systems, Gottfried Wilhelm Leibniz Universität Hannover, Hanover, Germany
E-mail: christian.hansen@imes.uni-hannover.de

Abstract – The energy demand of applications using inverter controlled electrical servo drives is considerably dependent on the applied trajectory profiles. In this paper, the improvement potential in supply energy consumption is analyzed with respect to the standard servo controller functionality. The optimal choice of trajectory profiles and execution times for point-to-point movements offers distinct energy savings. From the presented results, a number of general trajectory parametrization advices of even roughly known system characteristics can be derived. Even with little effort and without knowledge of exact system parameters, considerable energy savings are achievable.

Keywords – energy efficiency, servo inverter drive, optimal trajectory profile, energy losses

1 Introduction

Electrically actuated automation systems, e.g. in manufacturing and other industrial sectors, exhibit significant energy consumption. Since energy prices are continuously rising and political guidelines require improvements in efficiency to reduce CO\textsubscript{2} emissions, industrial companies are forced to reduce energy demands.

Electromechanical drives are utilized in versatile industrial applications, e.g. robotic manipulators, lift drives, or other material handling devices. Generally, the layout of the servo drive and the choice of hardware components (e.g. motors and inverters) preferably follow the mechanical loads to be actuated (masses, inertia, external forces, etc.) and the desired movement dynamics. A basic efficiency of the system is achieved by reasonable dimensioning (i.e. by avoiding unnecessary overload capacity) and an energetically compatible selection of hardware components, since available servo motors and inverters individually offer quite good efficiency.

Often, servo drives perform point-to-point (PTP) movements that are typically commanded using standard trajectory profiles (e.g. the Double-S-Velocity profile [1]) and are often driven with maximum dynamics. However, besides an energetically efficient design of the servo drive system, also the utilized trajectory characteristics have distinct influence on the consumed grid energy, since appearing energy losses in all power transmitting elements are significantly dependent on the sequence of operating points.

In [2], a reduction of energy losses in electrical drives is achieved with the application of an optimal control problem. A comparable approach is presented in [3], where a dynamic programming technique is used to reduce resistive copper losses in servo motor windings. However, friction losses in the application’s mechanics or any further source of energy dissipation in the servo drive are not considered. Furthermore, the implementation of such advanced control approaches is not state of the art in industrial applications. A selection of standard trajectory profiles is applied to different manipulators and energy demand and smoothness of jerk are evaluated in [1]. But, besides the effort, i.e. the absolute actuation torques, neither explicit consideration of different reasons for energy losses nor the possibility of energy recuperation are included.

The scope of this paper is the identification of possible improvement potential in energy demand by a proper choice of standard trajectory profiles and execution times with utilization of classical cascade control of current, velocity, and position [4]. Hereby, a wide range of different servo drive applications performing PTP tasks is addressed. The subsequent investigations deliver evidence that may be used as a guideline for proper movement parametrization of specific actuation tasks on different mechanical loads with utilization of standard hardware equipment and established path planning approaches.

The article is organized as follows. In section 2, relevant sources of energy losses in the servo drive hardware components are discussed and classified. Then, a variety of standardized trajectory profiles is revised in section 3 and the marked differences are outlined since these are relevant for the subsequent analysis. Section 4 presents versatile simulations for different PTP movements using various profile characteristics and system parameter combinations. Finally, section 5 subsumes the results and the extracted hints and advices for proper and intuitive movement parametrization.
2 Electromechanical servo drive

An electrical servo drive typically comprises a mechanical load that is actuated by an electromechanical motor (e.g. asynchronous or synchronous three phase machines). Due to the required speed and torque variable operation with high dynamics, the motor is operated by a programmable logic controller (PLC) with an inverted rectifier module consisting of IGBT-modules. The servo inverter is connected to a supply module via a DC link (see Figure 1). On the DC side, a certain electrical capacitance exists in the supply and the inverter module, and, therefore, a small energy storage capacity. The considered servo drives are capable of four quadrant operation and, consequently, recuperating electrical energy from the mechanical motion, e.g. during deceleration phases. In most applications, a brake chopper and a brake resistor are utilized to dissipate the recuperated energy. Special supply module types are capable of energy recovery to the power network but the are rarely used due to higher procurement costs.

\[ P_{\text{mot}}(t) = P_{\text{mot,DC}}(t) + P_{\text{mot,inv}}(t) + P_{\text{mot,mec}}(t). \]

The different sources of energy losses of the different system components are discussed in the following section.

2.2 Energy losses

With selection of high-quality components and reasonable dimensioning, high efficient servo drive applications can be realized. However, besides the mechanical load, all power transmitting elements of the servo drive also produce certain energy losses during operation. Manufacturers usually name the efficiency factors \( \eta_{\text{mech}} \) for drive operation at nominal rating. Typical values for three phase permanent magnet synchronous motors (PMSM) and power inverters are above 0.9 and 0.95, respectively [4]. However, the efficiency distinctly deviates for different operating points, defined by the motor velocity \( \dot{\varphi}(t) \) and motor torque \( \tau(t) \). Therefore, the resulting energy losses vary during the axis movement, depending on the operating point, the related electrical currents, the DC link voltage levels, the temperature, etc. The combined overall efficiency of the complete servo drive application and the calculation of the supply energy demand for given trajectories is subject of the following sections.

2.2.1 Mechanical load

The mechanical power \( P_{\text{mec}} \) of a rotational axis with total inertia of the motor axis \( J \) and transmission gears (if existing) is:

\[ P_{\text{mec}}(t) = \dot{\varphi}(t) [ J \ddot{\varphi}(t) + f_v \dot{\varphi}(t) + f_s \text{sgn}(\dot{\varphi}(t))] , \]

where the friction coefficients \( f_v \) and \( f_s \) represent all dissipative effects due to viscous damping and Coulomb
friction, respectively. Energy dissipation due to friction may arise in bearings, gear transmissions, sealing washers, oil bath lubrication, etc. As can be seen in (3), friction loss power is proportional to the drive velocity \( \dot{\varphi} \) for Coulomb friction and proportional to the square of the drive velocity \( \dot{\varphi}^2 \) for viscous damping.

### 2.2.2 Synchronous motor

In the considered three-phase electromechanical motor, where the mechanical movement is provoked by the interaction of magnetic fields in the rotating armature and the stator, different effects cause dissipation of energy and, hence, have impact on the motor’s efficiency. PMSMs possess a constant armature field, induced by high-performance magnets that are typically mounted on the armature’s surface. The stator field on the contrary is induced using copper windings and field-orientated current control. The motor input power \( P_{\text{mot}} \) is:

\[
P_{\text{mot}}(t) = P_{\text{mech}}(t) + P_{\text{t, mot}}(t),
\]

where \( P_{\text{t, mot}} \) combines all energy losses produced by the synchronous motor. A motor shaft rotation leads to continuous reversal of magnetism and, therefore, to iron losses due to magnetic hysteresis and eddy currents. Friction and windage losses in the motor are identically handled as described for the mechanical load, see section 2.2.1.

Since the motor torque \( \tau(t) = k_m i_{\text{mot}}(t) \) is approximately proportional to the motor current by the motor constant \( k_m \), resistive losses in the stator windings are assumed as proportional to the square of the motor current \( i_{\text{mot}}^2 \) and, therefore, proportional to the square of motor torque \( \tau^2 \). In contrast, the hysteresis and eddy current losses are approximately proportional to the motor shaft velocity \( \varphi \), and proportional to the square of the motor shaft velocity \( \varphi^2 \), respectively [7]. Further loss effects, e.g. due to harmonics, insufficient isolation, or field weakening are not considered since negligible small.

### 2.2.3 Power electronics

The total input power of the servo inverter \( P_{\text{inv}} \) is:

\[
P_{\text{inv}}(t) = P_{\text{mot}}(t) + P_{\text{t, inv}}(t) + P_{24V}(t),
\]

where the total servo inverter losses \( P_{\text{t, inv}} \) mainly result from resistive losses in conductor paths and switching losses in insulated gate bipolar transistors (IGBTs) during the pulse width modulated (PWM) three-phase voltage generation to feed the PMSM. Further demands are covered by a 24 V power supply \( P_{24V} \), e.g. for the PLC electronics, I/O-interface, or the optional mechanical motor holding brake.

Hence, resistive losses are considered proportional to the square of the total inverter current \( i_{\text{inv}}^2 \), which also has approximate proportionality to the motor torque \( \tau \), whereas IGBT switching losses are proportional to the constant PWM switching frequency, e.g. 8 kHz. Further losses in the servo drive system basically result from resistive conduction losses in the diode rectifier of the supply module and from leakage currents in the DC link capacitors due to limited isolation. However, experiments showed that these are comparatively small and, therefore, neglected in the following. Finally, also neglecting the small DC link capacitance, the supply module input power \( P_{\text{sup}} \), which represents the power consumption of the complete servo drive system, and the dissipated brake chopper power \( P_{\text{chp}} \) are:

\[
\begin{align*}
P_{\text{sup}}(t) &= P_{\text{inv}}(t) H[ P_{\text{inv}}(t) ], \\
P_{\text{chp}}(t) &= -P_{\text{inv}}(t) H[- P_{\text{inv}}(t) ],
\end{align*}
\]

where the Heaviside step function \( H \) is defined as:

\[
H[n] = \begin{cases} 
0, & n < 0 \\
1, & n \geq 0
\end{cases}
\]

### 2.3 Energy loss categorization

Based on the preceding sections, all decisive energy losses of the utilized servo drive components are classified in three groups: constant energy losses \( P_{\text{t, con}} \), motor speed-sensitive energy losses \( P_{\text{t, vel}} \) and torque-sensitive energy losses \( P_{\text{t, trq}} \). Hence, in comparison to (2), the total energy losses \( P_t \) can also be expressed as:

\[
P_t(t) = P_{\text{t, con}}(t) + P_{\text{t, vel}}(t) + P_{\text{t, trq}}(t).
\]

Table 1 subsumes the mentioned dependencies. To calculate energy losses of the servo drive components properly parametrized model equations have been defined and validated (see [6]), and are utilized in the following.

### 3 Trajectory profiles

In this research, PTP movements of single axis servo drives are addressed. Hence, the axis position profile \( \varphi(t) \) of a trajectory starting at time \( t_0 \) and ending at time \( t_1 \) is bounded by the following constraints:

\[
\begin{align*}
\varphi(t_0) &= \varphi_0, \\
\varphi(t_1) &= \varphi_1, \\
\dot{\varphi}(t_0) &= \dot{\varphi}(t_1) = \dot{\varphi}(t) = \ddot{\varphi}(t) = 0.
\end{align*}
\]

Hence, a variety of different trajectory profile types comes into consideration. To limit the number of profile variations and to comply with the requirements for practical implementation, some specifications are defined for the utilized trajectories. First, the position profile \( \varphi(t) \)
Table 1. Qualitative categorization of energy losses by proportionality to the operating point (ϕ, τ)

<table>
<thead>
<tr>
<th>Loss type</th>
<th>Speed</th>
<th>Torque</th>
<th>Const.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coulomb friction</td>
<td>∼ϕ</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Viscous damping</td>
<td>∼ϕ²</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Resistive</td>
<td>-</td>
<td>∼τ²</td>
<td></td>
</tr>
<tr>
<td>Hysteresis</td>
<td>∼ϕ</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Eddy currents</td>
<td>∼ϕ²</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Coulomb friction</td>
<td>∼ϕ</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Viscous damping</td>
<td>∼ϕ²</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Holding brake</td>
<td>-</td>
<td>const.</td>
<td></td>
</tr>
</tbody>
</table>

should be continuous up to the second time derivative \(\ddot{\varphi}(t)\), i.e. the acceleration. Hereby, the jerk is limited \((|\dddot{\varphi}(t)| \leq \dddot{\varphi}_{\text{lim}})\) to prevent the mechanics from damage and reduce wearing [8].

3.1 Trajectory profile types and categorization

The following list describes a selection of different trajectory profiles, that comply with the above mentioned requirements:

P1) **Trapezoidal acceleration profile**: With utilization of the trapezoidal acceleration profile (a.k.a. Double-S-Velocity profile), the time-optimal PTP movement is achieved, since always jerk, acceleration, or velocity limitation is exhausted during the complete movement [5, 8].

P2) **Sinusoidal profile**: For the sinusoidal profile trigonometric functions are used to describe the acceleration and deceleration phases and a linear polynomial for the constant velocity phase. The profile enables to specifically stimulate or suppress certain frequency bands to avoid vibrations and oscillations of the mechanical set-up.

P3) **Polynomial profile**: The polynomial function of degree 5 that is characterized by a simple mathematical expression that can easily be parametrized for different boundary constraints (see (9)).

All the named trajectory profile types are displayed in Figure 2. Here, the individual time-optimal profiles with minimum execution times \(T_{\text{def},i} = t_{P,i} - t_0\) with \(t_0 = 0\) s are given for an axis position displacement of 10 revolutions, when the velocity limit or acceleration limit, or both are exceeded. As can be seen, the trapezoidal acceleration profile has the shortest trajectory time \(t_{P1}\) while the sinusoidal \((t_{P2})\) and the polynomial \((t_{P3})\) profiles possess longer minimum-time since the dynamics limitations are not continuously exploited.

The intended trajectory profile analysis also requires the calculation of valid profiles also for specifically defined longer movement durations \(T_{\text{def}} > T_{\text{min}}\). In this context, it has to be pointed out that, for the trapezoidal acceleration profile (P1) two different methods can be applied to increase the movement duration for a given axis position displacement. For the first method (P1a), the time optimal trajectory profile is generated first. Then, a scaling factor is defined by the quotient of the defined movement duration \(T_{\text{def}}\) and the minimum trajectory time \(T_{\text{min}}\) as \(f_s = T_{\text{def}} / T_{\text{min}}\). With simultaneous scaling of the profiles in dimension of time using factor \(k_s\), for the position, velocity, and acceleration profiles follows:

\[
\varphi_{T_{\text{def}}} (f_s t) = \varphi_{T_{\text{min}}} (t) f_s^0 ,
\]

\[
\dot{\varphi}_{T_{\text{def}}} (f_s t) = \dot{\varphi}_{T_{\text{min}}} (t) f_s^{-1} ,
\]

\[
\ddot{\varphi}_{T_{\text{def}}} (f_s t) = \ddot{\varphi}_{T_{\text{min}}} (t) f_s^{-2} ,
\]

where \(\varphi_{T_{\text{min}}} (t)\) is the motor angle position profile for the minimum movement duration \(T_{\text{min}}\) and \(\varphi_{T_{\text{def}}} (t)\) is the motor position profile for the defined movement duration \(T_{\text{def}}\). The second method (P1b) generates a trapezoidal acceleration profile maintaining maximum acceleration but, instead, reaches a lower velocity level with rising movement duration.

Figure 3 displays the four resulting trajectory types (P1a, P1b, P2, and P3) with the same position displacement of 10 axis revolutions for a predefined movement duration of \(T_{\text{def}} = 1\) s. Apparently, all the resulting profiles fulfill the given displacement and duration requirements but obviously provide different movement characteristics. While the trapezoidal profile (P1b) and the sinusoidal profile (P2) still reach maximum acceleration values and, therefore, have lower velocities, the scaled trapezoidal profile (P1a) and the polynomial (P3) accelerate at lower levels but instead achieve higher velocity levels.

3.2 Expected influence on energy losses and supply

Based on the mentioned profile characteristics highlighted in the previous section, the relationship of torque and speed-sensitive as well as constant energy losses must be analyzed for the given servo drive application. Depending on the system parameters (friction coefficients, inertia, etc.) as well as the desired position displacement, the ratio of losses may change with the choice of trajectory profile type. Hence, for specific PTP tasks of different servo drive applications, the utilization of a particular trajectory profile type as well as an optimal movement duration should lead to the most efficient movement with minimal energy losses. In this context, the following three assumptions can be concluded:
1) Utilization of trajectory profiles (P1b) and (P2) leads to high torque-sensitive energy losses due to high acceleration values but moderate velocity-sensitive energy losses due to lower speed levels.

2) Utilization of trajectory profiles (P1a) and (P3) leads to high velocity-sensitive energy losses due to high speed levels but moderate torque-sensitive energy losses due to lower acceleration values.

3) Velocity (and acceleration) levels decrease with a longer movement duration and, therefore, speed and torque-sensitive losses are reduced, while constant energy demands are gaining in influence.

On the basis of the aforementioned assumptions, the existence of an energetically optimal trajectory profile type and movement duration is expected for a variety of different servo drive applications, possessing specific system parameters (e.g. friction coefficients, inertia, movement dynamics/position displacement). The examination of the listed assumptions is the objective of section 4.

4 Trajectory profile variation

In the following investigations, a PTP movement with a position displacement of $\Delta \varphi = \varphi_1 - \varphi_2 = 75 \cdot 2\pi$ is consulted. As described in section 2, the system model of a rotational servo drive axis with inertia as well as Coulomb and viscous friction effects is utilized. The parameters of the systems mechanics as well as the energy loss model parameters have been identified following standard parameter identification methods and verified using measurements that were presented in a previous publication [6]. The temperature level of current conductors and friction partners is assumed to be stationary in warmed up condition. The utilized system parameters and the dynamics limitations of the servo drive are collected in Table 2. The acceleration limit $\dot{\varphi}_\text{lim}$ for the trajectory profile calculation is obtained from $\dot{\varphi}_\text{lim} = \tau_{\text{lim}}/J$. Additional motor torque due to friction during the axis movement is accepted since the servo drive provides high overload capability $\tau_{\text{ult}}$.

For the given system dynamics, a minimum trajectory time of $T_{\text{min}} = 1.221 \, s$ results under utilization of the trapezoidal acceleration profile (P1). To simulate the energy demands for different situations, $n = 20$ particular movement durations $T_{\text{def},j}$ are defined using an exponential relation:

$$T_{\text{def},j+1} = T_{\text{min}} \cdot 10^{\frac{j}{n-1}}, \quad j = \{0, 1, \ldots, n-1\},$$

leading to a maximum servo drive movement duration variation beginning from $T_{\text{def},1} = T_{\text{min}} = 1.221 \, s$ up to $T_{\text{def},20} = 10 \cdot T_{\text{min}} = 12.210 \, s$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inertia</td>
<td>$J$ (kgm$^2$)</td>
<td>0.011</td>
</tr>
<tr>
<td>Coulomb friction</td>
<td>$f_c$ (Nm)</td>
<td>0.100</td>
</tr>
<tr>
<td>Viscous damping</td>
<td>$f_v$ (Nm/s)</td>
<td>0.002</td>
</tr>
<tr>
<td>Speed limit</td>
<td>$\dot{\varphi}_\text{lim}$ (rad/s)</td>
<td>628.32</td>
</tr>
<tr>
<td>Torque limit</td>
<td>$\tau_{\text{lim}}$ (Nm)</td>
<td>20.00</td>
</tr>
<tr>
<td>Ultimate torque</td>
<td>$\tau_{\text{ult}}$ (Nm)</td>
<td>30.00</td>
</tr>
</tbody>
</table>

The resulting velocity profiles of the four different profile variants for three exemplary movement durations ($T_{\text{def},1}$, $T_{\text{def},10}$, and $T_{\text{def},20}$) are displayed in Figure 4. Note that the polynomial (P3) and the sinusoidal profiles (P2) for fast movements ($T_{\text{def},1} \approx T_{\text{min}}$) violate the velocity and/or the acceleration limits and, therefore, are excluded. Also note that the trapezoidal acceleration profiles (P1a and P1b) are identical for $T_{\text{def},1} = T_{\text{min}}$. 
The following sections present simulation results showing the influence of the trajectory profile choice and the movement duration on the servo drive system’s supply energy demand. Therefore, in section 4.1, the different trajectory profile variations are applied to the validated virtual rotational axis using a variation of movement durations $T_{\text{def},i}$ as defined in (11). To motivate the transferability of the presented results, additional variations of other system parameters, e.g. the friction coefficients, the inertia, and the movement dynamics (by variation of the axis displacement) are presented in section 4.2.

### 4.1 Results for validated test rig model functions

According to the specifications defined in section 4, the trajectory profile types are generated for increasing movement durations for a fixed position displacement of 75 axis revolutions. Figure 5 shows the corresponding supply energy demand values. In addition, the absolute amounts of energy losses are given, separated for different causes (compare section 2.3): constant loss power, speed and torque-specific losses.

As can be seen in the upper plot, obvious differences in energy demand exist depending on the chosen profile type and the applied movement duration. For all utilized profile types, the energy demand possesses a minimum between the movement times from 4.6 to 5.2 s, which is approximately $\{3.77 \ldots 4.26\} \cdot T_{\text{min}}$. In the lower plot, the different sources of energy losses are given for all trajectories. Obviously, the high energy demands of fast manipulator movements are mainly caused by high torque-specific losses, induced by high currents in the conductor resistances of the servo drive system (compare Table 1). However, with increasing movement duration, the relation between torque and speed-specific losses changes and the speed-specific losses become dominant. With a view to the upper plot, the change in loss dominance, together with an increasing influence of constant energy losses, also involves a change of energetically optimal group of trajectory profile. Apparently, for faster movements the group of trajectory profiles (P1b and P2) are more efficient while for slower movements, the other group (P1a and P3) becomes favorable. Finally, the energy demand of the utilized profiles increases again with longer execution time due to the linearly rising influence of constant power demands (of electronics, brakes, etc.). The absolute minimum energy demand for the given PTP task and the rotational axis is achieved with a movement duration of $T_{\text{def}} \approx 4.26 \cdot T_{\text{min}} \approx 5.2$ s and utilization of the scaled trapezoidal acceleration profile (P1a), closely followed by the polynomial (P3).

### 4.2 Variation of system parameters

For the motivation of general statements about an advantageous choice of the profile type and the optimal
movement duration for a favorably wide range of different servo drive applications, the same calculations are repeated for a set of parameter variations. Therefore, the Coulomb friction and viscous damping coefficients $f_c$ and $f_v$ are incrementally increased using a factor $f = \{1, \ldots, 10\}$ in 20 equidistant steps, based on the validated system model parameters utilized in section 4.1. In this manner, the same energy demand analysis is accomplished for similar servo drive axes, that possess higher friction influence. The results are presented and discussed in section 4.2.1. Furthermore, section 4.2.2 presents the energy demand results for increased inertia of the mechanical load. Finally, a variation of movement dynamics is consulted in section 4.2.3 by increasing the desired position displacement under retention of the movement durations $T_{\text{def},i}$ defined in (11) at the beginning of this section.

### 4.2.1 Variation of friction coefficients

In comparison to the lower plot in Figure 5, the surface plots in Figure 6 present the proportion between constant, speed and torque-specific losses also with a variation of the system’s friction coefficients. As expected, the velocity dependent losses gain in importance with increasing friction parameters, while the torque-specific relations remain more or less untouched.

![Figure 6. Energy losses, separated for constant ($E_{\text{con}}$) as well as velocity- ($E_{\text{vel}}$) and torque-dependency ($E_{\text{trq}}$)](image)

As shown in Figure 7, for fast servo drive movements, the trapezoidal acceleration profile with maximum acceleration (P1b) (and for a smaller range the sinusoidal profile (P2)) demands the lowest energy supply (independent from the friction intensity). For systems with low friction influence ($f \leq 7$), the energetically optimal profile selection switches to the group of profile types P1a and P3 (here, to the scaled trapezoidal acceleration profile P1a) after an approximate movement duration of $T_{\text{def},i} \approx 2 \cdot T_{\text{min}} \approx 2.5 \text{s}$, since lower acceleration levels reduce the torque-specific energy losses. The yellow line highlights the optimal movement duration which increases with the friction variation.

![Figure 7. Supply energy demand $E_{\text{sup}}$ for different profile types, movement durations $T$ and friction coefficient variation by factor $f$ (bottom view)](image)

Since the surface plots of the supply energy results are relatively close, Figure 8 presents the relative difference in supply energy, to emphasize the possible savings potential by adequate choice of the trajectory profile. The results show that especially for fast PTP movements of the servo drive a high relative difference in supply energy demand exists of about 20 to 40\%. For slower movements a relative difference of 5 to 10\% remains, which also illustrates the high importance of a reasonable trajectory profile choice.

![Figure 8. Relative difference in supply energy demand between the most efficient against the mean value and against the worst case](image)
4.2.2 Variation of mechanical inertia

Figure 9 shows the supply energy demand results $E_{sup}$ for a variation of the mechanical inertia using a factor $f = \{1 \ldots 2\}$ in 20 equidistant discrete steps. For higher inertia many of the profile types reach the torque ultimate limits. As shown in the plot for fast axis movements, the same trajectory profiles, i.e. profile P1b followed by profile P2 are the most efficient choice. A change of the best profile with increased inertia parameter does not appear, since compared to the increased friction effects in section 4.2.1, the higher inertia leads to increasing torque-specific energy losses, which are smaller for trajectory profile group including types P1a and P3. Again, the optimal movement duration increases slightly with higher inertia, marked by the yellow line. The resulting values of relative energy demand show similar savings potential.

4.2.3 Variation of axis dynamics

The last simulation example presents the supply energy demand results and corresponding optimal trajectory profile types for an increasing axis displacement for the given variation of movement durations (see Figure 10). Since the values of minimum movement time increases with the desired displacement $\Delta \varphi$, a rising number invalid trajectories for small times $T_{def,i}$ appears with increasing displacement. The factor $f = \{1 \ldots 3\}$ is incrementally increased in 20 steps. Rising axis displacement leads to higher velocity levels and longer acceleration phases, to cover the longer distances within the given movement duration. This leads to both, increased speed and torque-specific energy losses, while the constant losses remain on their ratio. The resulting map of optimal trajectory profiles is similar to the previous one (compare section 4.2.2).

4.3 Discussion of Results

Although during the presented simulation, different system parameters are variegated, comparable trends become apparent. As can be seen in Figures 7, 9, and 10, for fast PTP movements (close to $T_{min}$) of the given rotational servo drive axis, the most efficient trajectory is the trapezoidal acceleration profile (P1b). However, for all test series, the optimal profile type switches to the sinusoidal profile (P2) after an approximate extension of the movement duration by factor 2 ($T_{def} \approx 2 \cdot T_{min}$). With slower movements, the energetically optimal trajectory profile finally switches to the scaled acceleration profile (P1a), with exception of parameter configurations with higher friction influence. Here, the dominant source of energy loss changes, resulting in a preference of profiles, that achieve lower velocity levels. The polynomial trajectory profile only offers best system efficiency in exceptional cases (not shown).

Furthermore, there exists a clear trend for the energetically optimal movement duration that is similar for all parameter configurations. For all test series, the optimal movement duration for the given PTP task (including the increased position displacement in section 4.2.3) appears in a range between $T_{def} = \{3.5 \ldots 4.5\} \cdot T_{min}$. Of course, the factor is dependent on the amount of constant power demands effecting the raise of energy supply for longer movement duration. In all these cases, the optimal profile type was the scaled trapezoidal acceleration profile (P1a).
5 Conclusion

The influence of different standard (jerk bounded) trajectory profiles for PTP tasks on the total supply energy demand of servo drive applications has been presented. Since with different trajectory profiles, the servo drive’s operating points (motor speeds and torques) during the movement change, different causes of energy losses are evoked to a greater or lesser extent. Furthermore, the influence of constant energy losses on the complete system’s energy balance must not be neglected. Based on a validated system model parametrization for a rotational servo drive axis, different parameters have been adapted for variable friction coefficients, load inertia, and movement distances.

The results show the significant difference in total energy demand of a servo drive while changing the motion profiles and the movement duration within reasonable limits. Due to the influence of constant energy demands, there exists an energetically optimal movement duration which is different from minimum time.

During the parameter variation series, repetitious characteristics became apparent, resulting in a number of general advices for the choice of trajectory profile. For all examined parameter series, a constant order of favorable profile types turned out as energetically optimal with increasing movement duration, only with exception of the case where a chance of the dominant energy loss category appears. Here, a deliberate choice of trajectory profile is recommended. Generally, already slightly increased execution times lead to the energetically optimal movement with considerable savings and, hence, to efficiency improvement.

Future works will focus on the examination of additional parameter variations and the combination of different parameters to variate. Furthermore, a wider range of different servo drive applications must be investigated, e.g., by the inclusion of different non-linear system mechanics and the influence of external forces.

Acknowledgment

The authors gratefully acknowledge support by research grants from the German Research Foundation (DFG; OR 196/4-2) and reviewers’ comments.

References


Implications of Legal Frameworks on Construction and Demolition Waste Recycling – a Comparative Study of the German and Australian Systems

C.-R. Wonschik\textsuperscript{a}, J. Brennan\textsuperscript{b}, G. Ding\textsuperscript{b}, A. Heilmann\textsuperscript{a} and K. Vessalas\textsuperscript{b}

\textsuperscript{a}Faculty of Automation and Computer Sciences, University of Applied Sciences, Germany
\textsuperscript{b}School of Civil Engineering, University of Technology Sydney, Australia

Email: rwonschik@hs-harz.de, Jane.Brennan@uts.edu.au, Grace.Ding@uts.edu.au, aheilmann@hs-harz.de, Kirk.Vessalas@uts.edu.au

Abstract -

This comparative study between German and Australian legislation demonstrates that legal frameworks impact on the way in which recycling systems work. Both Australia and Germany operate as Federations and the autonomy of states influences common federation wide practices and standards. In Germany's case however, it is obliged to comply with European Union guidelines which result in German federal legislation being binding for all German states and to common industry practices across all of Germany. Purely industry regulated systems are not always sufficient to cater for societal and environmental needs, and political intervention can sometimes be necessary to achieve desired outcomes. The construction and demolition (C&D) waste recycling industry is a good example. In Australia C&D waste recycling is mostly industry regulated, while the state has greater influence in Germany. A statistical analysis illustrates legislative impact on recycling outcomes. Nonetheless, any legislative efforts can also have effects contrary to the intended ones. A study of such cases is conducted and other influencing factors also considered. In conclusion, the study outlines the importance of interstate coordination and regulation; and the need for the incorporation of industry requirements and other potentially influencing factors into the legal frameworks in order to meet desired outcomes.

Keywords -

Construction and Demolition (C&D) waste; recycling; legislation; Australia; Germany

1 Introduction

The intense activity in the construction sector during the last decades worldwide has generated huge amounts of construction and demolition (C&D) waste. C&D waste includes a wide range of mostly inert materials, such as: maintenance materials, road construction and excavation materials; but can also contain hazardous waste types such as asbestos, PCBs or PAHs, which can be present in significant proportions when buildings are renovated or demolished.

It can be assumed that the differences in the amount of C&D waste between different countries derive from the: differences in building tradition; poor quality of available data; unequal levels of control and reporting of C&D waste in States; legislative systems; and differences in definitions [1].

In an attempt to correct the serious effects which the C&D waste can have on the environment, important developments have been incorporated into International and European laws, which aim to promote the culture of reuse and recycling [2]. Furthermore the recovery of waste is an opportunity for the protection of the environment. It mitigates climate change by saving energy during the material recovery, conserving natural resources by substitution and preventing illegal dumping. For example in the Sydney basin, there is a need to substitute virgin sand by other materials in the process of concrete production due to a now limited availability. The recycling and processing of C&D waste to substitute natural sand has been found to be a suitable alternative.

The following paper reports on a comparative study between the Australian and the German legislative systems and their effects on recycling rates and conservation of resources and embodied energy.
2 Objects of investigation

Australia and Germany are both well developed countries. Nonetheless, there are differences between the two in the field of waste management. Table 1 and Table 2 gives an overview of the stages of development of both countries. Since Australia and Germany have very similar human development index and gross domestic product (GDP) values [3], their stages of development are comparable. However, there is a considerable gap between the two countries’ waste recycling outcomes, particularly with regard to C&D waste recycling, although comparable recycling technologies are used. For the same time period in Germany 88 percent of C&D waste was recovered and reused in further applications or recycled, whereas in Australia this was only the case for 55 percent [4]. It is noticeable that Germany, in contrast to Australia, has adequate natural resources such as construction sand available within feasible transportation limits.

Table 1 Country Data of Australia 2008-2012 [3, 4]

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (Mill.)</td>
<td>22.919</td>
</tr>
<tr>
<td>Population urban (%)</td>
<td>89.4</td>
</tr>
<tr>
<td>Human Development Index [0-1]</td>
<td>0.938</td>
</tr>
<tr>
<td>GDP per capita (US $)</td>
<td>34,548</td>
</tr>
<tr>
<td>Carbon dioxide emissions per capita (tonnes)</td>
<td>18.6</td>
</tr>
<tr>
<td>C&amp;D waste (Mill. tonnes)</td>
<td>19.00</td>
</tr>
<tr>
<td>C&amp;D waste per capita (tonnes)</td>
<td>1.206</td>
</tr>
<tr>
<td>C&amp;D waste recovery rate</td>
<td>55.26</td>
</tr>
</tbody>
</table>

Table 2 Country Data of Germany 2008-2012 [3, 4]

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (Mill.)</td>
<td>81.991</td>
</tr>
<tr>
<td>Population urban (%)</td>
<td>74.1</td>
</tr>
<tr>
<td>Human Development Index [0-1]</td>
<td>0.920</td>
</tr>
<tr>
<td>GDP per capita (US $)</td>
<td>34,437</td>
</tr>
<tr>
<td>Carbon dioxide emissions per capita (tonnes)</td>
<td>9.6</td>
</tr>
<tr>
<td>C&amp;D waste (Mill. tonnes)</td>
<td>201.47</td>
</tr>
<tr>
<td>C&amp;D waste per capita (tonnes)</td>
<td>2.457</td>
</tr>
<tr>
<td>C&amp;D waste recovery rate</td>
<td>88.00</td>
</tr>
</tbody>
</table>

The study presented in this paper investigates the impact of different legislative systems on C&D waste recycling rates. It will also consider how far international and national legislation applies to the national recycling system(s).

As a general measure, the recycling system is divided into three stages. The first stage represents the collection and salvaging operations of C&D waste. The second stage focuses on the C&D waste treatment, whereas the third stage includes products, applications and disposal. The simplified recycling system model is illustrated in figure 1.

Figure 1 Simplified recycling system model

3 Australian & German legislation

3.1 Germany

Construction and demolition waste management in Germany is a mature and well integrated sub industry within the broader German construction market.

In 2011, German construction and demolition activity generated 201.47 Megatons of waste, which is composed of two thirds of excavation material, approximately one third of building and road demolition waste, and a smaller fraction of 2 percent of mixed construction site waste. Despite these large amounts of waste, only 12 percent of this material was disposed of in landfills, while the remaining 88 percent was recovered and reused in further applications or recycled [5]. In Germany, C&D waste is normally collected source-segregated as a result of selective demolition. The separation of materials at the demolition site through selective demolition or other means is often the most effective way to ensure a clean, uncontaminated product [6].

Germany’s high material, energy, labour and waste disposal costs favour the economics of recovering, reusing and recycling as much C&D waste as possible. In 2011 the costs for landfill were between 100 to 170 Euro per tonne (appr. AUD 150-250), expect for excavated soil, which has a lower cost. These high landfill costs stem from strict regulative environmental specifications in Germany such as the regulation on waste dumps and long-term deposits. Hence, a multi-barrier concept preventing negative environmental impacts is required for landfill operations. Additionally, strong waste management systems have long been required by laws and regulations at all levels of government in order to minimise the impact of C&D waste in the waste stream [4].

The Federal Republic of Germany is composed of 16 states (Länder) with relatively broad powers and responsibilities over their regions. Generally, the
Federal level of government establishes laws that the Länder must implement and administer. These federal laws and regulations are mainly based on international and European legislation. The Basel Convention and the Kyoto Protocol are the most important international agreements influencing the German legislation waste management.

The Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and their Disposal, short Basel Convention, is an international agreement established on 22 March 1989 to initiate a transnational waste management framework. It sets rules to control transboundary movements of wastes that are hazardous to human health and the environment, and their disposal at an international level. The Convention became effective on 5 May 1992 and has 180 countries participating today (May 2013), including Australia and the European Union (EU) [4]. The goal of the Basel Convention is to minimise the movements of certain hazardous wastes, which can also be found in C&D wastes. The Conventions also strive to reduce the amount of the generated hazardous wastes, and its toxicity, by assisting developing countries directly on the source of generation. [7]

The Kyoto Protocol was passed on 11 December 1997 and is an additional protocol for the configuration of the climate convention framework of the UNFCCC. It came into force on 16 February 2005 and sets binding goal values for the emission of greenhouse gases within the developed countries. As at 2011, 193 states worldwide, including Germany and Australia, had signed the Kyoto Protocol. [7]

In addition, the resolution of many environmental issues in Germany, as a member of the European Union (EU), has to comply with EU standards which are implemented through national laws. This includes dealings with hazardous and non-hazardous waste within the European Waste Legislation. Along with this, the European environmental legal frameworks contain guidelines and acts for various topics, including a Waste Shipment Act, an End Of Life Vehicle Directive, a Package Directive, a Sludge Directive, and a Waste Oil Directive. By establishing an overall guideline for the handling of waste, the EU contributes to the development of high standards of waste management in Europe. Besides the guidelines, there are also compulsory directives, which regulate and monitor the management of waste within the EU. The most important laws and acts are described below.


It contains important subsections regarding the hierarchy of waste (figure 2), the waste management, permits and registration, plans and programs for future waste treatments. It defines and determines key policies and concepts for waste and its recovery and recycling. The goal of the Directive 2008/98/EC is to preserve and protect the environment, to diminish the harmful effects waste has on the environment and human health within the EU. [8]

![Figure 2 European waste hierarchy](image)

The Directive 2008/98/EC came into force on 12 December 2008 and had to be implemented within the following two years.

The Regulation (EC) No 1013/2006, regulates and monitors the movements of waste between the member states of the EU, as well as the importation and export of waste to and from countries outside of the EU. It also directs and observes the notification procedures, giving a clear overview about the type of waste treatment (recycling or disposal), the waste movement directions (import or export), the role of the affected states (sending, receiving, or transferring state), or the type of waste (for example hazardous or non-hazardous wastes). The regulation was opened to signature on 14 June 2006 and entered into force on 12 July 2007. [8]

Another important Regulation of the European Parliament is the Regulation (EC) No 2150/2002, it establishes a legal framework for the creation of statistics on waste within the EU. The goal is to provide the EU and all its member states with continual, comparable and reliable data streams in order to regulate and supervise the transposition of the acts and directives on waste management established by the European Parliament and the Council. It provides data streams and elaborates statistics on the production, the recovery and recycling, and the disposal of waste within the EU by collecting and estimating data and by claiming statistics. [8]

Besides the more general regulations and directives mentioned above, which Germany is obligated to implement in domestic legislation by law, there are
several other EU efforts which aim to regulate specific areas in waste management. For example, the Council Directive 1999/31/EC requirements for landfills, specifically on surface water, groundwater, soil, air and human health, as well as the Directive 2000/76/EC on preventing or reducing air, water and soil pollution caused by the incineration or co-incineration of waste, or the European Commission’s Green paper on the management of organic waste.

In order to comply with all European directives and regulations, the Recycling and Waste Management Act of Germany, the so called Gesetz zur Förderung der Kreislaufwirtschaft und Sicherung der umweltverträglichen Bewirtschaftung von Abfällen, short Kreislaufwirtschaftsgesetz (KrWG), was enacted. This Act promotes closed substance cycle waste management and ensures environmentally compatible waste disposal. Enacted in 2005, the Recycling and Waste Management Act of Germany bans untreated waste from being disposed. This has led to higher recycling rates, since this prohibition pressed for the development of a functional recycling system. In contrast, the economic influence of landfill fees on the C&D recovery rate appear to be only marginal. Figure 3 illustrates this using the example of Hamburg. It shows that disposal costs have only a minor influence on the high waste recovery rates. Other drivers such as limitations in landfill space or the preservation of resources had a more significant impact. Landfill levies to avoid the disposal of C&D waste have never been charged in Germany. Nevertheless, over a long period the landfill costs were high and contributed to the high C&D waste recovery rates baseline.

The Recycling and Waste Management Act of Germany is complemented and specified by several ordinances and acts, for example:

- regulation on European waste listing
- regulation on requirements regarding removal and usage of mature timber
- regulation on the disposal of commercial wastes and of certain building and demolition wastes
- regulation on refuse economy concepts and waste balances
- PCB/PCT Waste Ordinance
- regulation on dumps and long-term camp
- regulation on the storage of waste

The last two items concerning the dumping and storage of waste highlights the specific nature of the legislation which even deals with the distinction between dumping and backfilling. For example backfilling of excavated soil in open cut mines, is a recovery option, because the material is used to stabilise the quarry, which is seen as utilisation and not disposal. In practice however, these regulations can have negative effects. For example, a recycling facility usually has to pay for the backfilling of treated material thus enabling the open cut mines to offer their excavated virgin material and resources at lower prices than they otherwise would have been able to. This in return leaves recycled materials more expensive than virgin materials which can severely impact on German recycling facilities and the markets for their products.

In general, looking at the simplified recycling system, illustrated in figure 1, it can be assumed that the federal German governmental legislation focuses mainly on the 1st and 3rd stage. The primary responsibility for ensuring the proper treatment of C&D waste, influencing the 2nd stage of the simplified model, is mainly in the hands of local authorities, but is also partially legislated by the German government and the Länder (e.g. immission control law (BImSchG), Environmental Impact Assessment Law (UVPG)). Meanwhile, the 16 Länder are responsible for the implementation and establishment of more specific extensions to the waste management legislation. They are also responsible for the enforcement of regulations meant to achieve C&D waste goals set by higher levels of government, particularly the Federal government of Germany and the EU and to oversee the proper operation of waste treatment and disposal facilities.

At the local level, municipalities are responsible for the administration and issuing of demolition and construction permits that now occasionally include detailed deconstruction plans and detailed recycling specifications of the building’s materials. While local authorities are responsible for arranging the collection, recycling and disposal infrastructure of household waste, commercial waste such as C&D waste is solely the responsibility of the waste’s owners. The local authority ensures this responsibility is met according to federal and state legislation, and is responsible for initiating prosecution against offenders. Any commercial use of the processing, recycling and disposal infrastructure
operated by the municipality is paid for directly by the user.

3.2 Australia

Construction and demolition waste management in Australia is a young and moderately integrated industry within the broader Australian construction market. In 2011, Australian construction and demolition activity generated 19 Megatons of waste [4]. Currently, no consolidated data about the specific composition and origin of C&D waste is available at the national level. Only 55 percent of C&D waste was recovered and reused in further applications or recycled and 45 percent of this material was disposed of in landfills [4].

Compared to Germany, Australia has a considerably lower material and energy recovery rate. In the greater area of Sydney, for example, C&D waste is often collected as mixed waste. Lower waste disposal costs appear to limit the economics involved in recovering, reusing and recycling C&D waste. In 2011 the costs for landfill disposal were between AUD $40 to AUD $130 per tonne [4]. The landfill design is comparable to the German multi-barrier concept.

One likely explanation for this situation is the fact that the Australian Government does not directly legislate the management of C&D waste [9]. The management of environmental issues, including all waste streams, is largely the responsibility of Australian state and territory governments. Exceptions to this general principle are where international treaties are involved (i.e. the Basel Convention, Kyoto Protocol) or developments are deemed to be of significant environmental importance to the nation [4].

Waste management and resource recovery in Australia is dependent on the regulatory framework of a particular State or Territory. Thus, the approach commonly adopted by the Australian Government is one of multi-stakeholder engagement and the introduction of multi-party agreements. These may be supported by underpinning legislative measures in instances where all parties support the need for such fall-back legislation at a jurisdictional level [10].

Looking at the simplified recycling system, illustrated in figure 1, it can be seen that the federal Australian governmental legislation focuses mainly on the 3rd stage. Examples of this kind of legal framework are the Australian and New Zealand Government Framework for Sustainable Procurement and the National Road Pavement Guidance. Both policies regulate product design and associated specification, manufacture and application [10], but influence the 1st and 2nd stage of the recycling system only indirectly. The first two stages of C&D waste recycling are mostly industry regulated.

C&D waste recycling rates are strongly dependent on State or Territory legislation. Table 3 illustrates the resource recovery rates and future targets by jurisdiction of the 8 Australian states. It is noticeable that some states and territories (e.g. New South Wales (NSW), South Australia (SA), Australian Capital Territory (ACT)) have significantly higher recovery rates than the other states. On closer observation a link between legislation and resource recovery can be identified. A comparison of South Australia and Western Australia legislation documents this in a remarkable way.

The state of South Australia has one of the highest developed legislative systems with respect to C&D waste management in Australia [cf. 9]. For instance the Zero Waste SA Act of 2004 influences the C&D waste management in SA by promoting C&D recycling infrastructure and business development for recycling C&D waste materials. This results in increased C&D waste avoidance and resource recovery by C&D waste generators, and C&D waste recovery by reprocessors and landfill operators.

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>VIC</th>
<th>QLD</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>recovery rate</td>
<td>73%</td>
<td>53%</td>
<td>37%</td>
<td>77%</td>
</tr>
<tr>
<td>Target recovery</td>
<td>76%</td>
<td>80%</td>
<td>50%</td>
<td>90%</td>
</tr>
<tr>
<td>rate</td>
<td>by 2014</td>
<td>by 2014</td>
<td>by 2014</td>
<td>by 2015</td>
</tr>
<tr>
<td></td>
<td>ACT</td>
<td>WA</td>
<td>NT</td>
<td>TAS</td>
</tr>
<tr>
<td>recovery rate</td>
<td>81%</td>
<td>29%</td>
<td>&lt;1%</td>
<td>15%</td>
</tr>
<tr>
<td>Target recovery</td>
<td>Overall</td>
<td>100%</td>
<td>Overall</td>
<td>No</td>
</tr>
<tr>
<td>rate</td>
<td>80%</td>
<td>by 2015</td>
<td>Inert</td>
<td>specific targets</td>
</tr>
<tr>
<td></td>
<td>(no</td>
<td>waste</td>
<td>(no</td>
<td>specific</td>
</tr>
<tr>
<td></td>
<td>specific</td>
<td>(mainly</td>
<td>specific</td>
<td>C&amp;D</td>
</tr>
<tr>
<td></td>
<td>C&amp;D</td>
<td>target)</td>
<td>C&amp;D</td>
<td>target)</td>
</tr>
</tbody>
</table>

In contrast to the situation in SA, the Government of Western Australia established a comparable act (Waste Avoidance and Resource Recovery Act) like SA with strict division targets and associated regulations commencing 1 July 2008 – four years later than SA. This lag in time causes significant differences in recycling rates between the two territories, because it normally takes several years, before additional recycling capacity reaches the market.

Another reason for the lower recovery rates is the enormous range of landfill levies within the eight Australian states and territories. Levies have been created as an additional charge and instrument to the normal tipping fee ($15-$30), and to enhance the recovery and
recycling efforts. Waste levies correct a market failure by making recycling cost competitive against landfill disposal [11]. These landfill levies exist in NSW, Victoria, SA and ACT. The other states either scrapped a levy regulation altogether (Queensland (QLD)), made it voluntary (Tasmania (TAS)), limited it to special regions (Western Australia (WA)) or never adopted a levy regulation in the first place (Northern Territory (NT)). Table 4 illustrates the current status of the landfill levy across Australia. The influence of an increasing levy on the C&D recovery rate is illustrated in figure 4. It shows that high disposal costs can be associated with a high landfill levy leading to high waste recovery rates [13]. A point in case is Germany, where the high landfill costs have led to one of the highest C&D waste recovery rates worldwide.

Table 4 Actually Landfill Levy Rates [13]

<table>
<thead>
<tr>
<th></th>
<th>NSW</th>
<th>VIC</th>
<th>QLD</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levy (AUS$/tonne)</td>
<td>122.40 (rural &amp; urban)</td>
<td>26.60 (rural, municipal)</td>
<td>no levylevy (levy of $35.00 scrapped in 2013)</td>
<td>23.50 (rural &amp; urban)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>46.60 (rural, industrial)</td>
<td></td>
<td>47.00 (metropolitan)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>53.20 (metropolitan &amp; provincial)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>121.90 (com. waste)</td>
<td></td>
<td>12.00 (only for waste from Perth)</td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td></td>
<td>no levylevy by the 3 regional waste groups (approx: 2.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4 Discussion

From the various values presented in the previous sections, it can be seen that there is a direct correlation between high C&D disposal costs and legislative incentives such as the landfill levy in Australia. Figure 5 shows clearly that the recovery rates are directly proportional to the landfill levy imposed in different states, apart from Southern Australia which has well integrated legislation for C&D waste recycling.

It does however have to be noted that there might be a slight distortion of figures due to “exporting of waste” to other states in order to avoid high levies. For example, it is common to dispose NSW waste in Queensland landfills, which prove cheaper even when considering transportation costs. This highlights the issue of inconsistent legislation across state boundaries. The German system is more consistent due to overarching European legislation and directives; Australia would require federal guidelines to overcome issues such as “waste exporting” to other states which can be counterproductive to the goal of increasing recovery rates. In Germany on the other hand, the use of virgin material is partially increased due to backfilling operations that provide revenue to open cut mines and hence causing prices of virgin materials to be lower than recycled materials. This is also counterproductive to the goal of reusing C&D waste and makes recycling of this kind of waste less viable. These examples demonstrate the need to develop consistent and overarching legislation with the policy goal of ensuring higher recovery rates and reuse of recycled C&D waste materials and its associated products. Regular revision of legislation is required under consideration of the current situation in the C&D waste market; this will also require consistent data collection across states in Federal countries such as Australia and countries, such as Germany, in legislative unions such as...
the EU. Federation can cause tension when individual states insist on their own ways of managing their C&D waste, which can prove counterproductive to other states’ efforts. The impact of Germany’s participation in the EU means that it is bound to European directives and legislation, which provides a consistent system across all German states.

5 Conclusion

The study demonstrates that the demolition of buildings and civil infrastructure usually create significantly more waste material than construction activities. Traditional demolition practices in which all building materials are mixed together create a waste stream which is difficult and costly to recycle has been almost eradicated in Germany today. Local landfill fees have a strong influence on the economic viability of alternative demolition practices. In the absence of strong government regulations, as in the case of Australia, landfill fees are the primary factor influencing demolition processes.

While the recycling processes used to sort mixed materials and reduce contamination vary between countries and regions, it is clear that the separation of materials at the demolition site through selective demolition or other means is often the most effective way to ensure a clean, uncontaminated product. To enhance the selective demolition and collection of C&D waste, national laws, acts and guidelines are necessary. However, demolition materials can sometimes be mixed together depending on the use identified for them. Concrete, bricks and ceramics for instance, are often ground up together for use as fill in road building or other construction projects. Such a use allows these materials to be mixed during the demolition process, although other materials may still need to be sorted out.

References


[13] Ramsay P. J. Waste Authority of Western Australia - Review of the Regulatory Framework for the
Predicting Energy Usage Using Historical Data and Linear Models
Majeed Safa\textsuperscript{a}, Jeremy Allen\textsuperscript{b}, and Mahdi Safa\textsuperscript{c}

\textsuperscript{a}Department of Agricultural Management and Property Studies, Lincoln University, New Zealand
\textsuperscript{b}Energy Solution Providers Ltd, Auckland, New Zealand
\textsuperscript{c}Department of Civil and Environmental Engineering, University of Waterloo, Canada
E-mail: Majeed.safa@lincoln.ac.nz, jeremy.allen@espnz.co.nz, msafa@uwaterloo.ca

Abstract:
This paper presents a method to predict energy usage, based on weather conditions and occupancy, using a multiple linear regression model (MLR) in research office buildings. In this study, linear regression models of four research office sites in different regions of New Zealand were selected to show the capability of simple models to reduce margins of error in energy auditing projects. The final linear regression models developed were based on monthly outside temperatures and numbers of full time employees (FTEs). Comparing actual and predicted energy usage showed that the models can predict energy usage within acceptable errors. The results also showed that each building should be investigated as an individual unit.

Keywords: Energy prediction, energy auditing, linear regression model, office buildings, energy saving, optimization

1. Background:
Recent years have witnessed a remarkable increase in energy costs and increasing environmental concerns [1]. The optimal operation of office buildings is essential for reducing energy. Optimizing the energy consumption of office buildings requires robust energy monitoring and energy auditing systems [2]. In many commercial projects energy savings are estimated based on a comparison of current usage and the previous years data [3, 4]. Due to environmental conditions, changing occupation and other causes, energy usage in different years could change significantly. Therefore, modelling energy usage based on historical data has the potential to provide more accurate energy saving estimations. The objective of this study was the development of a new forecasting model to predict office building energy usage to measure energy savings during energy auditing projects.

Offices and the retail sector are the most intensive energy consumers in the non-domestic building sector and it is estimated that over 50% of energy usage was for non-domestic buildings [5]. In New Zealand, standard office buildings mainly consume electricity for their operation. A survey of energy sources by the Building Research Association of New Zealand (BRANZ) showed that 11% of non-residential buildings consumed natural gas while only 3.5% of non-residential buildings used diesel and/or fuel oil, mainly for heating [6]. The major electrical consumption in the buildings surveyed was for lighting, air-conditioning, plug loads and hot water [5, 6].

There have been several studies that analysed electricity consumption in different buildings. However, due to the lack of current advanced metering technologies, in most early energy modelling studies monthly electricity bills were used to analyse energy consumption in buildings [7].

As the first step to analyse energy usage in buildings, it was important to understand which factors were more important for energy consumption in different buildings. Kavousian, et al. [7] identified four main categories for energy consumption in residential buildings: weather and location, appliances and electronic stock, physical characteristics of the buildings, and occupancy. Nonetheless/However??, Issacs, et al. [6] classified the businesses in their survey based on business sector and activities, staff numbers, client numbers and operating periods.

Office buildings are very different in terms of their design, construction, occupancy and activity, which makes it too difficult to identify small numbers of them to represent the majority of office buildings [8]. Comparing different studies showed that environmental parameters were the main factors in most studies for estimating energy usage in buildings. Temperature, humidity and lux level were the main environmental parameters which can directly affect energy consumption in buildings.
In energy auditing projects and sustainability plans, predicting energy usage is a very important challenge. Several optimization methods have been applied for the energy usage estimation and a variety of modelling methods have been developed over the last decade [10-18].

2. Models:

In this study, four research buildings, including offices and laboratories, of the same research institute in different New Zealand regions, were investigated (Figure 1). Each site was investigated individually based on available historical monthly data.

The multiple linear regression model was selected and developed for predicting energy consumption. This model has been extensively used in energy modelling projects [9, 19-25]. Compared with nonlinear models, linear regression models are easier and more practical for solving the different problems [21].

For use in the model, it was necessary to select a limited number of variables without any selective bias [26]. A simple model with the highest $r^2$ was designed through a combination of forward, backward and stepwise regression adjustments. Terms were maintained in the final model if they were significant at $p=0.05$ [27]. In the first step, the relationship between energy consumption and each input variable was tested with a simple linear regression using the $r^2$ as the decision criterion. Then, a multiple linear regression model was developed for predicting the energy consumption as:

$$Y=a_0+a_1V_1+a_2V_2+\cdots+a_nV_n$$  (1)

where $a_0-a_n$ are the regression coefficients and $V_0-V_n$ are the independent variables.

The model was in a linear form to represent the linear relationships of the dependent variable with the independent variable. After running the model, predictions on the validation data were estimated. In this study, the model was developed with a minimum number of variables to capture energy consumption in as simple a form as possible. After investigating several variables, occupancy (full time employees) and outside temperature were selected as the independent variables with which to develop the models. The monthly temperature was collected from a national weather database (NIWA) and property managers provided the number of employees.

Several methods of error estimation were proposed. The mean square error (MSE) over all training patterns (Eq. 2) was the most commonly used error indicator. MSE was very useful to compare different models; it showed a network’s ability to predict the correct output. The MSE can be written as:

$$MSE = \frac{1}{2N} \sum_{i=1}^{N} (t_i - z_i)^2$$  (2)

where $t_i$ and $z_i$ are the actual and predicted outputs for the $i^{th}$ training pattern, and $N$ is the total number of training patterns [21, 28]. The root mean square error (RMSE) is another error estimation method, which shows the error in the units of actual and predicted data.

In this study, several models were developed and compared to find the best fit between the predicted and actual data. As shown in Figure 2 the final model of the Palmerston North, Christchurch, and Dunedin sites were developed based on 30 months of available historical data and the model of the Hamilton site was developed based on 18 months of data. The Christchurch model was affected by the 4 September 2010 Canterbury earthquake, which significantly reduced energy usage during September and October; therefore, the data of September and October 2010 were removed from the model. It was notable that the models developed, based only on outside temperature for all sites, were accurate with a very high correlation coefficient between actual and
predicted energy usage. However, comparisons between models with one independent variable (temperature) and two independent variables (temperature and FTEs) showed that the FTEs can improve the correlation coefficients between actual and predicted energy usage.

You need to label these figures from A to D

![Graph A](image)

![Graph B](image)

![Graph C](image)

![Graph D](image)

Figure 2. Predicted and actual energy usage in the four research offices.

3. Results

A comparison of the models showed that the monthly energy usage in research office buildings can be predicted by temperature and occupancy data. The energy usage mostly depended on weather conditions. The low usage in summer would be influenced mostly by outside temperature and Christmas and New Year holidays (summer in the Southern Hemisphere). It should be noted that energy usage in buildings would also be affected by a number of other factors.

Figure 2 shows that the predicted and actual data are matched in most months; however, for different numbers of days per month, for different numbers of holidays per month and some other factors can influence the model and the predicted data.

Multiple linear regression models could be fitted to the energy consumption data and accounted for around 89%, 81%, 79%, and 76% of the variance in the sites (a to D) studied, respectively (Figure 3). The final RMSEs were calculated as 2751.6, 9508.2, 1904.6, and 1411.8 kW.h for Palmerston North, Christchurch, Dunedin, and Hamilton research office buildings, respectively (Table 1).

<table>
<thead>
<tr>
<th>Site</th>
<th>Equation</th>
<th>MSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmerston North</td>
<td>$87561.6+ (31.68 \times \text{Tem}) - (3405.34 \times \text{FTE})$</td>
<td>7571201.0</td>
<td>2751.6</td>
</tr>
<tr>
<td>Christchurch</td>
<td>$262840.9-(5830.67 \times \text{Tem}) - (337.17 \times \text{FTE})$</td>
<td>90405451.5</td>
<td>9508.2</td>
</tr>
<tr>
<td>Dunedin</td>
<td>$31043.8-(1116.2 \times \text{Tem}) - (758.86 \times \text{FTE})$</td>
<td>3627612.3</td>
<td>1904.6</td>
</tr>
<tr>
<td>Hamilton</td>
<td>$55496.3-(953.168 \times \text{Tem}) - (496.2 \times \text{FTE})$</td>
<td>1993100.6</td>
<td>1411.8</td>
</tr>
</tbody>
</table>

Table 1. Model equations and MSE and RMSE (kW.h)
Figure 3. Correlation between the actual and predicted energy usage (1000 kW·h) in studies research offices

The differences between the models showed it would be too difficult to develop a model estimate even in similar buildings. However, the development of the models was simple and accurately compared actual and predicted energy usage and estimate energy auditing in buildings. The final linear models were very simple (two variable regression model) compared with complex nonlinear modelling methods and can even be used in MS Excel.

This modelling method can be used by building owners, property managers and consultants to monitor and manage energy usage in existing buildings. The hypothesis of this study can be used in other similar projects. Based on the introduced models, several models have been developed in different commercial buildings. For example, similar modelling methods have been used to forecast energy usage in a number of swimming pools and libraries using outside temperature and numbers of visitors.

4. Conclusion

This study presents a modelling method to predict energy usage to estimate energy savings in energy auditing projects in existing office buildings, when investigated separately. The results showed that the linear model with simple independent variables can predict energy usage within acceptable errors. The main challenge of this method was finding accurate data in an acceptable period of time.

References


Labor Management in Masonry Construction: A Sustainable Approach

L. Florez\textsuperscript{a} and D. Castro-Lacouture\textsuperscript{b}

\textsuperscript{a,b}\textsuperscript{School of Building Construction, Georgia Institute of Technology, 280 Ferst Drive, Atlanta, GA 30332, U.S.A.}
E-mail: lflorez3@gatech.edu, dcastro@gatech.edu

Abstract -

Work in masonry is divided in different tasks and performed by sub-contractors. Sub-contractors specialize in a given skill to accomplish a specific construction task and usually work with crews. Due to this specialization, multiple crews with different skills are present on the jobsite at any one time and the sub-contractor is responsible of configuring crews to produce good quality work by means of capitalizing skills. Even though crews work independently within the range of their own skills, they need to properly sequence and coordinate their work to complete tasks and eliminate disruptions. One of the problems masonry contractors face is the need to design crews, that is, determine the number of crews and the composition of each crew to be effectively used in the construction process to maximize workflow. This study proposes the framework for a decision support system to assist contractors in the allocation of crews in masonry projects. The proposed system can be a valuable tool to assist masonry contractors in the process of predicting the number of workers and allocating workers.

Keywords -
Labor Management; Masonry Construction; Social Sustainability, Simulation Optimization

1 Introduction

Project-oriented organizations such as construction companies use temporary organizations to perform work [1, 2]. Projects are limited in time and scope and every time a new project starts the configuration and allocation of workers changes [3]. This creates a dynamic and transient work environment frequently seen in masonry construction projects. In masonry, workers are transferred between different projects and even between different roles to accomplish a specific task. This temporary nature of the work influences the motivation of workers and creates pressure due to the uncertainty of future assignments, variability of work durations, and lack of opportunities for career growth [3, 4, 5]. Furthermore, the size and number of projects change constantly, making it difficult for contractors to predict future workers and for workers to achieve a work-life balance.

This paper proposes the framework of a decision support system (DSS) for sustainable labor management that is practical to implement in masonry construction projects. The DSS addresses both the challenges faced by contractors when predicting future labor and allocating workers and workers’ needs of career development and labor stability. It is expected that the DSS will help contractors efficiently manage masonry workers in a more ethical and sustainable way.

2 Masonry Construction

Masonry activities such as scaffold installation, mixing mortar, laying block, cutting block, and grouting are labor dependent and require a large number of workers with diverse skills. Work is physically demanding (Spielholz et al, 2006) and masons as well as mason helpers (laborers) often lift heavy materials and stand for long periods of time (Boschman et al, 2011). A number of studies have identified masonry features that may impact masonry crew’s performance. These include excessive block cutting, numerous corners (Sanders and Thomas, 1991), numerous openings (Hassanein and Melin, 1997), the use of non-adjustable scaffolds, and the size of masonry units (Mortlock and Whitehead, 1970). Additionally, masons work outdoors and are subject to poor weather conditions that reduce their work activity (BLS, 2014).

Masonry contractors divide work in different wall sections and usually work with crews (Ng and Tang, 2010) to accomplish a specific task. Tasks are dictated by what area of the building is accessible and available for crews to work in, and what pipes are already in place or what openings are marked and ready so that masons
can work around. When a wall section is completed, often times, crews are re-assembled and the contractor is responsible of configuring and assigning crews with different skills (Ng and Tang, 2010) to produce good quality works. Due to this task division, multiple crews are present on the jobsite at any one time. Even though crews work independently within the range of their own skills, contractors need to properly staff crews to complete tasks and maximize workflow. Therefore, to optimally assign workers, contractors need to consider not only tasks and activities but also workers’ skills.

By considering workers’ needs and expectations, construction projects may not only benefit current workers but may also help attract new workers to fill these positions. Due to this task division, multiple crews are present on the jobsite at any one time and the contractor has to constantly assemble crews and properly sequence their work to maximize workflow.

2.1 Roles

Field observations indicated that masonry crews have a number of roles that need to be filled to build a wall. The masonry roles presented below were identified during extensive observations conducted in a masonry project and are complemented with the roles identified in Memarian and Mitropoulos [13]. The observations were conducted in a multi-storey building of mixed-use space with a parking structure and come from non-union crews. Labor rules on union jobs may change the job-sharing observed in this research. Typically in a block masonry crew there are seven major roles:

- The superintendent: responsible for staffing. Hires workers and allocates field personnel to the projects.
- Project superintendent (often-called foreman): monitors crews, layouts the walls, plans and distributes tasks, monitors workers, orders materials and equipment, performs quality control, and coordinates working areas with other foremen (electrical, plumbing, mechanical).
- Masons: lay block, install rebar and wire for reinforcement, and set up door frames. If needed they also perform support work, e.g., grouting and wall layout.
- Forklift operator: delivers materials from the storage areas to the work areas which include blocks, scaffold, mortar container, and grout container. In projects with small access areas, it often uses a warehouse lift in combination with the forklift. In big projects, the number of operators increases due to the number of workers and/or the number of floors.
- Saw operator: cuts block.
- Laborer on the mixer: makes grout and mortar.
- Laborers: perform support tasks including mixing mortar in the buckets and delivering mortar to the mortar boards, transporting and stacking the block, erecting and dismantling scaffold, grouting, and handling wire and rebar for reinforcement.

Some workers may have skills that will allow them to work in different roles that require those skills. For instance, some masons can be assigned to layout a wall (measuring and marking the walls) or to a wall to lay blocks. Additionally, laborers may have the skills to be assigned to look over a mason, operate the saw, operate the forklift, or work in the mixer. This role swapping is possible as long as the laborer possesses the skills and knowledge to do so. For instance, the forklift operator needs a license to operate machines, the saw operator needs to know how to work on the saw and the one in the mixer needs to know the ratio of material to make mortar and/or grout. It is important to consider there are some basic operators that must be present in the jobsite at all times i.e. one laborer in the saw, one laborer mixing the mortar, and one laborer operating the forklift. This last one increases with the number of floors and/or the number of crews. The number of masons is determined by the workload, that is, what walls are ready for the masons to work in (in terms of space) and what walls have the ductwork, plumbing and conduits in place (in terms of availability). The number of laborers for support (to look over a mason) is often determined by the mason to laborer ratio the contractor wants to use.

2.2 Crew Makeup

Masonry contractors follow a process to design and configure crews. The process is applied by considering a set of rules that impact the whole crew design and ultimately helps contractors determine the number of workers per crew and the composition of each crew [10]. The rules presented in this study were determined from the observations conducted in the masonry project and are coupled with the rules identified in Hassanein and Melin [10].

- Rule 1- Crew composition: In general contractors don’t have a typical crew, but they have some guidelines as to the worker’s ratio in a crew that they use and try to maintain. This ratio dictates the typical number of laborers per number of masons that is used when forming crews [10].
- Rule 2- Crew operators: There are some basic operators that need to be present in the jobsite. There always has to be a laborer in the saw, a
laborer making grout and/or mortar, and a laborer operating the fork lift. This last one increases with the number of crews present in the jobsite and/or the number of floors.

- **Rule 3- Crew size:** In regards to the crew size, there are a maximum number of workers that are assigned to a non-working foreman. Typically foremen determine the maximum number of workers they feel comfortable handling to guarantee an adequate control [10].

- **Rule 4- Crew control:** The foreman has a working assistant foreman (often times called labor foreman) that monitors crews, builds scaffold, and helps layout the job. The number of assistant foremen is dictated by the foreman and depends on the number of crews.

- **Rule 5- Learning curve:** When forming crews, the foreman (or assistant foreman) tends to keep the same masons in a crew due to the learning curve. Masons that have worked together know what they are doing and have identified ways to collaborate.

- **Rule 6 -Compatibility of labor:** The foreman (or assistant foreman) considers a compatibility factor, that is, how masons get along to keep production up. Masons have different ways to work and get things done. Some masons work well together, but some do not get along or the way they work is not compatible. The foreman tries to form crews with workers that are compatible because masons that work well together are more efficient and this can improve quality and increase throughput [14].

- **Rule 7- Quality of work:** When the quality of work is a factor, experienced masons (i.e., journeymen) are assigned to wall sections that require a high demand of craft work (e.g., openings, corners, details), whether non-experienced masons (i.e., apprentices) are assigned to non-craft work (e.g., straight walls). The foreman, or assistant foreman, knows which worker is good at what, based on the skills.

### 2.3 Management of labor

Masonry contractors usually manage a program of different project types that are unique, custom-made, and temporary. These features add uncertainty and novelty as no project is completed using the same approach. Consequently, contractors need to work in a cross-functional way and integrate resources (labor, equipment, and capital) in such a way so that they realize the maximum benefit for the operations. By integrating resources, contractors align resources to achieve the desired quality within the given time frame [15].

Management of labor is one of the key factors in masonry to balance production and quality. Forecasting human resource requirements and human allocation are the two approaches used for the management of labor. Determining the number of workers to complete the projects while balancing cost and production is the primary objective of predicting labor. On the other hand, human allocation deals with designing and configuring crews to assure workers have the skills required by the projects to complete tasks.

#### 2.3.1 Forecasting labor

In construction, the resource demands for successive projects might be significantly different and that makes it difficult for contractors to plan for the required number of workers. Often times construction companies employ temporary workers for the periods of increased workload, but it takes time to find skilled and well-qualified workers. Consequently, predicting the labor power demand is essential to facilitate decision making and labor planning in construction [16]. Labor planning allows companies to have knowledge (in advance) of the possible consequences, which allows them to better deal with tight schedules and project demand. By predicting labor, a construction firm can assess long term staffing needs, ultimately increasing productivity and the quality of work.

There are several approaches that have been developed to predict labor in construction. Wong et al [16] used dynamic econometric modelling techniques to establish the relationship between the aggregate demand for manpower and a group of inter-related economic variables. They found that construction productivity and construction output are the most significant factors determining the demand of labor power. Wong et al [17] used multiple regression analysis to help establish labor demands in construction. Their regression models used variables such as project cost, project complexity attributes, project type and physical site conditions to predict labor. Cheng et al [18] developed a team human resource planning (THRP) method for deploying labor power in process reengineering using simulation. The method determined the maximum loading of projects the original labor power can carry and also identified the range of labor power required for the expected project loadings by combining process reengineering and a simulation approach.
2.3.2 Human allocation

Human resource allocation in construction is the process of assigning crews of workers to tasks [20]. Tasks may require several crews with diverse skills to be completed and crews need to be scheduled to ensure an efficient output and adequate control [10]. This allocation process in masonry construction is challenging. Often times, crews complete their work on a section, but then have to come back later to complete the other part of the work. Due to this characteristic, every time a wall section or part of a wall section is completed, the labor configuration is reorganized. This results in temporary crews that need to be constantly moving. One of the problems masonry contractors face is the need to allocate crews to maximize workflow while satisfying contractors’ constraints and worker’s needs.

There are several approaches that have been developed to allocate human resources in construction. Al-Bazi and Dawood [19] presented a strategy to allocate crews of workers in the precast concrete industry using genetic algorithms-based simulation modeling. Lin [20] proposed a decision-making model for human resource allocation in remote construction projects. El-Rayes and Moselhi [21] developed an optimization model that uses dynamic programming for repetitive construction projects. Maxwell et al [22] presented a stochastic simulation program to measure the elapsed time and activity cost of each candidate crew. Their proposal uses an optimization rule to determine the best crew configuration. A review of the aspects and modeling approaches in personnel allocation and scheduling can be found in Brucker et al. [1].

3 Workers’ needs

Workers place a great value on requirements such as involvement, respect, and sense of personal growth (Lingard and Sublet, 2002). From the contractor’s point of view, a project that offers continual employment allows the contractor to maintain an acceptable skill level and helps generate a sense of commitment to the job from the workers (MacKenzie et al, 2010). In addition, it decreases the possibility of schedule overruns and leads to less deviations from the normal workflow (Lee et al, 2004). From the worker’s point of view, a job that gives stable work increases employment duration and provides opportunities for career growth (Loosmore et al, 2003). Furthermore, a change often implies not only a change in pay grade, but also the need to get familiar with new workers and practices. That is, workers prefer jobs that will keep them busy all the time. This situation increases labor skill level. Workers prefer to work with skilled peers because they are good at what they do. This decreases the possibility of rework and helps keep a normal workflow.

The proposed way to increase employment duration is based on the labor stability indicator developed in Florez et al. (2013). By considering labor stability, a construction project smoothes the allocation of workers to avoid drastic measures and minimize the variation of the number of workers. Therefore, a stable workforce provides benefits to the workers in terms of employment duration (Florez et al, 2013) and benefits to the contractor in terms of workflow (Lee et al, 2004).

This dynamic work environment is regularly seen in masonry construction: project demand is unpredictable, which forces managers to constantly adjust the number of workers and reconfigure crews to get the projects done. Workers are often transferred to different projects and even between different roles in a project according to the type of project and the necessary skills. This temporary organization and dynamic environment can influence the motivation and impose pressure and stress to workers. Workers cannot be sure what kinds of projects they will be assigned to, the location of the projects, or the co-workers they will work with [3]. Furthermore, given the peaks in workloads and the uncertainty of work, it is difficult for workers to achieve a work-life balance.

Given this temporary environment, there is a need for managers to assure an ethical treatment and the well-being of workers. However, project management has traditionally focused on planning operations with little attention being paid to the workers. The organizational perspective of project oriented organizations in sake of profit maximization and client demands usually prevails over the needs of workers [3].

People, unlike other resources, have their own needs and requirements beyond the financial compensation for their work. Workers place a great value on requirements such as involvement, respect, and sense of personal growth. Because of their needs and requirements, workers may represent the most difficult resource for organizations to manage, but when managed effectively can bring considerable benefits [4].

For instance, a project that links project assignments to career development allows contractors to not only develop staff for its future projects, but also helps workers develop skills and opportunities for career growth [5]. At the same time, contractors benefit through increased work control by better using education and competence development that increases job satisfaction and motivation [3]. Projects have better productivity rates where employee work experience is enhanced by assembling teams that consider managers and co-workers as well as tasks and roles [3, 4, 5]. By
considering their needs and expectations, construction projects may not only benefit current workers but may also help attract other workers.

This study proposes a DSS to forecast labor needs and allocate crews of workers in masonry projects that considers simultaneously both contractor’s requirements and workers’ needs.

4 Proposed Tool

The proposed strategy for sustainable labor management in masonry considers two subsets. The first subset forecasts the number of workers such that the contractor can determine its labor needs. The second subset allocates those workers to projects such that the workflow is maximized while considering contractor’s requirements and workers’ needs.

The methodology for developing the decision support system (DSS) comprises three phases: data-analysis, simulation, and optimization. Figure 1 shows a high-level decision process diagram of the DSS for sustainable labor management in masonry. The data-analysis phase characterizes the projects in terms of duration and size. To represent the behavior of projects, the methodology adjusts probability distribution to the demand of projects in a contractor firm. The simulation phase aims to predict the number of workers that are needed to complete the projects within the given time constraints. The input to this phase includes the probability distributions of the activity durations (with the maximal labor power for an activity), which allow generating multiple scenarios using random numbers drawn from the probability distributions. The optimization phase aims to allocate workers to crews and crews to projects in order to maximize workflow. The input to this phase includes the labor prediction of the second phase, which allows generating multiple scenarios. The output is a detailed schedule of the times to start tasks, the number of workers and the skills of the workers needed as well as the crew configuration under the optimal schedule.

4.1 Data analysis phase

Because activity duration with maximal labor power is the input for the simulation phase and subsequently the optimization phase, the data analysis is the key phase of the methodology with two objectives. The first objective was to identify the types of activities required in the construction of masonry walls and the second is to characterize the duration of the activities. Generally, the duration of an activity is simplified as a linear function of the number of workers in an activity, that is, the duration will decrease as more labor power is assigned. In the masonry industry, the duration of an activity usually decreases with the number of workers in a crew, but there is a point of maximum labor power. After that point is reached, the productivity rate does not increase, that is, more workers cannot further reduce the duration [18].

The duration of any activity needed in the construction of masonry walls is defined as the shortest operation time with a given labor power for an activity. To examine the actual duration of all activities, this study characterized the activities and captured labor productivity as a function of crew size using self-organizing maps [23] and time series analysis [24].

On the other hand, not only the identification of activities and their duration has a major effect on the methodology, but also the accuracy level is related to the duration’s fit level. Therefore, the methodology captures the duration by fitting probability distributions. Fitting methods allow selecting the parameters that produce the best fit to the collected data. With this fitted distributions, it is possible to replicate the duration of activities in the next phases of the methodology.

4.1.1 Simulation model

The proposed discrete-event simulation model (shown in Figure 2) represents the construction of a masonry wall, where the workers are assigned to different activities to complete a wall. Figure 2 shows the workflow of the CMU process. The workflow starts with the layout process, which is the process of measuring and marking the walls. The project superintendent usually marks the walls but sometimes an experienced mason (that knows how to read the blueprints) does this process. The second process is to prepare the wall so when the masons get to the wall they can start laying block. The forklift operator delivers materials to the area while a laborer is in charge of placing the resources and accommodating the blocks and mortar boards.

The third process is the actual process of laying block and it is the mason’s primary activity. The forklift operator is directly involved in this process since he has to deliver the mortar to the area in buckets. The laborer that is looking over the mason coordinates the delivery with the operator and is responsible of placing mortar in the boards on a timely basis to make sure the mason always has mortar to spread. The fourth process is grouting and this process involves completely filling the cells with grout to create a solid concrete wall. The fifth process is cleaning and finishing the wall. This process involves joining both the head joints and bed joints and brushing the wall to clean the mortar outside the joints.
Finally, the mason determines if the wall is at ‘scaffold height’ or if the maximum height has been reached. If it is scaffold high, the mason stops laying block and then the laborer proceeds to build the scaffold and prepare the wall so that the laborer can continue laying block in the next portion of the wall. Scaffolding involves erecting and dismantling the scaffold. The combination of laying block and scaffolding continues for as long as the maximum height of the wall is reached.

4.2 Optimization phase

This phase consists of running a series of optimization models capturing different labor requirements with the goal of maximizing workflow in a masonry program. To do so, a deterministic integer programming (IP) model which is fed with the predicted parameters from the simulation phase will be used. These parameters are random number generated using the output of the data analysis phase, the duration of activities and the number of workers probability distributions. Therefore, the optimization model input is the complete schedule and duration of all the activities. This detailed schedule represents a (random) labor requirement realization, drawn from the probability distributions derived in the data analysis phase.

The IP model obtains the optimal allocation of workers that maximizes the workflow for a particular realization of labor demand, taking into account contractor’s requirements and workers’ needs. In other words, the allocation of workers considers the rules that for crew design (Section 2.2) and also specific needs that workers have for developing skills and labor stability. Then, with a large number of randomly generated scenarios, the optimal allocation of workers sheds light into the right number of workers to have in payroll and the configuration of crews.

The data input module collects the information on the rules that contractors use to makeup crews. The reader is referred to the sub-section 2.2 for a detailed explanation of the rules. The data input module also includes for every time period, information on the availability of workers. Table 1 shows the skills for each worker. The binary parameter takes the value of 1 if a worker has that skill; it takes the value of 0, otherwise.

<table>
<thead>
<tr>
<th>Worker</th>
<th>Mason</th>
<th>Saw</th>
<th>Forklift</th>
<th>Mixer</th>
<th>Laborer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worker1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Worker2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Worker3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Worker4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Note that some workers have different skills that will allow them to work in activities that require those skills. For instance, worker 1 is a mason that can operate the saw, and can be allocated to either a wall to lay blocks or to the saw to make cuts.

Once all the data input is entered, the optimization model allocates labor to the projects. The model’s objective is to allocate crews of workers so that the workflow, that is, production in terms of number of masonry units is maximized. The allocation process consists of assigning the time needed to be working in a specific task to each of the workers in the crew. Each task demands a certain number of crews and each crew demands a certain number of workers with different skills. To build the schedule, the model uses binary decision variables to define the times each worker is working in a specific task. The model determines which worker should be working in which crew and doing which task.

5 Discussion

The crew allocation process in masonry construction is challenging. Masonry is labor-intensive and often crews have to sequence their work to avoid disruptions and maximize production. Multiple crews with different skills are present on the jobsite at any one time and it is difficult to determine when each crew should be working and the workers that need to be assigned to a crew.

The proposed model aims to help contractors allocate crews of workers in masonry projects. The model integrates a simulation and a modeling approach in an attempt to help alleviate some of the issues faced by masonry decision-makers in their day-to-day practices. Typical contractors’ requirements are included to address realistic scenarios experienced by masonry contractors in the jobsite. These requirements dictate the rules contractors use to design and makeup crews based on experience. In addition, the model considers workers’ needs that helps workers stay longer on the project and increase continuity of job assignments. With the model, contractors are able not only to determine the optimal starting times for each of the crews, but also to quantify the number of crews and ultimately the number of workers needed to build a masonry project.

These new considerations should prove useful to masonry contractors and enable them to optimize the allocation of multiple workers in order to maximize workflow. An optimal schedule that considers the needs of workers and contractors contributes to formulate strategies to make project management more sustainable and increase the benefits achieved by both workers and contractors.
Figure 1. Decision support system (DSS) for labor management

Figure 2. Workflow mapping diagram of the simulation process
References


Modelling October 2013 bushfire pollution episode in New South Wales, Australia

Hiep Nguyen Duc*, Sean Watt*, David Salter*, Toan Trieu*

*Office of Environment & Heritage, NSW, PO Box 29, Lidcombe, NSW 2141
E-mail: hiep.duc@environment.nsw.gov.au

Abstract -
From mid-October 2013 to the beginning of November 2013, New South Wales (NSW) experienced a significant sustained bushfire spring period in the Blue Mountains, the Central Coast and the Southern Highland causing property damages in those areas, and high pollution (particle and ozone) episodes over the metropolitan areas of Sydney.

This paper studied the extent of the air pollution due to the October 2013 bushfire using data from various sources as well as meteorological modelling to understand the course of bushfire and its effect on air quality in Sydney. Satellite images of smoke and burned areas obtained from MODIS Terra/Aqua satellites together with predicted wind from the WRF (Weather Research Forecast) model at different height levels and the measured meteorological and air quality data collected from the monitoring stations in the Sydney region will be used to give a complete picture of the extent of the effect of bushfire on air quality in the Sydney and surrounding areas.

Keywords -
NSW bushfire, WRF meteorological model, particle air pollution, MODIS Aqua/Terra satellites.

Introduction
Australia, except in the north east, is mostly a dry continent. The weather, especially rainfall, is driven by the four main climate drivers: the ENSO (El-Nino Southern Oscillation), the SAM (Southern Annular Mode), the PDO (Pacific Decadal Oscillation) and the IOD (Indian Ocean Dipole). On the East Coast of Australia, the ENSO strongly affect the rainfall pattern. During the El-Nino period, drought occurs throughout most of regions in New South Wales and Victoria. And bushfires have been a part of life in Australia for millions of years, particularly in the south east of the country which is regarded as perhaps the most bushfire-prone region on the planet. Serious fires occurred frequently and depending on location with return period ranging from 10 years for the Blue Mountains, NSW to 13 years in Victoria (Cunningham, 1984).

On October 17th 2013, NSW experienced one of its worst bushfire day in 100 years, with major fires threatening property across NSW, including the Central Coast, Southern Highlands and the Blue Mountains. On October 23rd, the fires in the Blue Mountains had destroyed 208 properties, of which 193 homes were destroyed in the lower Blue Mountains town of Wimnalee.

Three distinct phases of activity took place across this bushfire episode; from October 17-18, from October 23-27 and on November 2nd. Whilst spring bushfire episodes are uncommon, they are not unexpected in eastern NSW, as spring is part of the peak bushfire season (BOM, 2009). However, a significant proportion of media attention surrounding these fires revolved around the potential role of climate change in their occurrence. Whilst it is still difficult to attribute anthropogenic climate change to any particular climatic or meteorological event (Peterson et al, 2013), Australia’s Climate Commission notes that parameters responsible for extreme bushfire weather in the country’s south east have increased over the last third of a century. These conditions are anticipated to increase the frequency and intensity of bushfires in Australia as the climate changes (Hughes and Steffen, 2013).

As a direct consequence of bushfire activity, air quality in the Sydney Metropolitan area deteriorated as concentrations of PM<sub>2.5</sub>, PM<sub>10</sub> (Particulate Matter less than 2.5 and 10μm) and ozone (O<sub>3</sub>) increased. These observed augmentations in pollution concentrations are synchronous with the three phases of activity noted above.

The elevated concentration of air pollutants during high pollution episodes, such as bushfires, is strongly associated with high mortality and hospitalisation rates due to respiratory and cardiac diseases in the exposed population.
It is important to study these high pollution days in detail as this provides the basis to further develop the health exposure studies and to verify the development and prediction of the air quality modelling system.

The use of NOAA AVHRR and MODIS Terra/Aqua satellite images, observed data from monitoring sites in the Sydney Metropolitan area and wind modelling from the WRF meteorological model resulted in the ability to determine the meteorology responsible for the fire conditions and the fire’s effect on air quality. The satellite images also allow the determination of the burned areas from which the emission rates of various air pollutants can be estimated based on the fuel load and emission factors.

The meteorological and emission modelling data will then be used to drive the air quality dispersion modelling system which is currently under development at the Office of Environment and Heritage, NSW (OEH, NSW). This air quality modelling system, based on a chemical transport model (CTM), will be able to predict the ozone and particles (PM$_{10}$ and PM$_{2.5}$) in temporal and spatial domains under various meteorological and emission scenarios.

The bushfire of 2013 is a good case study to test and verify the air quality modelling system for the Sydney airshed.

**Data sources for the bushfire studies**

The NSW Office of Environment & Heritage (OEH NSW) operates and maintains a comprehensive ambient air quality monitoring network in NSW. Most stations measure the following air pollutants: particles (PM$_{10}$ and PM$_{2.5}$), visibility (nephelometer), oxides of nitrogen (NO$_2$ and NO), ozone, sulphur dioxide and carbon monoxide.

The Sydney basin currently has 14 monitoring stations scattered throughout the Sydney metropolitan region from the coastal area in the east to the edge of the Blue Mountain in the northwest and in the west (Duc et al., 2012). During late spring and summer period, usually in the morning after sunrise, onshore sea breeze flow from the east and north east flowing across Sydney towards the southwest causes an elevated level of ozone in the southwest and west of Sydney in the afternoon (Hart et al., 2006). In the evening and night-time, drainage flow of cold air from the mountains in the west is directed to the coastal east and from southwest to the north.

Figure 1 shows the concentration of particles (PM$_{10}$, PM$_{2.5}$, and visibility) and gaseous air pollutants (ozone, nitrogen dioxide, sulphur dioxide and carbon monoxide) during the bushfire period of 17th October to 11th November 2013. All particles standards (daily PM$_{10}$ at 50 µg/m$^3$, daily PM$_{2.5}$ at 25 µg/m$^3$, nephelometer back scattering index of 2.1 Bsp) were exceeded during this the bushfire period.

Among the gaseous pollutants standards (hourly ozone at 10 ppb or parts per hundred millions, nitrogen dioxide at 12 ppb, 8 hourly carbon monoxide at 9 ppb or parts per million and sulphur dioxide at 20 ppb for hourly and 8 ppb for daily) only the ozone standard was exceeded.

![Figure 1](image-url)
The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

Figure 2 - Bushfire on 26/10/2013 as captured by NOAA AVHRR satellite and MODIS thermal hot spot detection.

Data analysis of air pollutants from ground-based monitoring stations and WRF model

The end goal of the project is to model the air pollution from the October 2013 bushfire period in NSW, Australia. However, in order to do that, the amount of biomass burned needs to be calculated. In the first instance, this involves the collection of burn-scar and hotspot data. This was done using shape file images and Terra and Aqua MODIS GeoTiff images for the burn-scar and hotspot data respectively – both from NASA.

The hotspot data was collected for two purposes. The first was to gain an understanding of the total number of hotspots in the month of October, whilst the second purpose was to determine the number of hotspots that occurred on each day of the bushfire episode – from 17th October to 3rd November.

In order to see if these two data sets matched up with the modelled wind meteorology from the WRF (Weather Research Forecast) model, both maps were layered together in ArcGIS. To gain a better understanding of the spatial variation of the hotspot and burn-scar data, these were layered with roads, towns and suburbs of the surrounding locales.

The next step of this process is to analyse the 1 hour average of a number of the pollutants from each of the OEH’s 15 monitoring sites around Sydney. The parameters collected for analysis were PM$_{2.5}$, PM$_{10}$, ozone ($O_3$).

The data analysis is to be undertaken using the OpenAir open-source tools for the analysis of air pollution data. This is a package library in the R statistical program. Having identified the sites of highest concentration for each pollutant, polar plots and concentration plots where constructed using OpenAir. Following this, the source of the pollutants was tracked using the back trajectory function. The GoogleMapsPlot function was used to spatially identify areas of greatest concentration within the 15 monitoring sites in the Sydney Metro area.

Summary analysis of air pollutant data

Time series and summary statistics of air pollutants are analysed using Excel spreadsheet and R statistical package for all the days during the bushfire period. Figure 3 shows the plot of ozone concentrations at all stations in Sydney.

Figure 3 – Ozone concentration as measured at Sydney monitoring stations during the bush fire period. Ozone exceeds the standard (10 ppm) on the 21 Oct 2013 and very high (above 8 ppm) on the other days.

Figure 4 – Particle concentrations (hourly averages) as measured at Sydney monitoring stations during the bush fire period.

To accommodate the spatial variability of the concentration, the plots of concentration statistics at the stations are overlayed on the Sydney map using OpenAir. An example of the spatial plot is given in Figure 4.
Figure 4 – Panel (a) shows the 24 hour maximum PM10 across the 16 monitoring sites in the Sydney region for the 23rd November using the GoogleMapsPlot function in Openair. Panels (b) and (c) show the same for ozone (O3) and nitrogen dioxide (NO2) respectively.

Weather Research Forecasting (WRF) Model

There are a number of regional meteorological models that are available for modelling the meteorology over the Sydney and surrounding areas. The CCAM (Cubic Conformal Atmospheric Model) and the TAPM (The Air Pollution Model) models were developed by CSIRO (Commonwealth Scientific and Industrial Research Organisation) in Australia and are commonly used in a number of countries. The Weather Research Forecast (WRF) model is developed by NCAR (National Center for Atmospheric Research) in the US. As a community model, it is very popular and is widely used by many organisations in the world. The WRF model is used by many organisations in the world including national weather bureau in some countries for real-time forecasting. In this paper, we will use the WRF for modelling the October 2013 bushfire period.

WRF is a mesoscale weather model prediction system. It is a 3-D model with dynamic advection, physics including various convection, radiation, boundary layer and cloud microphysics options. As the WRF model is a mesoscale regional model, its use is to downscale the coarse global forecasting model (GFS) such as NCEP GFS or Global Climate Model (GCM) prediction to a finer resolution applicable to a region (eg Eastern Australia).

We used the WRF with the defined options for the forecasting of meteorological variables at different height during the bush fires period of 17 Oct to the 3rd November 2013. The defined options are single-Moment (WSM) 3-class simple ice scheme for microphysics, Rapid Radiative Transfer Model (RRTM) scheme for longwave radiation, Dudhia scheme for shortwave radiation, Monin-Obukhov similarity scheme for surface layer physics, Noah Land-Surface Model for land surface, YSU scheme for boundary layer, Kain-Fritsch (new Eta) scheme for cumulus cloud physics and 4 soil layers in land surface model.

The domain is 74 by 61 grids with grid size at 37.5 km in x and y direction and 30 vertical levels with pressure at top is 5000hP.

The initial and boundary conditions and coarse data are the NCEP GFS global analysis data downloaded from NCEP GFS to drive the WRF model run.

An example of the WRF prediction which is
compared with satellite images is given in Figure 5 and 6. Figure 5 shows the satellite image on 23/10/2013 at 3 UTC (1pm local time) and Figure 6 shows the WRF predicted wind at cloud level overlayed with satellite image using the IDV software.

Figure 5 - Bushfire on 23/10/2013 at 3 UTC time as captured by NOAA AVHRR satellite and MODIS thermal hot spot detection

Figure 6 - The cloud level (~3000m) wind on 23 Oct 2013 at 3UTC as predicted from the WRF model and is overlayed on the satellite image and hot spots from MODIS on the same day and time.

It can be seen from Figure 5 and 6 that the WRF prediction of wind at cloud level is fairly accurate. Furthermore the spatial pattern of the maximum concentration of air pollutants at Sydney stations as shown above follows closely the path of plumes in the North West direction.

Radiosonde upper air data as measured at Mascot airport on the coast show a temperature inversion at about 300 hPa (~9000m) in the morning at 6am. The smoke cloud was expected at most at this inversion height.

Conclusion

The extent of the air pollution due to the October 2013 bushfire has been shown using data from the monitoring stations and satellite images. Meteorological modelling using WRF give predicted wind field at different high levels.

The WRF wind predictions are compared with satellite images of smokes and burned areas obtained from MODIS Terra/Aqua satellites. The results show that the WRF wind prediction is reasonably accurate. This allows us to use the WRF predicted wind field at different hour to show the extent and evolution of the bushfire plumes during the October 2013 bushfire period in the surrounding areas of Sydney.

The next step in the study is to estimate the daily emission rates of various pollutants (gases and particles) from the burned areas as derived from satellite images and the development and verification of a chemical transport model, using the bushfire period of October 2013 as a case study, for the Greater Metropolitan Areas which include Sydney, the Lower Hunter (north of Sydney) and the Illawarra (south of Sydney).

References


Towards Measuring the Impact of Personal Control on Energy Use through the Use of Immersive Virtual Environments

Arsalan Heydarian^a, Joao Carneiro^b, David Gerber^c and Burcin Becerik-Gerber^d

^a,b,c,d Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, USA
E-mail: heydaria@usc.edu, jcarneiro@usc.edu, dgerber@usc.edu, becerik@usc.edu

Abstract -
Recent studies have focused on increasing energy efficiency in commercial buildings through technological means (e.g., efficient HVAC systems, sensors and sensing systems). However, most studies underestimate the impact of occupants' behavioural choices. Lighting systems account for approximately a fifth of the total electricity consumption in the US; commercial buildings account for 71 percent of such consumption. This paper focuses on human behaviour related energy consumption by investigating the impact of personal control on lighting use in office environments. To effectively examine human energy consumption behaviour, alternative 3D design models of an office are created using an immersive virtual environment to visualize different lighting control features. Participants are brought into these immersive virtual environments by wearing Head-Mounted Displays and are asked to interact within these environments and perform a defined task. Participants were then allowed to control and change the room’s lighting settings based on their preferences in order to perform their assigned task. Participants were then allowed to control and change the room’s lighting settings based on their preferences in order to perform their assigned task. Unique to our experimental design is the use of immersive virtual environments, enabling measurement and control of a series of design feature isolations and combinations. The work presents the impact of decisions made both during design and operation of buildings on occupants’ energy related behaviour. The experiment demonstrated that when participants are provided with personal controls for the blinds and the artificial light, there is no significant difference in their preferences between natural and artificial lighting; however participants are significantly more likely to open the blinds remotely if they are only provided with a personal control for the blinds.

Keywords –
Immersive virtual Environments; Energy-use behaviour; Parametric Design; Human Building Interaction

1 Introduction
Buildings consume roughly 44 percent of the energy produced in the United States [1]. Global energy consumption and CO₂ emissions have significantly increased in the recent years by 49 percent and 43 percent, respectively, and these figures are estimated to annually increase by 2 percent and 1.8 percent, respectively [11,20].

In order to reduce the energy consumption of commercial buildings and CO₂ emissions in general, many researchers are investigating ways to make buildings more sustainable, intelligent, and responsive to occupants, environment, and the societal needs [4,14]. One potential way to improve the performance of buildings is to increase the interaction between buildings and their occupants. For instance, by allowing users to have more control over the available natural lighting in their office, the daily lighting energy-use could possibly be reduced.

Buildings are designed based on standard set-points and assumptions about occupants’ behaviours and comfort levels that are thought to provide satisfaction to occupants. However, research studies have shown that occupants are not always satisfied and such set-points do not guarantee occupant comfort [3,13-15]. One method to incorporate occupants’ behaviour in to the buildings’ systems is through simulating defined behaviour models [9,21]. Simulations are commonly used to estimate the occupant activities and needs. However, due to the complexity of human behaviour and the diversity among building occupants, simulations are usually not a realistic representation of all occupants in a building [14,25]. Therefore, an accurate measurement of occupant behaviour could be an influential factor in reducing the energy consumption in buildings [14,25].

Lighting systems account for 18 percent of the total electricity consumption in the United States, where 71 percent of this consumption is from the commercial buildings [12,20]. Previous research has indicated that a large amount of energy can be saved with incorporating
well-designed lighting controls to the building systems and more importantly by understanding the occupant’s lighting needs and preferences [7]. In order to gain information on occupant behaviour and resultantly better understand the occupants’ energy-use behaviour, the authors examine the impact of end-user’s control on the available lighting (adjusting lighting levels both artificial and natural lighting) in a single-occupancy office room. To effectively examine the end-user behaviour, the authors have used an immersive virtual environment (IVE) to create alternative design environments of the office room with various lighting settings and controls that the users can interact with in order to measure the impact of user control on energy use. The participants were put in a dark room and were asked to read a passage. They were given a set of manual and semi-automatic lighting control options to adjust the room’s available lighting in order to be able to fully read the passage. This paper measures the impact of having personal control on energy use through the use of immersive virtual environments.

2 Background

Previous researchers have studied occupants’ lighting preferences when they are provided with different options of natural lighting vs. artificial lighting [22,24]. For instance, [24] showed that employees strongly prefer natural lighting and an outdoor view in an office environment. Other researchers have studied the effect of windows sizes and occupants’ preferences about window types [19]. Many researchers have explored the effect of lighting control systems and the reaction of occupants towards such control systems [10,16]. Such studies have suggested that occupant satisfaction is increased with semi-automatic and manual modes of operations [23]. For instance, in the fully automated systems, the researchers studied the effect of sensor-controlled settings, in which the sensors determined how much natural lighting and artificial lighting should be available based on the availability of the natural light through the windows and the time of the day. In the semi-automated systems, the users were given limited control to manually adjust the sensor-controlled available lighting (i.e., dimming down/up the artificial lighting) [6]. Prior studies do not investigate the impact of having personal controls, though which occupants can increase the illuminance level from where they are located (impact of convenience) when performing a task, such as reading a passage. They also do not consider the energy consumption behaviour of occupants.

With the advent of virtual reality, augmented reality, and computer science fields, such as artificial intelligence and human-computer interaction, in recent years, Architecture, Engineering, and Construction (AEC) professionals have also access to such technologies more than ever. Such technologies provide AEC professionals with opportunities such as evaluating alternative designs [18], interacting and improving 3D models, communicating among parties [17,18], and more importantly studying human behaviour and preferences. Understanding human behaviour and preferences during the design phase allows architects and engineers to develop designs that would fit the end-users’ needs the best, resulting in a higher satisfaction and comfort. For instance, [5] brought healthcare organization end-users (e.g., doctors, nurses, etc.) to an IVE in order to present different designed environments and get their feedback to make the necessary changes and adjustments, resulting in a more improved design based on the end-users’ needs.

In this paper, in order to evaluate the effect of different lighting control options on energy consumption, an IVE was used to create alternative models with different settings, controls, and lighting settings, while providing realistic representations of physical environments. In another study, the authors investigated how human performance, perception and behaviour differ in an immersive virtual environment compared to an actual physical environment [8] and found no noticeable difference in terms of human performances between the two environments. The use of IVEs gives the researchers an opportunity to create environments with various control settings and evaluate end-users’ behaviour and preferences given different scenarios. This process might also significantly help the AEC professionals during the design phase of buildings to ensure their design not only meets the end-users’ preferences but also is more energy efficient.

3 Methodology

This paper examines human energy consumption behaviour and the impact of personal control on using different light sources (artificial vs. natural) in an office environment.

To evaluate the participants’ energy consumption behaviour, two parameters were measured: (1) participant preference (natural vs. artificial) and (2) the impact of availability of a personal control on preference. Parameters were measured based on the choices the participants made in choosing the source (manual light switch, personal light switch, manual window blinds, personal window blinds) to increase the brightness of the room along with responses to the set of questionnaires asked to the participants.

3.1 Experiment and Hypothesis

In this experiment, three possible lighting settings (no light, natural light, and artificial light) for an office were designed within an IVE to evaluate the participants’ behaviour when they were given the option to increase
the brightness (Figure 4). The participants were randomly assigned to 1 of 3 experimental groups that varied the options available to increase the brightness in the office (Figure 1). In group 1, participants only had the options of manually opening the blinds or manually turning on the light switch, requiring the participants to physically move within the IVE to either turn on the light switch or open the blinds (Group 1 in Figure 1). In group 2, the participants had the options of manually opening the blinds or manually turning on the light switch but they also were provided with a “personal remote control” that automatically opens the blinds; this control was set on the desk where the participants had to perform an assigned task (Group 2 in Figure 1). In group 3, the participants not only had the options to manually turn on the light switch and open the blinds, but they also were provided with two “personal remote controls” that could open the blinds or turn on the artificial light while they were sitting at their desk next to where they had to perform an assigned task (Group 3 in Figure 1). Figure 2 shows the different control options that each participant had to make the room brighter.

### Artificial Lighting System | Natural Light / Blinds
---
**Group 1** | Manually turning the lights on | Manually opening the blinds
**Group 2** | Manually turning the lights on | Manually or remotely opening the blinds
**Group 3** | Manually or remotely turning the lights on | Manually or remotely open the blinds

Figure 1 - Experimental Groups

Four hypotheses were developed to compare ‘within’ group and ‘cross’ group behaviours:

**H1**: There is no significant difference between participants’ choice of manually turning on the light switch and opening the blinds in group 1.

**H2**: Participants choose natural lighting significantly more than the artificial lighting when they are provided with a personal control only for opening the blinds (group 2).

**H3**: There is no significant difference between participants remotely turning on the light or remotely opening the blinds (group 3).

**H4**: There is a significant difference between remotely increasing the room’s brightness (both for artificial and natural) and manually doing so (group 3).

### 3.2 Model and Apparatus

In order to create a realistic model of an office space, an actual office room at the University of Southern California’s campus was selected and all dimensions of this office room were measured. At first, a base structure of the room was designed in Revit© 2013. The Revit model was then imported to Autodesk 3ds Max© to create a realistic rendering of the room by adding lighting, shadows, reflections, furniture, and materials. The 3ds Max© file was then imported to Architecture Interactive© (the IVE software), in which the participants were able to fully navigate and interact with the models and objects within the model. To make the models more realistic and interactive, animations for opening the blinds, clicking on the remote controls, and turning on the light switch were designed; this allowed the participants to have a more realistic interaction with the model.

![Figure 2 - Participants Lighting Control Options](image)

The system configuration used for this experiment composed of a Microsoft© Xbox Kinect, an Oculus HMD, a tracker, a Microsoft© Windows graphics...
workstation with NVIDIA® 3000M graphics card. To increase the sense of presence and to allow participants to realistically interact with the IVE, a Kinect was used to track the body’s displacement (3 Degrees-of-Freedom - DoF), a Head Mounted Display (HMD) was used to track the head rotation (3 DoF), and a tracker was used to navigate through the room, providing 4 DoF. Figure 3 shows the procedure for creating the models and the apparatus used for this experiment.

3.3 Procedure

In order to test the hypotheses, an experiment was conducted with 30 participants. The participants were undergraduate and graduate students at the University of Southern California between the ages of 18 to 36 years old. All participants completed a consent form. The participants had none or limited prior experience with IVEs. Since the participants were given the option to choose between the personal controls that were depicted in light and dark blue colours (Figure 2d), they were asked if they had normal or corrected visual acuity through a questionnaire for the purpose of this pilot experiment. Once the participants reviewed and agreed to the consent form, they were trained on how to navigate within an IVE, using a model different from the environment used in the actual experiment. During the training, they were instructed to find objects, navigate to different sides of the room using the tracker, and grab and move objects from one location to another. Once the participants felt comfortable with the IVE, they were instructed to remove the HMD and asked about their general feelings to ensure there was no motion sickness or headache caused by the IVE environment. Once the participants were ready, they were asked to put the HMD back on and were put in the experimental environment (the virtual office).

In order to eliminate any order effect, participants were given a random number when they entered the office room that corresponded to one of the experimental groups (10 participants per each group – see Figure 1). In each environment, the participants were instructed to navigate in the room and sit behind the desk and read a passage placed on the desk. At first the room was designed to be dark enough so the participants could not read the passage but were able to navigate and see the furniture in the room (Figure 4a). Then they were instructed on the possible choices they had in order to make the room brighter. For instance, if a participant was in group 2, he/she were instructed that he/she could (1) walk towards the door and manually turn the light switch on, (2) walk towards the window and manually open the blinds, or (3) click on the personal control right next to the passage to automatically open the blinds (Figure 2c).

![Figure 3](image3.png)

**Figure 3 – Modeling procedure and apparatus**

![Figure 4](image4.png)

**Figure 4 - Different Brightness Options Available in the Room. (a) Dark room with no available natural or artificial lighting, (b) bright room lit by natural light and (c) bright room lit by artificial light.**

In order to eliminate any order effect, participants were given a random number when they entered the office room that corresponded to one of the experimental groups (10 participants per each group – see Figure 1). In each environment, the participants were instructed to navigate in the room and sit behind the desk and read a passage placed on the desk. At first the room was designed to be dark enough so the participants could not read the passage but were able to navigate and see the furniture in the room (Figure 4a). Then they were instructed on the possible choices they had in order to make the room brighter. For instance, if a participant was in group 2, he/she were instructed that he/she could (1) walk towards the door and manually turn the light switch on, (2) walk towards the window and manually open the blinds, or (3) click on the personal control right next to the passage to automatically open the blinds (Figure 2c).
Once he/she chose one of the options available in his/her group, he/she either turned on the artificial lights or opened the blinds (see Figure 4). The participants were then asked to read the passage on the desk to complete the experiment.

4 Results

To examine the impact of personal control on using different light sources (artificial vs. natural) and test the a priori hypotheses, ‘within’ group and ‘cross’ group comparisons were performed.

In Group 1, three participants chose to use the light switch manually and seven participants chose to open the blinds manually. In Group 2, nine participants chose to remotely open the blinds, one participant chose to manually open the blinds, and no participants turned on the light switch manually. In Group 3, two participants chose to remotely turn on the lights, six participants chose to remotely open the blinds, one participant chose to manually open the blinds, and one participant chose to manually turn on the light switch. Table 1 summarizes the details of the experimental data.

4.1 Within Group Comparison

The ‘within’ group comparison examined the effects of different lighting control options within each group. The Null Hypothesis (H0) for each group was that participants would choose each option equally at chance. If participants do not have a prior preference of the lighting options and/or are not influenced by the presence of personal controls, they are expected to choose the options randomly, leading to an approximately equal chance of choosing each option.

4.1.1 Group 1

The Null Hypothesis (H0) specifically for this group was that 50 percent of the participants would choose to manually turn on the light switch while 50 percent would choose to manually open the blinds. To test the H0, a Chi-square ($\chi^2$) test confirmed that the percentage of participants that opened the blinds manually did not significantly differ from the percentage that manually turned on the light switch, $\chi^2 (1, N=10) = 0.21, p > 0.05$. See Table 2 for more detail.

4.1.2 Group 2

The (H0) for group 2 was that 33.3 percent of the participants would choose to manually turn on the light switch, 33.3 percent would choose to manually open the blinds, and 33.3 percent would choose to remotely open the blinds. A Chi-square ($\chi^2$) test confirmed that the percentage of participants who opened the blinds remotely significantly differed from the percentage that manually on turned the light switch, $\chi^2 (1, N=10) = 0.01, p > 0.05$. See Table 3 for more detail.

4.1.3 Group 3

The Null Hypothesis (H0) for this group was that 25 percent of the participants would choose to manually turn on the light switch, 25 percent would choose to manually open the blinds, 25 percent would choose to remotely open the blinds, and 25 percent would choose to remotely turn on the artificial light. Since the number of expected participants is less than 5 for each condition, Yates’ correction is applied to the Chi-square ($\chi^2$) test to make it a more conservative test. The first test compared the use of remote control for both blinds and artificial light. The Yates’ chi-square test confirmed that the percentage of
participants that opened the blinds remotely did not
significantly differ from the percentage that remotely
turned on the artificial lights, $\chi^2 (1, N=8) = 0.15, p > 0.05$. See Table 4 for more detail.

<table>
<thead>
<tr>
<th>Table 4 - Group 3 $\chi^2$ Analysis (Remote)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote Blind vs. Remote Light</td>
</tr>
<tr>
<td>Chi-Square</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
</tr>
<tr>
<td>p-value</td>
</tr>
<tr>
<td>Yates' chi-square</td>
</tr>
<tr>
<td>Yates' p-value</td>
</tr>
</tbody>
</table>

The second test compared the use of manual control (both for artificial lighting and blinds) vs. remote control (both for artificial lighting and blinds). This group comparison shows whether the participants were more likely to use the remote options more than the manual options. The chi-square test confirmed that participants used the remote options marginally significantly more than the manual options, $\chi^2 (1, N=10) = 0.057, p \approx 0.05$. See Table 5 for more detail.

<table>
<thead>
<tr>
<th>Table 5 - Group 3 $\chi^2$ Analysis (Grouped)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual vs. Remote (grouped)</td>
</tr>
<tr>
<td>Chi-Square</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
</tr>
<tr>
<td>p-value</td>
</tr>
<tr>
<td>Yates' chi-square</td>
</tr>
<tr>
<td>Yates' p-value</td>
</tr>
</tbody>
</table>

4.2 Cross Group Comparison

The within-group comparison in group 2 revealed that participants chose the personal control option for natural light significantly more than the manual control option for artificial light. However, in group 3, when a remote control is added for artificial lighting, the participants did not choose the natural light personal control option significantly more than the artificial light personal control option. Therefore, comparing these two groups, having only a personal control for the blinds is far more effective than having a personal control for both the blinds and artificial lighting. This is indicated by the significant chi-square p-value for the remote blinds in group 2 (Table 3) and the non-significant chi-square p-value in group 3 (Table 4).

4.3 Questionnaires

Through a set of questionnaires, the authors examined the knowledge of the participants about green building features and their familiarity with IVE. On a scale of one to seven [2] (one being not environmentally friendly at all and seven being very environmentally friendly), participants rated their concern for the environment as 4.86 on average. Meanwhile, 13 people were familiar with the term LEED (Leadership in Energy and Environmental Design), the only word indicative of the participants’ environmental knowledge in the series of words given to them. This indicates that the pool of participants could be slightly more knowledgeable about green buildings and energy efficient features than an average person.

The participants were also asked about how realistic they thought the IVE looked/felt and whether they thought the model was a good representation of an average office space, on a scale of one to seven (one indicating the model to be very unrealistic), they gave the modelling of the room a 5.22 on average. With regards to their familiarity with IVE, only three participants indicated that they have previous knowledge about the term; these participants indicated that they had an average amount of experience with virtual gaming.

5 Limitations and Future Work

Although served as a first step toward the research goals, this study had several limitations. There was a small sample size of 30 participants (10 for each group). In future studies, the number of participants will be increased. Also, most of the participants were engineering students, whom could be more environmentally friendly than the average population; in future studies a more diverse pool of participants will be used to reduce any bias in their decision making.

In future studies, a ‘fourth’ group will be added. This group will provide participants with the options of manually opening the blinds and manually turning on the light switch but also provide the artificial light personal control option without the natural light remote control option. We will then compare the effects of the remote control for artificial lighting on participants’ energy consumption behaviour to investigate whether participants will choose to open the blinds manually or prefer to semi-automatically turn on the artificial lights.

Some of the features of the models could also be improved (shadows, reflections, objects in the room) that could provide the participants with more realistic view of the office environment. As part of the future work, the authors will further investigate the effects of different design choices on human energy-use behaviour using IVEs. The authors will also use IVEs to explore the integration of multi-agent systems to impact building
design performance and occupant satisfaction.

6 Conclusion

Understanding occupant behaviour and their impact on the buildings’ energy consumption is an important avenue of research in order to reduce building’s electricity consumption. The authors aimed to explore the impact of personal control on lighting use in office environments using alternative 3D design models within immersive virtual environments. This paper demonstrated occupants are significantly more likely to use natural lighting if they are only given a personal control for blinds, but not more likely to use natural lighting if a personal control is available for both blinds and artificial light. Additionally follow-up questionnaires revealed that the participants chose the personal control feature for the blind, just because it was ‘easier’ and more convenient for them, which shows that such features could potentially not only be integrated to the existing operational buildings but more importantly could also be part of the design options for during the design phase of future commercial buildings. The use of IVE enables us to measure and control a series of design feature isolations and combinations to further understand what features are more effective compared to one another. This paper reveals an important application IVEs to increase the interaction between building design and construction and user behaviour and satisfaction.

7 Acknowledgment

This project is part of the National Science Foundation funding under the contract 1231001. Any discussion, procedure, results, and conclusion discussed in this paper are the authors’ views and do not reflect the views of National Science Foundation. Special thanks to all of the participants and to the researchers that contributed to this project; specifically to Alex Coco and Samantha Kaplan for their contributions in preparing and running the experiments.

References


Automatic Power Managing and Monitoring System
Applying for Underground Mines in Vinacomin

Tuan Anh Vu and The Nam Vu

Institute of Mining Science and Technology, Vietnam
E-mail: tuananhvu2984@gmail.com, thenamvu@gmail.com

Abstract-
In the situation of energy deficiency, the management of the economical and efficient power utilization plays an important role in the national energy policies and has been considered as an essential requirement for the coal production enterprises, especially in the underground ones. The report represents the features of the automatic power managing and monitoring system at Ha Lam Coal Company and the potentiality of its application to other underground mines in Vinacomin.

Keywords-
Management, monitoring, system, underground mine, automatic, energy efficiency, power, energy saving, Vinacomin.

Introduction
Electricity supply system has been playing an important role for stable production activities in mining industry. Due to the complicated conditions as well as the strict requirements to the flame and explosion-proof safety in hazardous environment, the power supply system of the underground mines is shown in Figure 1.1.

Centralized power monitoring and managing system had been designed with the target for the collecting and storing the measured signals to the Dispatch center. From this data, the operators could have the comprehensive assessments following real-time parameters such as: Current, voltage, power quality, power factor, capacity, efficiency as well as monthly bill. The electric management system comprises four parts (Figure 1.2):

1) Management Software with convenient user interface. It is the electricity monitoring and control center.
2) Smart Collector to connect with the PC via optical fiber. It collects data directly or indirectly from the smart terminals and send the center's instruction and settings to the smart terminals.
3) Smart Terminals: The smart terminals are sensor sockets integrated with repeater function. Each smart terminal can establish a star network with itself.
4) Measurement instruments: Measure real-time parameters such as current, voltage, power quality, power factor, capacity etc. The measurement instruments be able to connect each other following RS-485 Modbus.

In order to deal with the demand of power monitoring and its management which would help to improve management efficiency as well as reducing breakdown times related to the increase energy consumption performance. The author would like to describe the design and application results of automatic power monitoring and managing system implemented in underground mines belonging to Vinacomin.

System Description
Centralized power monitoring and managing system had been designed with the target for the collecting and storing the measured signals to the Dispatch center. From this data, the operators could have the comprehensive assessments following real-time parameters such as: Current, voltage, power quality, power factor, capacity, efficiency as well as monthly bill. The electric management system comprises four parts (Figure 1.2):
The monitoring network can be established, modified or expanded freely and easily by just adding or removing Smart Terminals. The system has functions as following:

- The system can be developed to manage the cluster monitoring devices in which could monitor the power consumption progress of the machines in the workshops; therefor, the faults can be infected on-time following the alarms in correspondence that the operators have the timely intervention.

- Storing and analysis the active parameters on the system software that can be drawn again on which the operator command. In addition, the software could give the forecasts relating the overload, the faults as well as the demand of consumption power.

- Analyzing electric quality based on the criterion such as: voltage, harmonics or frequency. Besides, the system would assess supply source quality.

- Monthly automatic power consumption reports and bills that represent as printout graphs, charts or tables.

**The Application Of System In Ha Lam Underground Coal Mine**

Centralized power monitoring and managing system applying in Ha Lam underground Coal Mine had been designed to respond the scale as well as production characters which has the ability to monitor energy consumption all machines not only in underground but also in mine surface as following:

- Measurement the characteristic activities of machines in underground: The three-phase power meter manufactured in flame-proof structure for measurement purposing the local transformers supplying electricity to power consumption loads in the workshops. The meters have communicated via the protocol of Modbus RS-485. Additionally, the data from these meters had been transmitted to master computer via the connected optical fibers between the smart terminals and the smart collector. (Figure 1.3)

- Measurement the performed characteristic of machines at the surface: The digital panel meters manufactured in flame-proof structure according to IP-54 standard for measurement purposing the local transformers supplying electricity to power consumption loads in workshops. The meters have communicated via the protocol of Modbus RS-485. And the data from these meters had been transmitted to master computer by GPRS package transmitter. (Figure 1.4)

- The master computer: The master computer in which setup the software that can collect the integrated data both in underground and at the surface.

**Table 1. Technical specifications of the measurement meters**

<table>
<thead>
<tr>
<th>Type of electrical system</th>
<th>3-phase with or without neutral wire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>$\pm(1.5%FS+1\text{DGT})$</td>
</tr>
<tr>
<td></td>
<td>$\pm(0.5%FS+1\text{DGT})$</td>
</tr>
<tr>
<td>Rated input current</td>
<td>5 A</td>
</tr>
<tr>
<td>Rated input voltage</td>
<td>660 V</td>
</tr>
<tr>
<td>Measurement method</td>
<td>TRMS type</td>
</tr>
<tr>
<td>Current transformer</td>
<td>Prog.ratio from 1 to 999</td>
</tr>
<tr>
<td>Voltage transformer</td>
<td>Prog.ratio from 1 to 99.9</td>
</tr>
<tr>
<td>Serial port RS-485</td>
<td>2 or 4 wirings</td>
</tr>
<tr>
<td>Comm.protocol</td>
<td>Modbus/Jbus</td>
</tr>
<tr>
<td>Baud rate</td>
<td>9600 baud</td>
</tr>
<tr>
<td>Frequency</td>
<td>$\pm0.1 \text{HZ} (48 \div 62 \text{HZ})$</td>
</tr>
<tr>
<td>Harmony</td>
<td>$\pm3%FS$.</td>
</tr>
</tbody>
</table>
The Results And Discussions
The collected data have been drawn as graph containing 3-phase voltage, current and power consumption that generate to real-time. At the period of abnormal time that represents such as: over-voltage, low-voltage or overload, the operators can zoom in or zoom out to determine the detailed irregular intervals.

Besides, the system also has evaluated continuously the consumption power as well as machine operating hours. From this assessment, the managers would have reasonable operating load adjustments changing from rush hours to off-peak hours which decrease the energy consumption as well as energy expense.

Last but not least, the system will alert the faults, breakdown to the operator in the form of sound or display on screen; thus, they could detect to deal with these problems correctly and timely.

Conclusion
Centralized power monitoring and managing system applied in underground mines has been implemented in accordance with the actual problem of coal production so that having useful features for managing the power consumption. Additionally, the system is an effective tool for the logical exploitation organization as well as the typical solution for appropriate energy management.
References

[1] Vu The Nam. and Vu Tuan Anh. Industry and Trade ministry topic, the researching and design automatic electric management and monitoring system to improve energy management efficiency for the coal production, 2010.

Integrating Energy Simulation in Energy Saving Analysis of Taiwan’s Green Hospital Buildings

Po-Han Chen and Meng-Shen Kan

Abstract - Hospitals consume large amounts of energy due to its operation characteristic. The HVAC system designed in hospital buildings runs 24 hours and has special air requirements. This study aims to evaluate the energy-saving efficiency of HVAC system by studying two Taiwan’s green hospital buildings. Results show 40% energy saving efficiency based on integrating VFD (Variable Frequency Drive) chiller, VAV (Variable Air Volume) system and VVW (Variable Water Volume) system for an existed hospital buildings. Verification shows eQUEST simulation results are within margin of error of 7% compared to actual electricity consumption. The new -built hospital is evaluated based on appropriate parameter assumptions and life-cycle analysis (LCA). It is estimated to have 42% energy-saving efficiency based on the initial HVAC energy-saving design. Four cases are studied based on different chiller capacity designs, by implementing higher COP chillers, optimum option can save tremendous energy cost in a life-cycle basis. To sum up, this study demonstrates the integration of computer simulation and LCA in evaluating green hospital buildings, and the HVAC energy-saving performance for hospitals proposed in Taiwan’s green building policy is approximately 40%.

Keywords – Hospital Building, Energy Simulation, eQUEST, Life Cycle Analysis

1 Introduction

The number of hospital building in Taiwan has increased due to domestic care demand rises. Many systems in hospital are operated 24 hours, including electricity system, HVAC (Heat, Ventilation and Air-Conditioning) system, and emergency systems which should all be maintained stably to provide reliable service. Hospitals consume large amounts of energy. Among all the electricity consumption in hospitals, HVAC system accounts for nearly 50%, equipment use accounts for 30%, lighting use accounts for 20% [1].

HVAC system design and operation play an important role in managing energy consumption of hospital buildings. For existed hospitals under renovation or new-built hospitals, it could be helpful and energy efficient if HVAC system can be early evaluated by implementing energy simulation tool to propose the best HVAC design. Taiwan’s green building certification system EEWH has been established more than ten years. HVAC system design is one of the highest score which can be achieved by adopting energy-saving design.

Study of hospital energy use can be divided into two types: building envelope and energy concerning to internal use. Chou [2] studied regional hospitals based on building envelope and concluded that the EUI (Energy use Intensity) value of regional hospital buildings are 225 kWh/m² • year. In Taiwan’s green building certification published in 2012, the EUI value is 254 kWh/m² • year. With more sophisticated technology and equipment used in hospitals, the energy consumption also climbs up compared to years ago. Chen [3] studied energy consumption of hospitals by measuring the actual electricity use, and concluded the most energy-consuming area in hospital is operating room, see Table 1.

<table>
<thead>
<tr>
<th>Area</th>
<th>EUI (kWh/m² • year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out Patient Clinic</td>
<td>234.9</td>
</tr>
<tr>
<td>Rehabilitation</td>
<td>367.3</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>511.4</td>
</tr>
<tr>
<td>Radiology</td>
<td>514</td>
</tr>
<tr>
<td>Hemodialysis Unit</td>
<td>597.1</td>
</tr>
<tr>
<td>General Laboratory</td>
<td>695.7</td>
</tr>
<tr>
<td>Patient Room</td>
<td>276.5</td>
</tr>
<tr>
<td>Emergency Room</td>
<td>448.7</td>
</tr>
<tr>
<td>ICU</td>
<td>562.5</td>
</tr>
<tr>
<td>Operating Room</td>
<td>740.1</td>
</tr>
</tbody>
</table>

Energy-saving design in HVAC system have been
developed and are widely adopted by designers in recent years, including using VSD (variable speed drive) chillers, variable air volume system, and variable speed pump to increase energy use efficiency. The above systems are the most widely adopted HVAC energy saving designs. This study will discuss the energy saving performance based on the above three designs to evaluate the potential energy saving performance of hospital buildings by using energy simulation tool.

<table>
<thead>
<tr>
<th>System</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSD Chillers</td>
<td>Use variable speed drive to adjust motor speed when chillers are in part load capacity operation</td>
</tr>
<tr>
<td>VAV</td>
<td>Control temperature by varying the supply air volume</td>
</tr>
<tr>
<td>VWV</td>
<td>Use variable speed pump to allow pump to operate over a wide speed range</td>
</tr>
</tbody>
</table>

The concept of energy simulation began when need for environmental protection in construction industry increased. Energy simulation provides a tool to evaluate different design alternatives in early designing stage. Famous simulation tools include DOE-2, EnergyPlus, eQUEST etc. Different software has its own designed input interface and load calculation process. Energy simulation tool chosen in this study is eQUEST3-64 version, developers are constantly updating and adding new features to make sure programs are efficient in use. Singh [4] compared the difference in model built-up process and calculation results between eQUEST and DesignBuilder. All input data were the same in both software. However, there was constant difference of 7% in calculation result. The reason was finally tracked down to the area of the experimental models. Design Builder used user-defined dimensions as external boundary, and deducted the external wall thicknesses from the plan reducing the floor area. However, eQuest treated it as internal zone space and kept the full area as conditioned floor area.

Rallapalli [5] compared the simulation process and result between EnergyPlus and eQUEST. Results show that EnergyPlus provides more accurate calculation results than eQUEST, while eQUEST performs better in terms of calculation time. Ke [6] uses eQUEST in examining the energy-saving performance of an office building by applying IPMVP Option D verification process. It is concluded that eQUEST calculation results can be verified within 6% margin of error compared to actual data. Lee [7] concluded that compared to VAV and VWV system, energy-saving design in chillers performs the best energy-saving efficiency. The life cycle year of HVAC systems are usually 5 to 10 years. It is important to consider the time value of money when it comes to evaluating the investment benefit and payback period of HVAC systems. Chan [8] uses saving to investment ratio (SIR) evaluating the best design option for chillers in a building renewal project, energy cost is also considered to provide a comprehensive financial analysis for HVAC system.

2 Methodology

This study aims to use energy simulation tool eQUEST 3.64 to simulate a hospital building and verify with actual electricity data. Based on the experience familiarizing using eQUEST and model built-up process, a new-built hospital is evaluated to propose the ideal HVAC design based on the energy consumption and life cycle cost study. Research flowchart is shown in Figure 2.
2.1 Study Scope

Two hospitals chosen in this study are located in the center of Taiwan. Level of hospital is regional hospital (medium scale with the amount of inpatient bed over 250). This study focuses on studying the electricity consumption and energy-saving efficiency of HVAC systems, heat and natural gas energy are not discussed in this study.

2.2 eQUEST

Input data in eQUEST includes climate data, building footprint, space allocation (zoning), material property/placement of windows and doors, heat load and occupancy (lighting, equipment and people), HVAC system design, indoor temperature and schedule. Building footprint is collected from hospital management manager who provides the original CAD drawings of architectural and HVAC designs. Lighting and equipment density are collected from hospital’s central monitoring center, site visit and data from previous study [3]. Occupancy schedule is investigated through paper survey. The whole input process takes time to finish and check. Figure 3 shows the process of how eQUEST calculates thermal load.

![Thermal Load Calculation of eQUEST](image)

During the life cycle year, it is estimated to have 2% increase annually in energy cost, see Figure 5. In compounding interest calculation, i=2% is used in calculating the time value of money based on current statistical data provided by Taiwan’s financial authority [10].

![Average Electricity Price in Taiwan since 1992](image)

2.3 Life Cycle Analysis

Life cycle analysis in this study considers the investment cost, maintenance and repair (M&R) cost, replacement cost, and energy cost of HVAC system. Life cycle year uses 15 year based on the average life cycle year of chillers. Cost data are collected from HVAC system suppliers. Calculation model is shown in Figure 4.

![Life cycle analysis](image)

3 Case Study-Hospital A

Hospital A was launched since 2007 and was certified as green building with achieving four energy-saving indexes. Electricity cost accounts for 90% of total energy cost. Hospital A is 24-hours monitored in electricity and emergency use, and provide optimum computer control in HVAC system. Building
information of hospital A is shown in Table 4.

<table>
<thead>
<tr>
<th>Content</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Zone</td>
<td>Taichung, Taiwan</td>
</tr>
<tr>
<td>Orientation</td>
<td>N-S</td>
</tr>
<tr>
<td>Floor</td>
<td>B2-6F</td>
</tr>
<tr>
<td>Floor Area</td>
<td>32526 sqm</td>
</tr>
<tr>
<td>Conditioned Area</td>
<td>25328.7 sqm</td>
</tr>
</tbody>
</table>

### 3.1 Data Input

Data input is eQUEST takes time and effort to finish. Try and error method is used during model built-up process to investigate the setting that most fits with actual design. Heat transfer property of building envelope should be input in imperial unit in eQUEST. Table 5 shows input data in SI unit.

<table>
<thead>
<tr>
<th>Category</th>
<th>Content</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building Envelope</td>
<td>Exterior Wall</td>
<td>U=3.5 W/m².k</td>
</tr>
<tr>
<td></td>
<td>Curtain Wall</td>
<td>U=2.4 W/m².k</td>
</tr>
<tr>
<td></td>
<td>Interior Wall</td>
<td>U=0.68 W/m².k</td>
</tr>
<tr>
<td></td>
<td>Roof</td>
<td>U=0.75 W/m².k</td>
</tr>
<tr>
<td></td>
<td>Window</td>
<td>U=5.68 W/m².k</td>
</tr>
<tr>
<td>Lighting Power</td>
<td>Medical Area</td>
<td>15W/m²</td>
</tr>
<tr>
<td>Density</td>
<td>Public Area</td>
<td>12W/m²</td>
</tr>
<tr>
<td></td>
<td>Corridor</td>
<td>10W/m²</td>
</tr>
<tr>
<td></td>
<td>Patient Room</td>
<td>11W/m²</td>
</tr>
<tr>
<td></td>
<td>Office</td>
<td>15W/m²</td>
</tr>
<tr>
<td>HVAC chillers</td>
<td>Centrifugal</td>
<td>420RT*2</td>
</tr>
<tr>
<td></td>
<td>Screw</td>
<td>250RT*2</td>
</tr>
</tbody>
</table>

Building footprint defined in the model is drawn according to the conditioned area line of each zone. There’re no tall buildings around hospital A, sunlight affected by surrounded tall buildings can be ignored. Interior occupancy is mainly divided into 12-hour for office use, and 24-hour for medical use. Model of hospital A is shown in Figure 6.

### 3.2 Verification

Verification result of hospital A is shown in Table 6. Actual data is the average electricity consumption of each month between year 2009 to 2012. Electricity consumption in 2007 and 2008 are excluded due to the uncertainty of space use in the first two years of building operation. Simulation results are within 7% margin of error, which proves that eQUEST can perform reliable simulation result with fully understanding model’s operation characteristic and adequate data assumption.

<table>
<thead>
<tr>
<th>Actual (1000kWh)</th>
<th>Simulated (1000kWh)</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>404.2</td>
<td>419.9</td>
<td>3.9</td>
</tr>
<tr>
<td>415.4</td>
<td>398.8</td>
<td>-4.0</td>
</tr>
<tr>
<td>502.8</td>
<td>470.9</td>
<td>-6.3</td>
</tr>
<tr>
<td>498.8</td>
<td>512.2</td>
<td>2.7</td>
</tr>
<tr>
<td>603.7</td>
<td>603.2</td>
<td>-0.1</td>
</tr>
<tr>
<td>631.0</td>
<td>607.0</td>
<td>-3.8</td>
</tr>
<tr>
<td>639.5</td>
<td>647.1</td>
<td>1.2</td>
</tr>
<tr>
<td>657.8</td>
<td>630.4</td>
<td>-4.2</td>
</tr>
<tr>
<td>627.2</td>
<td>591.7</td>
<td>-5.7</td>
</tr>
<tr>
<td>578.1</td>
<td>592.2</td>
<td>2.4</td>
</tr>
<tr>
<td>502.3</td>
<td>470.6</td>
<td>-6.3</td>
</tr>
<tr>
<td>465.6</td>
<td>454.4</td>
<td>-2.4</td>
</tr>
</tbody>
</table>

### 3.3 HVAC Energy-Saving Analysis

Three energy-saving designs: VSD chillers, VAV system, and WWV system are discussed in this study to investigate the energy-saving performance of HVAC system. In eQUEST, variable speed drive in chillers is set up by inputting coefficients of performance curve to simulate VSD chillers. VAV ventilation system and WWV system is set up by selecting eQUEST program setting from constant to variable. Energy saving efficiency is calculated by formula (1).

\[
ESP = \frac{(WSED - BL)}{BL} \times 100\% \quad (1)
\]

ESP: Energy-saving percentage  
BL: Electricity consumption of baseline model  
WSED: Electricity consumption without HVAC energy-saving design  

Baseline model is set up exactly the same as hospital A with energy-saving design in HVAC system being implemented. In order to calculate how much energy is saved, energy-saving design is excluded in another model to calculate the energy-saving efficiency of hospital A. Simulation result shows the electricity
consumption of HVAC system in hospital A increases 38% annually, almost 20% of total electricity consumption of hospital A. Compared to summer period, HVAC system saves more energy during non-summer time because energy-saving design performs better when systems are in part-load supply.

![HVAC energy-saving efficiency of Hospital A](image)

### 4 Case Study-Hospital B

Hospital B is under designing stage. Compared to hospital A, it is designed to be a mix-use hospital with 8th to 10th floor used as dormitory. In this part, floor 8 to 10 will not be discussed because the air-conditioning system is not served by central HVAC chillers. Hospital B contains less conditioning area than hospital A. Building information of hospital B is shown in Table 7.

#### Table 7. Hospital B - Building information

<table>
<thead>
<tr>
<th>Content</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Zone</td>
<td>Taichung, Taiwan</td>
</tr>
<tr>
<td>Orientation</td>
<td>N-S</td>
</tr>
<tr>
<td>Floor</td>
<td>B2-10F</td>
</tr>
<tr>
<td>Floor Area</td>
<td>31130 sqm</td>
</tr>
<tr>
<td>Conditioned Area</td>
<td>11522 sqm</td>
</tr>
</tbody>
</table>

#### 4.1 Baseline Model Data Input

Baseline model defined in hospital B will be without energy-saving design, which is different from what is defined in model A. Due to the limitation in collecting building information data from hospital B, building envelope is set up according to ASHRAE standard 90.1-2010 [11]. Lighting and equipment density are input according to Chen’s study [3], who has measured the actual electricity consumption of different areas in hospital. The total capacity of chillers is designed to be 600RT. Model of hospital B is shown in Figure 8. Hospital B processes large amount of curtain wall to allow more daylight in the room, which can be expected to reduce lighting power during daytime.

#### Table 8. Data input of hospital B

<table>
<thead>
<tr>
<th>Category</th>
<th>Content</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exterior Wall</td>
<td>U=3.3 W/m².k</td>
<td></td>
</tr>
<tr>
<td>Curtain Wall</td>
<td>U=1.46 W/m².k</td>
<td></td>
</tr>
<tr>
<td>Interior Wall</td>
<td>U=0.68 W/m².k</td>
<td></td>
</tr>
<tr>
<td>Roof</td>
<td>U=0.78 W/m².k</td>
<td></td>
</tr>
<tr>
<td>Lighting Power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Area</td>
<td>13.9 W/m²</td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>16.1 W/m²</td>
<td></td>
</tr>
<tr>
<td>Patient Room</td>
<td>11.3 W/m²</td>
<td></td>
</tr>
<tr>
<td>Office</td>
<td>12 W/m²</td>
<td></td>
</tr>
<tr>
<td>Equipment Power</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient Clinic</td>
<td>42.3 W/m²</td>
<td></td>
</tr>
<tr>
<td>Operating Room</td>
<td>65.2 W/m²</td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td>346.5 W/m²</td>
<td></td>
</tr>
<tr>
<td>Patient Room</td>
<td>15.9 W/m²</td>
<td></td>
</tr>
<tr>
<td>Office</td>
<td>52.1 W/m²</td>
<td></td>
</tr>
</tbody>
</table>

#### Figure 8. Hospital B model in eQUEST

Baseline model simulation result is shown in Table 9. Total electricity consumption is estimated to be 2424500 kWh each year. This will be used as comparison basis for simulating different HVAC design alternatives.

#### Table 9. Hospital B – Baseline Model Simulation Results

<table>
<thead>
<tr>
<th>Month</th>
<th>Electricity consumption of HVAC system (1000kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>164.6</td>
</tr>
<tr>
<td>2</td>
<td>154.9</td>
</tr>
<tr>
<td>3</td>
<td>184.6</td>
</tr>
<tr>
<td>4</td>
<td>198.5</td>
</tr>
<tr>
<td>5</td>
<td>223.6</td>
</tr>
<tr>
<td>6</td>
<td>223.9</td>
</tr>
<tr>
<td>7</td>
<td>242.3</td>
</tr>
<tr>
<td>8</td>
<td>233.8</td>
</tr>
<tr>
<td>9</td>
<td>223.1</td>
</tr>
<tr>
<td>10</td>
<td>222.5</td>
</tr>
<tr>
<td>11</td>
<td>175.5</td>
</tr>
<tr>
<td>12</td>
<td>177.0</td>
</tr>
<tr>
<td>Total</td>
<td>2424.5</td>
</tr>
</tbody>
</table>
4.2 Design Alternatives

Chillers have huge impact in HVAC energy consumption. This study uses the total capacity of baseline chillers (600RT) to generate different capacity combinations. Taiwan’s green building regulations also indicate that at least two chillers should be used in hospital buildings. Four design alternatives are simulated, see Table 10. All designed chillers are centrifugal chillers with VSD.

<table>
<thead>
<tr>
<th>Case</th>
<th>Chillers (RT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>300+300</td>
</tr>
<tr>
<td>2</td>
<td>400+200</td>
</tr>
<tr>
<td>3</td>
<td>200*3</td>
</tr>
<tr>
<td>4</td>
<td>500+100</td>
</tr>
</tbody>
</table>

Energy efficiency calculated in hospital B is different from hospital A. Baseline model defined in hospital B is input without any energy-saving design in HVAC system. Energy efficiency is calculated using formula (2).

\[ ESP = \frac{(BL - ESD)}{BL} \times 100\% \quad (2) \]

ESP: Energy-saving percentage  
BL: Electricity consumption of baseline model  
ESD: Electricity consumption with HVAC energy-saving design

Simulation results shown in Table 11 indicates that case 4 has the lowest HVAC energy consumption. Average energy-saving efficiency is expected to be 40%.

<table>
<thead>
<tr>
<th>Case</th>
<th>Simulated Result (1000kWh)</th>
<th>ESP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1432.9</td>
<td>40.9%</td>
</tr>
<tr>
<td>2</td>
<td>1480.7</td>
<td>38.9%</td>
</tr>
<tr>
<td>3</td>
<td>1500.9</td>
<td>38.1%</td>
</tr>
<tr>
<td>4</td>
<td>1355</td>
<td>44.1%</td>
</tr>
</tbody>
</table>

4.3 Life Cycle Analysis

In this part of the study, life cycle analysis is used to compare different design alternatives and provides recommendation from financial point of view. Maintenance & repair cost is calculated every 5 years. Table 12 shows the life cycle cost of each design alternative. Case 4 has the lowest life cycle cost compared to other three cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>Investment Cost</th>
<th>Total life cycle cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8848.4</td>
<td>66710.7</td>
</tr>
<tr>
<td>2</td>
<td>8208.4</td>
<td>66196</td>
</tr>
<tr>
<td>3</td>
<td>8848.4</td>
<td>67570.3</td>
</tr>
<tr>
<td>4</td>
<td>8548.4</td>
<td>62055</td>
</tr>
</tbody>
</table>

Table 13 shows the payback period of each design case. Four cases have average 4.3 years to based on the energy cost which can be saved.

<table>
<thead>
<tr>
<th>Case</th>
<th>Payback period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

5 Conclusion

This study integrates energy simulation in evaluating hospital buildings. Using eQUEST in simulating complex hospital buildings is reliable with 7% margin of error. Study results also conclude that HVAC energy-saving design proposed by Taiwan’s green building regulation can largely lower the electricity consumption of HVAC system, saving nearly 40% of energy consumption. The life cycle year of proposed design is four years. For future study, automatic data input in eQUEST can be developed to increase project evaluation efficiency.
A Study on Total Electricity Consumption of Medical Buildings

Energy consumption characteristics of hospital buildings

A Comparison of EnergyPlus and eQUEST Whole Building Energy Simulation Results for a Medium Sized Office Building

A Comparison of EnergyPlus and eQUEST Whole Building Energy Simulation Results for a Medium Sized Office Building.

A Comparison of EnergyPlus and eQUEST Whole Building Energy Simulation Results for a Medium Sized Office Building.

The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

References


Measurement Scheme and Automatic Prediction for Ground Vibration Induced by High-Speed Rail on Embankments

Yit-Jin Chen*, Yi-Jiun Shenb and Chi-Jim Chen

*Department of Civil Engineering, Chung Yuan Christian University, Chung-Li, Taiwan
bCECI Engineering Consultants, Inc., Taiwan, No. 323, Yangguang St., Neihu District, Taipei, Taiwan
cDepartment of Computer Science and Engineering, National Taiwan Ocean University, Keelung, Taiwan
E-mail: yjc@cycu.edu.tw, yjs@ceci.com.tw, nyy86@yahoo.com.tw

Abstract - Experience shows that ground vibration induced by high-speed trains can reach levels that cause environmental problems, such as human annoyance, possible damage to old and historical buildings, and interruption of sensitive instrumentation and processes. In engineering, understanding the characteristics of the vibration source, path, and influence distance on various influence factors is essential. Thus, a reliable vibration measurement scheme for highly sensitive vibration areas is proposed to acquire the characteristics of ground vibration induced by passing trains. From a wide variety of field measurements and monitoring using vibration sensors, the main characteristics that affect overall vibration levels include train speed, ground condition, measurement distance, and supported structure. Based on the database of measurement results, the most widely used kernel learning algorithm, support vector machine (SVM) algorithm, is adopted in this research to predict the vibration levels induced by high-speed trains on embankments. Analysis results show that the developed SVM model can predict vibration levels with an accuracy rate of 70%-85% for four types of vibration levels, namely, overall, low, middle, and high frequency ranges. The methodology for developing an automatic prediction system for ground vibration levels induced by high-speed trains is also described in this paper.

Keywords - Environmental Vibration; Ground Vibration Measurement; High Speed Rail; Embankment; Ground Vibration Prediction; Support Vector Machine

1 Introduction

Many countries plan to build high-speed rail (HSR) systems because they enable train services to be faster and more convenient than motorcars and aircrafts. The HSR system has proven to be effective at bridging the gap between vast geographic distances in a relatively short time span, and the benefits of HSR system go beyond travel and eco-friendliness [1]. However, experience also shows that the ground or building vibration induced by train systems can reach levels that cause human annoyance, possible damage to old and historical buildings, and interruption of sensitive instrumentation and processes. Therefore, a reliable vibration measurement scheme is essential to obtain the key characteristics of ground vibration induced by passing trains. Moreover, the development of an automatic ground vibration prediction system is needed, especially in high vibration-sensitive areas.

Among all the possible influence factors for near-field and far-field vibrations, numerous authors [2–6] have concluded that train speed, ground condition, and frequency dependence are the most important factors in evaluating the vibration behavior of high-speed trains. Chen et al. [5] measured the ground vibration induced by Taiwan high-speed trains on embankments with a wide variety of influence factors, including train speed, ground condition, and embankment height. The study included near-field vibration characteristics and far-field vibration propagation. However, a convenient assessment model is also useful for engineers to predict possible vibration effects in the preliminary design stage. Therefore, systematically understanding ground vibration characteristics and establishing a simple vibration assessment for high-speed trains on embankments are important.

In this study, a reliable vibration measurement scheme for highly sensitive vibration areas is proposed. Extensive measurement data from Taiwan high-speed trains on embankments are used to establish ground vibration characteristics. A simple and widely used prediction model [7, 8], the support vector machine (SVM) technique, is then established based on these characteristics as a preliminary ground vibration assessment. Both measured and predicted vibration levels are compared to verify the reliability of the prediction model. The methodology of developing an automatic prediction system for ground vibration levels induced by high-speed trains on embankments is provided.
2 Vibration Measurement Scheme

To perform train-induced ground vibration measurements, some issues must be considered when planning the vibration measurement scheme. They include in-situ conditions and trainset. By understanding the characteristics of these issues, a vibration measurement scheme and data analysis methods are proposed for highly sensitive vibration areas. A background of the important issues, measurement scheme, and analysis methods are briefly summarized in the following sections.

2.1 Important Issues in Train-induced Ground Vibration Measurement

2.1.1 Weather and Site Conditions

Weather conditions may affect vibration measurement results. Weather conditions include air pressure, relative humidity, wind speed, wind direction, rainfall, temperature, and earthquakes [9]. These conditions must be recorded properly when vibration measurement results are evaluated. The qualities and uncertainties of the vibration measurements are closely related to weather conditions. The instruments for vibration measurements must be specified according to weather conditions.

The selection of measurement sites is another critical issue. Free-field measurement sites can be selected with careful consideration to minimize interaction effects. Localized effects on the measurements of man-made features of the particular location should also be minimized. Therefore, sites cannot be on localized topographic features, such as hills, ridges, or valleys, which may create unexpected ground vibrations. Sites should also not be in or on locally anomalous soft soils. Moreover, selected sites should be kept away from potential sources of vibration, such as large motors, large pipelines, large masts, heavy vehicle traffic, and industrial activities.

2.1.2 Train Operation Condition

Train speeds are known to affect the time history and frequency content of the free-field vibration response at various distances from track center. Thus, information on train operation conditions must be obtained when measuring train-induced ground vibration. This information includes train speed and the time when the train passes by certain marked targets. Photoelectric sensors mounted on the parapet of bridges can be used to obtain train speeds and the time a train passes. The signal is “on” during the passage of a train, and the signal is “off” before and after the train passes, as shown in Figure 1. Train speed can be calculated by obtaining the length of the train or distance between two photoelectric sets. The time that a train passes the photoelectric sensor can be obtained using the time history of the signal records. Given that ground vibration measurements have many uncertainties, measuring several passages per train category is recommended.

![Figure 1. The signal during the passage of a train (LS-A and LS-B are two sets of photoelectric sensors with the distances of 25 m)](image)

2.2 Trainset

The trainset of the Taiwan high-speed trains consists of 12 train cars with 10 cars for passengers and two cars as locomotives. The lengths of the passenger car and motive car are 25 and 27 m, respectively; thus, the total length of the trainset is 304 m. The trainset configuration and related dimensions of Taiwan high speed rail (THSR) are shown in Figure 2.

![Figure 2. Configuration of trainset for Taiwan high-speed trains](image)

2.3 Measurement of Ground Vibration

The ISO 804118 requirements are the main reference for installing measuring equipment and selecting measurement points. Basic measuring equipment, such as accelerometers, integrator, and a data acquisition system, is used for this study. The following procedure is used for installing sensors that attach to the ground:
A pit with proper dimensions in which accelerometers can be installed is excavated.

The standard sand is placed on the bottom of the pit to even the excavated surface.

The excavated surface is compacted, and the surface is assured in the horizontal level.

Three-dimensional accelerometers that connect to a steel plate as a firm base are placed on the ground.

The accelerometer direction is set as follows: the X-direction is the direction in which the train moves; the Y-direction is perpendicular to the direction in which the train moves; and the Z-direction is for the direction of gravity.

A set of measuring instruments and equipment is shown in Figure 3. The measured vibration accelerations include X (longitudinal), Y (transverse), and Z (vertical) directions. Only the vertical component (Z direction) is used in the subsequent discussion because some codes and studies [4–6, 10, 11] suggested the Z direction for analysis to simplify the process of vibration impact assessment in general situations.

Adjacent environmental conditions are also essential to avoid any possible interruption during measurement, and ensure that all analysis data are of good quality. Microphones and digital video recorders are installed as auxiliary instruments to record noise and any activity, and such information can be used to evaluate whether other vibration sources interfere with the measurement.

The vibration measurement plan includes near-field and far-field measurements. To establish the near-field vibration database in a consistent reference plane, the distance of the near-field vibration was set at approximately 25 m from the track center. For far-field measurement, four to five measurement points in each site, which were in a straight line and perpendicular to the train alignment, were used to simultaneously measure ground vibration when trains passed through the specific location. The first measurement point (the nearest point from the HSR alignment) was located approximately 25 m from the track center, whereas the distance of the last measurement point (the farthest point from the alignment) was 200 m or so, which was dependent on field conditions. The remaining points were located at average intervals. Figure 4 presents the typical schematic layout of the measurement site. Before measuring, all equipment was synchronized.

Figure 3. Measuring equipment including (1) accelerometers, (2) integrator, (3) and (4) data acquisition system

Figure 4. Typical schematic layout of measuring site

2.4 Data Analysis Method

Based on the characteristics of ground vibration induced by high-speed trains, a range of amplitudes (10–100 dB ref. 1 micro-inch/sec) and frequencies (1–100 Hz) is needed for the assessment. To evaluate the frequency effect, the frequency domain of a one-third octave band for the center frequency range of 1–100 Hz is adopted to describe the velocity vibration level in decibels (dB). The ground vibration level (VL) is expressed in terms of its root-mean-square (RMS) velocity. The RMS velocity level is computed using the following steps:

1. Fast Fourier Transform (FFT) is used to transfer the velocity of time history, \( y(t) \), to the frequency domain. The power spectrum density function (PSDF), \( S_y(f) \), is then calculated as follows:

\[
S_y(f) = \frac{2|Y(f)|^2}{T}
\]  

where \( |Y(f)| \) = FFT amplitude, \( T \) = time interval of \( y(t) \), and \( f \) = frequency (Hz). The suitable time interval (8 s in this study), which covers ground excitation during the passage of the train, is selected from the time history record.

2. PSDF is accumulated from the lower band to the upper band:

\[
E_y(f_c) = \int_{f_l}^{f_c} S_y(f) \, df
\]
where $f_l$, $f_u$, and $f_c$ are the lower band, upper band, and center frequencies, respectively. $E_s(f_c)$ represents the energy summation from $f_l$ to $f_u$. The frequencies of $f_l$, $f_u$, and $f_c$ are based on the definition of the one-third octave band in ANSI [12].

(3) The RMS of $\sigma_y(f_c)$ is calculated:

$$\sigma_y(f_c) = \sqrt{E_s(f_c)}$$

(4) The RMS velocity level (VL), which is represented by dB, is calculated:

$$\text{VL (in dB)} = 20 \log_{10} \left( \frac{\sigma_y(f_c)}{\sigma_0} \right)$$

where the referred velocity in this study is $\sigma_0 = 10^{-6}$ in/s ($= 2.54 \times 10^{-8}$ m/s).

Furthermore, the overall vibration level of one-third octave bands is used to evaluate the total vibration energy [1–9]. The overall vibration level can be transferred from the RMS vibration level of each one-third octave band using the following calculation:

$$\text{VL}_{\text{OA}} = 10 \sum_{k=1}^{n} \log_{10} 10^{\text{VL}(f_k)/10} = \sum_{k=1}^{n} 10 \log_{10} \left( 10^{0.1 \text{VL}(f_1)} + 10^{0.1 \text{VL}(f_2)} + \ldots + 10^{0.1 \text{VL}(f_n)} \right)$$

where $\text{VL}_{\text{OA}} =$ overall vibration level in decibels, $f_k =$ center frequency of each one-third octave band (1 … 100 Hz for the frequency of $f_1$ … $f_n$), and $\text{VL}(f_k) =$ vibration level for each center frequency.

### 2.5 Database for Analysis

A series of measurements was performed for this study. These measurements comprised different ground conditions, embankment heights, and train speeds. Table 1 lists the basic information for these measurements. The measurement distances ($D_1$–$D_4$) were 22–203 m from the track center. These selected embankments have various heights ranging from 3.6 m to 6.8 m.

A wide variety of ground conditions and ground shear wave velocities ($V_s$) were considered. The ground conditions included sand/silt/clay soils (silty clay and sandy silt), gravel, and rock (sandstone). The ground conditions of these embankments ranged from soft to hard ground.

Ground shear wave velocity was used as an indicator to describe “soil stiffness.” The value of $V_s$ increases with increasing soil stiffness. The National Center for Research on Earthquake Engineering (NCREE) measured the ground shear wave velocity throughout Taiwan. The site where NCREE measured the ground shear wave velocity is adjacent to the location of the ground vibration measurement in the present study. Both have relatively similar ground conditions and geotechnical parameters. The average $V_s$ ranges from 170 m/s to 650 m/s. The values of average $V_s$ were obtained from the ground surface to 10 m deep, based on the suggestion by Yoshioka [6], because such values represent surface wave analysis. The database was considered a sufficient representative sample for evaluating ground vibration characteristics.

### Table 1. Basic information of measurement sites

<table>
<thead>
<tr>
<th>Site No.</th>
<th>Soil Type</th>
<th>Height (m)</th>
<th>$V_s$ (m/s)</th>
<th>$D_1$ (m)</th>
<th>$D_2$ (m)</th>
<th>$D_3$ (m)</th>
<th>$D_4$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>silty clay</td>
<td>5.3</td>
<td>170</td>
<td>28</td>
<td>53</td>
<td>78</td>
<td>101</td>
</tr>
<tr>
<td>2</td>
<td>sandy silt</td>
<td>5.2</td>
<td>230</td>
<td>33</td>
<td>68</td>
<td>93</td>
<td>118</td>
</tr>
<tr>
<td>3</td>
<td>gravel</td>
<td>3.6</td>
<td>430</td>
<td>34</td>
<td>69</td>
<td>149</td>
<td>203</td>
</tr>
<tr>
<td>4</td>
<td>sandstone</td>
<td>6.8</td>
<td>650</td>
<td>22</td>
<td>47</td>
<td>122</td>
<td>202</td>
</tr>
</tbody>
</table>

### 3 Vibration Measurement Results

After performing FFT in Section 2.4, the measurement results are shown in Table 2. The vibration level of 21 frequencies for the one-third octave band with 1–100 Hz is classified based on low (1–8 Hz), middle (10–25 Hz), and high (31.5–100 Hz) frequency ranges. These frequency ranges were selected based on an observation of the numerous practical results and related literature [2, 3, 13]. The overall vibration level for each frequency range can then be computed using Eq. (5).

Based on a wide variety of field measurements and monitoring with vibration sensors, the main characteristics, which influence the vibration level, include geological condition, attenuation distance, and train speed, which are shown in Figure 5. The following section is a brief summary of Figure 5.

(1) For measurement distance of 25 m, the near-field overall ground vibration level slightly increases with increasing train speed.

(2) The near-field overall ground vibration level decreases with increasing ground shear wave velocity. A softer ground leads to a higher near-field vibration level.

(3) The attenuation of ground vibration is obvious as the distance from the track center increases.

### 4 Prediction Method

Given that SVM algorithm [7, 8] is probably the most widely used kernel learning algorithm, it is
Table 2. Vibration measurement results

<table>
<thead>
<tr>
<th>Size</th>
<th>$V_L$ (m/s)</th>
<th>Speed (km/h)</th>
<th>$D_1$ (dB)</th>
<th>$D_2$ (dB)</th>
<th>$D_3$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$V_L$</td>
<td>$V_{L_{max}}$</td>
<td>$V_{L_{avg}}$</td>
<td>$V_{L_{max}}$</td>
<td>$V_{L_{avg}}$</td>
</tr>
<tr>
<td>170</td>
<td>150</td>
<td>63.5</td>
<td>69.5</td>
<td>66.4</td>
<td>64.5</td>
</tr>
<tr>
<td>170</td>
<td>200</td>
<td>66.6</td>
<td>67.0</td>
<td>66.1</td>
<td>64.8</td>
</tr>
<tr>
<td>170</td>
<td>300</td>
<td>68.4</td>
<td>69.6</td>
<td>67.8</td>
<td>65.9</td>
</tr>
<tr>
<td>170</td>
<td>195</td>
<td>68.0</td>
<td>68.0</td>
<td>66.6</td>
<td>65.0</td>
</tr>
<tr>
<td>170</td>
<td>200</td>
<td>66.6</td>
<td>67.0</td>
<td>66.1</td>
<td>64.8</td>
</tr>
<tr>
<td>170</td>
<td>300</td>
<td>68.4</td>
<td>69.6</td>
<td>67.8</td>
<td>65.9</td>
</tr>
<tr>
<td>170</td>
<td>195</td>
<td>68.0</td>
<td>68.0</td>
<td>66.6</td>
<td>65.0</td>
</tr>
</tbody>
</table>

Figure 5. Relationships of $V_{L_{max}}$ and (a) train speed, (b) ground shear wave velocity, (c) attenuation distance.

adopted in this research to predict the level at which high-speed trains induce vibration based on the database of measurement results in Section 3.

4.1 SVM Prediction Method

The proposed prediction method in this research is mainly based on the SVM technique, a widely used machine learning tool for reconfirmation. Machine learning is a subfield of applied statistics, which trains on a collected sample dataset and generalizes rules from previous experiences for later classifier applications. The training data with unknown probability distribution are usually applied to extract some general principles and distribution for future predictions on the new testing data. Several types of machine learning algorithms can be categorized based on trained inputs or desired outcomes, such as supervised, unsupervised, semi-supervised, and reinforcement learning mechanisms. Recently, the SVM kernel method, a supervised learning model, has become one of the most popular classification algorithms by training known features.

To construct such a classification model, both positive and negative data classes should be provided as training examples in advance, and a trained SVM model
is then constructed according to the selected features. In the feature space, all learning objects are divided by a hyper plane with a separable margin as wide as possible. The query objects are mapped into the same feature space and assigned to one of the two defined categories based on the locations of the testing object.

This study adopted an SVM classification tool (LIBSVM), which was developed by Lin [7, 8], to predict train-induced vibration levels. All measurement data with defined features were trained, and an SVM model was constructed to predict vibration levels. The selection of classification features, evaluation of different kernel transformation techniques, and prediction results on benchmark datasets are shown and discussed in detail in the following sections.

4.2 Prediction Process

In this research, ground shear wave velocity (V_s), train speed, and distance from track center (D) are defined as the features in the SVM model. The radial basis function is selected as the kernel type, and the function of multi-class classification is employed in this SVM system. Based on the measured results of possible minimum and maximum vibration levels, the output of the SVM model is classified as four groups of ground vibration levels, as shown in Table 3. The ground vibration levels for Group 1-4 are 30-40 dB, 40-50 dB, 50-60 dB, and 60-70 dB, respectively.

Table 3. Group of ground vibration level

<table>
<thead>
<tr>
<th>Group No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>vibration level (dB)</td>
<td>30-40</td>
<td>40-50</td>
<td>50-60</td>
<td>60-70</td>
</tr>
</tbody>
</table>

The training and prediction processes are shown in Figure 6. In the training state, the system reads the input data at the beginning. It then extracts and normalizes the data to the region (0, 1). After transforming the data to the normal form, the system inputs the data to the LIBSVM and begins the training process. The system used the n-fold cross validation, which chooses one fold as the prediction fold. Another fold is used as the training fold for building the hyper plane that is used for prediction. If the prediction results are better than before, the parameter is changed. This process is repeated several times until the system finds that the result is in the local maximum. After the training process, the system uses the hyper plane to predict and classify the vibration level.

4.3 Prediction Results

Table 4 presents the prediction results of ground vibration levels using the aforementioned prediction process. The analysis results show that the developed SVM model can predict vibration levels with an accuracy rate of 70%–85% for four types of vibration levels, namely, overall, low, middle, and high frequency ranges. The accuracy rate is defined as the number of correct classification divided by the whole test sets.

Table 4. Prediction results of vibration level

<table>
<thead>
<tr>
<th>Type of vibration level</th>
<th>low</th>
<th>middle</th>
<th>high</th>
<th>overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>accuracy rate</td>
<td>0.84</td>
<td>0.72</td>
<td>0.78</td>
<td>0.75</td>
</tr>
</tbody>
</table>

5 Automatic Prediction System

The methodology for developing an automatic prediction system for high-speed train-induced vibration levels is also proposed in this study. This automatic prediction system provides a relatively simple method for developing vibration level estimates of high-speed train-induced vibration that can be compared with the acceptability criteria given in Section 5.1. This proposed system can help engineers determine whether vibration-sensitive land uses are close enough to the proposed HSR system, as well as the impact from ground-borne vibration. An example is provided to demonstrate this automatic prediction system.
5.1 Vibration Impact Criteria

The vibration impact criteria in this system cover the human response to building vibration, as well as the building that houses vibration-sensitive instruments and tools. They are briefly introduced in the following sections.

5.1.1 Vibration Criteria for High-speed Ground Transportation (U.S.A.)

The vibration criteria for high-speed ground transportation [2] are presented in Table 5. The criteria are based on the maximum levels for a single event. The criteria account for the variation in land use and frequency of events. Table 5 shows that land use is classified into three categories: building, residence, and institution. The frequency of an event is defined as more or fewer than 70 vibration events per day.

5.1.2 Generic Vibration for Vibration-Sensitive Equipment

The Vibration criterion curves are commonly used in the design of facilities that house vibration-sensitive instruments and tools [13]. Vibration criterion curves from a set of one-third octave band velocity spectra have been labeled VC-A to VC-E. The criteria apply to people and vibration-sensitive equipment, as described in Table 6. The criteria for vibration-sensitive equipment are more restrictive than those for human responses, as observed from Table 6. A detailed review is needed for situations involving buildings with high sensitivity equipment near a railway.

5.2 Automatic Prediction Methodology

The vibration levels at specific buildings are estimated by reading output values from the proposed prediction system. They are also estimated to account for factors, such as ground shear wave velocity, train speed, and distance from the track center. Users must input the position of the alignment for high-speed rail, train speed, and ground shear wave velocity along the high-speed line. The system calculates its distance from the HSR and computes the predicted vibration level automatically.

5.3 Example of Automatic Prediction System

Example output results of the proposed automatic prediction system are shown in Figure 7. The vibration levels along the HSR corridor are shown in different colors for each group of vibration levels. In Figure 7, the black region expresses the alignment of the HSR. The red, yellow, blue, and green regions represent vibration levels 4, 3, 2, and 1, respectively. The ground vibration level gradually decreases from 4 to 1.

The engineer can check the vibration criteria in Section 5.1 based on land use, and determine whether the vibration level induced by HSR exceeds the vibration criteria. If the vibration levels in some regions exceed the vibration criteria, a more detailed assessment is needed when potential problems exist. A detailed analysis is then undertaken during the final design stage to accurately define the level of impact and design mitigation measures. However, detailed analysis is not usually required if the mitigation measure consists of relocating the alignment of HSR systems.
6 Conclusions

The measurement scheme and analysis methods of ground vibration induced by trains passing on an embankment structure have been presented. The related issues for ground vibration measurements have also been discussed. Based on a wide variety of field measurements and monitoring with vibration sensors, the main characteristics that affect the overall vibration level include train speed, ground condition, measurement distance, and supported structure.

The analysis results show that the developed SVM model can predict vibration levels with an accuracy rate of 70%–85% for four types of vibration levels: overall, low, middle, and high frequency ranges.

The methodology for developing an automatic prediction system for vibration levels induced by high-speed trains has been proposed in this study. An example has also been provided to demonstrate this automatic prediction system. The computer implementation for this proposed system will be developed as future work.

Since the proposed prediction model is based on field measurement results, more field measurement data should be obtained to improve the accuracy of this automatic system. The proposed automatic prediction model can help engineers improve the accuracy of preliminary analysis, and be applied to a large number of assessments.

Acknowledgment

This study was supported by the National Science Council, Taiwan, under contract number NSC 100-2221-E-033-073-MY3.

References


GIS-BIM Framework for Integrating Urban Systems, Waste Stream and Algal Cultivation in Residential Construction

Daniel Castro-Lacouture\textsuperscript{a}, Steven Jige Quan\textsuperscript{b}, and Perry Pei-Ju Yang\textsuperscript{c}

\textsuperscript{a}Associate Professor, School of Building Construction, Georgia Institute of Technology, Atlanta, GA 30332 USA
\textsuperscript{b}Graduate Research Assistant, Built Environment Sustainable Technology (BEST) Laboratory, Georgia Institute of Technology, Atlanta, GA 30332 USA
\textsuperscript{c}Associate Professor, School of Architecture, School of City and Regional Planning, Georgia Institute of Technology, Atlanta, GA 30332 USA
E-mail: dcastro@gatech.edu, steven@gatech.edu, perry.yang@coa.gatech.edu

Abstract
Among alternative energy sources for residential buildings, algae technology has emerged as a promising option due to its closed-loop configuration and the ability to produce biofuel energy while reducing waste stream flow and capturing carbon. Furthermore, this technology has the potential of integrating resource and waste management, and can be complemented with other alternatives, such as photovoltaic, wind or fuel cells. This paper provides a framework for integrating information from geographic information systems, building information models, construction schedules, construction cost estimates, and constructability reviews. The integration is aimed at designing an algae-powered residential building environment at the level of urban neighbourhood, in which the algae technology is taken as a design intervention to promote energy performance and carbon reduction within the urban system. This framework couples the design intervention with impact simulations influenced by geographic contexts, construction considerations, and digital building technology. By extending the system boundary from a closed algae cultivation system to an open neighbourhood-scale urban environment, urban renewable resources such as energy, water, material and carbon flows are connected to the algae cultivation process. The framework would further advance the possibilities for sharing information among planners, architects, engineers and construction managers for innovative closed-loop sustainable energy systems in residential construction. This approach will address challenges such as cost, governmental incentives, regulatory barriers, or need of research and development that could overcome limitations for automating predesign, design, construction and facility management processes.

Keywords –
Algae technology, GIS, BIM, residential construction

1 Introduction
There is a potential for the use of algal cultivation and processing to manage the liquid and solid waste from a residential community and recycle the carbon stream into energy of sufficient quality to power the dwellings in a sustainable manner \cite{1, 2}. Once maximum reductions are made with respect to resource utilization and maximum efficiencies are realized using green, conservation-driven design principles, the evaluation of the waste streams must be considered for opportunities for reuse, recycling, and treatment processes that maximize the extraction of resources. Novel sources of inputs may then be developed that take advantage of the holistic approach to resource management.

Primary residential waste streams include wastewater and municipal solid waste. In addition, thermal and other energy losses due to inefficiencies require additional demands on resource inputs. In addition to the current first-choice renewable energy generating technologies available to residential communities (e.g., solar photovoltaic, solar thermal, wind turbine), novel technologies are currently under development that may be able to address the desire for sustainable energy generation while, at the same time, address waste stream management issues through resource recovery. Figure 1 identifies selected materials and energy inputs and outputs for a typical single-family home. In addition to the selection of an appropriate strain of microalgae, algae cultivation requires four primary inputs: water, nutrients (primarily nitrogen and phosphorous), carbon (primarily carbon dioxide), and sunlight.
Solid wastes that were readily biodegradable are directed to the biogas digester where anaerobic bacteria convert the majority of the organic carbon into methane and carbon dioxide. These gases are combusted in a gas turbine where electricity and heat are supplied to the residence. The flue gases (now CO2 and residual N2) are sparged into the algae cultivation pond. The liquid waste from the digester (liquor) is very rich in organic nutrients and is disinfected using UV technology and the residual nutrients added to the algae pond. Harvested algae is processed to remove natural lipids, assumed to be 20% of the algae mass. That oil can be used directly in a diesel generator, again to provide electricity and heat to the residence, and the flue gas from the diesel generator is sparged into the algae pond. The remaining 80% of the algal biomass is then directed to the biogas digester, providing additional methane for the gas turbine. The 80% fraction of algal biomass may also be used directly as food or animal feed. Examples of feed options include small food-producing animals, like chickens or ducks, or aquaculture opportunities, such as shrimp or fish farming. The wastes from the food production would also be amenable to inclusion in the biogas digester to provide additional biomass for energy generation and nutrient recycling. Perspectives for the feasibility of microalgae technology has been discussed elsewhere [1, 2, 3]. However, a more integrative system of sustainable housing and development presumes the interrelationship of parameters and metrics between planning, architecture and building construction.
density and waste stream management will need to be established.

2.2 Hydrological cycle

Hydrological cycle: the hydrological balance and the sewage flows will affect the amount of water required to manage algae ponds. To prevent the algae ponds from drying out or becoming too salty, conditions that would kill the algae, a steady supply of freshwater is needed to replenish the evaporating water [4]. The hydrological conditions of detention, ground water discharge and the grey water reuse will be explored. The urban site tends to have higher level of impervious surfaces and lower level of infiltration, and thus contains less groundwater availability to contribute to the base flow between storms especially in the dry period [5]. However, urban sites normally contain higher wastewater flow that provides potential steady supply of grey water for re-use in the algae pond with on-site waste water treatment such as constructed wetlands. In suburban and rural settings, the hydrological cycle is to be designed and managed by low impact development that works with nature to manage storm water as close to its source as possible [6]. The design of algae pond tends to choose a landscape approach that is to be integrated in the landscape ecological network. The research also adopts land cover based performance measure to connect relationship between the land cover type information and the annual hydrological parameters such as annual mean surface runoff and infiltration [7]. The hydrological model helps to determine location of algae pond and to quantify the area required for production of algae.

2.3 Solar availability

Incoming solar radiation or insolation received from the sun is the primary energy source that drives algae biological processes. The algae need sunlight to photosynthesize. Topography and built environment are major factors that determine the spatial variability of insolation. Across the urban gradient from urban, suburban to rural, the amount of insolation decreases accordingly, as seen in Table 1. In urban settings, the amount of insolation is affected by not only topography but also by surrounding buildings. While in the rural setting, the main factor determining insolation is orientation, slope, aspect and elevation. The solar availability analysis will also provide the information to support the location decision of the algae pond.

2.4 BIM-GIS integration

The integration of algal cultivation system and urban system calls for a new information technology management system to facilitate the construction process. Such information management system is spatial and locational, and moreover it deals with issues across different spatial scales. Such information system should involve the technologies that could manage the system data with the urban spatial modelling.

However, there is still a big gap between the spatial information technologies that apply at the micro spatial scale and at the macro scale. At the building level as the micro scale in the urban system, BIM (Building Information Modelling) goes beyond the traditional CAD (Computer Aided Drafting) approach for 2D representation to a new paradigm for placing real information in the model database to virtually construct a building [8]. BIM is designed to support the building maintenance and the life-cycle development [9]. At the urban level as the macro scale, GIS (Geographic Information System) aims at managing the database with spatial reference [10]. Different purposes and emphasis on different scales of these two information technologies lead to their various capacities in dealing with information at different levels and their distinct methods in locating objects. First, BIM provides strong functions to handle the “indoor” data, at a relatively micro spatial scale, while GIS offers many tools to organize the “outdoor” data, e.g. streets, land parcels and terrains, at the macro spatial scale. Second, BIM places the building as the system boundary and models the relative coordinates of the objects within the system, while GIS assumes the Earth as the system and places every object with the geographic coordinate systems and the world map projection. Third, BIM emphasizes the life cycle of the building construction and therefore has a clear temporal dimension, while GIS focuses more
on managing the interaction of spatial, social and economic information at a given time. Fourth, and the most important, ArcGIS has the advantage of network analysis based on geo-locations, while BIM is relatively weak on this aspect. Besides, their data formats are based on different protocols and therefore cannot be directly opened by each other.

Due to these differences of the two informational technologies, they are usually utilized at separate academic or practice fields. ArcGIS is often used at urban planning, transportation planning and urban design analysis, while BIM is commonly applied to architectural design and building construction management, though it could be related to planning parameters at larger scales [8].

However, in the combined algae-urban system, the information system faces new challenges to deal with cross-scale and system-oriented issues, with the following questions at hand: where is the resource? How are algal cultivation systems of different levels integrated to buildings and neighbourhood and how do they perform? How can the resource (input), algae system and output be connected? How could we manage the construction process to realize such networks?

These questions call for a performance-based information system that integrates ArcGIS and BIM. In this system, ArcGIS and BIM generate both project-specified data and urban-based data to improve the accuracy of planning models. At the same, their respective analytical tools enhance the overall performance analysis capacity of the model. The integration invites the extension of boundaries in the model and the thoroughness of the performance assessment. The data availability and analysis applicability in ArcGIS and BIM are shown in Table 2 ("G" means that the analysis can be done with GIS, while "B" means with BIM).

The integration of ArcGIS and BIM has different forms and applications at three scales. Figure 4 shows the management of the algal cultivation system in building construction using the integrated ArcGIS-BIM system.

<table>
<thead>
<tr>
<th>Data</th>
<th>GIS data</th>
<th>BIM data</th>
<th>Solar avail.</th>
<th>Waste stream</th>
<th>Recycled water</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>building system</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>urban infrastructure</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>massing</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>building operation</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>fenestration</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>population</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>occupation</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>material</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Algal</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>cultiv. consum.</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>Constr. mgmt.</td>
<td>G</td>
<td>G</td>
<td>B</td>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. GIS-BIM informational system for managing algal cultivation system in residential construction

At the building scale, the BIM is intensively used to manage the data of material, fenestration, building-level algal system and to assess the waste stream, solar gain, energy consumption and input-output of the algal system. Then this information is transferred into ArcGIS as the node data in a network. At the building block scale, both ArcGIS and BIM are utilized. The dataset includes more information about the building block layout, the infrastructure network, the population and the algal pond operation data. The analysis extends from within the building to include the more accurate solar gain measurement with external shadings, as well as the network analysis with input and output of the algae system in buildings and in the ponds.

At the neighbourhood scale, the dataset and the analysis focus more on urban level. Using a GIS-based network analysis, the dynamic model can analyse both simple and complex waste stream and water recycle systems. GIS network analysis will provide opportunities for maximizing the distribution of waste and recycled water and minimizing construction costs. The analysis also supports all stages of waste stream...
management analysis, defining the topology of a sewer network, the specification of sewage flow contribution parameters, the allocation of sewage-contributing areas to sewer manholes, hydraulic analysis and displaying the analysis results, as illustrated in Figure 5.

Figure 5. GIS-based network analysis

The solar radiation tool allows mapping and analysing the effects of the sun over a geographic area for specific time periods. It accounts for atmospheric effects, site latitude and elevation, steepness of slope and compass direction of aspect, daily and seasonal shifts of the sun angle, and effects of shadows cast by surrounding topography or buildings. The solar radiation tool calculates the sum of the direct radiation and diffuse radiation through sky. It is based on spherical approach of 3D GIS analysis developed by one of the authors [11]. The output radiation rasters will have units of watt hours per square meter (WH/m²).

3 Integrating construction information

The integrated framework incorporates construction information inherent to the facility, such as schedule, budgeted cost, and also has the possibility of including as-built cost, as well as constructability and safety considerations.

3.1 Background and research questions

There has been extensive research on recognizing construction elements on site and measuring progress using information technology [9, 12-15]. Construction progress monitoring is possible through a cross-correlation of collected progress data using GIS and existing BIM information. This enhances the ability to determine progress by comparing the planned schedule and cost presented in the BIM model and actual progress, and helps field managers monitor the work progress conceptually. A robust system of performance metrics needs to back this process up. BIM comprises useful design information, overused during the design phase but overlooked during construction. Schedule and cost information provided by BIM and compared with GIS-based as-built data may provide decision makers with a useful tool for project monitoring with reliable progress indicators. Some research questions addressed by this approach include: what are the integration protocols required to compare BIM and GIS-based information from the construction of a single building, multifamily housing or a neighbourhood? Can Earned Value progress indicators be extracted from the comparison? How timely is the method to the decision-makers?

A 5D (3D+time+cost) IFC-based model could be created in which as-built information is connected to BIM allowing a baseline for progress monitoring. The as-built cost information would be available on demand. This part of the research would analyse an automated Earned Value Analysis (EVA) system for progress monitoring based on comparing as-built with as-planned physical progress. Some of the research questions addressed include: how cost information could be linked up to 4D models to represent a rich baseline for as-planned data in Earned Value Analysis (EVA)? To what level of detail, 4D model needs to represent the as-planned progress? How non-constructing activities are represented in the as-planned model?

3.2 Sources of preconstruction data

A detailed cost estimate for the residential construction featuring the algae-based power system will follow Construction Specifications Institute (CSI) Masterformat, and will be properly populated with cost data from standard manuals (e.g., RS Means, Timberline), local builders, specialist trade subcontractors, and material suppliers. Costs will be adjusted based on geographic location, and the standard manuals will account for the adjustment. Also, local contractors and material suppliers can provide relevant prices of construction materials, equipment, and labour. Scheduling information will be computed from the standards manuals using daily production values for labour and material installation, and expected durations from equipment productivity manuals. Specific tasks to be included in the estimate include:
1. Investigate the costs of installing and putting in operation the algae system, including the collection of wastewater (e.g., toilets, sinks, showers); delivery to the pond; transfer to the bioreactor; installation of bioreactor and all ancillary systems.

2. Establish a maintenance plan for all systems mentioned above and estimate training requirements for safe and on-going operation.

3. Calculate life expectancy of the system, as a whole or its parts and estimate the replacement cost.

4. Analyse entire value chain from a cost perspective; track and evaluate costing from the research and development phase of the system’s life, through to the decline and eventual conclusion of its life.

Additional considerations that impact cost are associated with the life cycle assessment, ownership, delivery, and technology implementation. Such considerations include:

1. Evaluate the system’s environmental aspects and its potential impacts in general and specifically. Assess the raw materials (inputs) needed to effectively operate the system and its probable releases to the environment.

2. Investigate possible reuse, maintenance and waste management.

3. Investigate the embodied energy of the system put in place and operating.

4. Analyse different deployment schemes for the system: plug-and-play, when the decision to invest in the system is made by the owner or occupier of the dwelling; company control: a group of consumers provide the site for the system to be installed and it is operated by an Energy Service Company (ESCo) or other third-party energy supplier; community micro grid: a group of individuals groups together to provide some or all of their collective energy needs. They own and operate the system.

5. Investigate different forms of ownership for end users and their implications in initial and running costs, maintenance and upgrades.

6. Establish training levels and parameters for the different deployment schemes.

7. Study the delivery of the service and rate structures.

8. Most systems being evaluated are not mainstream, from water separating toilets, sinks, etc., to the retention ponds for wastewater and the bioreactor itself. Investigate the safety implications of installing such a system in a residential environment. For example: maintaining a wastewater pond in proximity to residential dwellings is not the same as a storm water retention pond. Determine the real and perceived safety issues related to the installation and operation of such a system of energy generation.

9. Determine the different levels of control needed to operate the system safely and who is better qualified to handle the risk.

10. Management: what are the personnel, cost and liability implications of maintaining safety while operating the system.

11. Design (differences with existing systems affected): study the design and construction implications and code issues related to implementing the design in a typical residential environment.

12. Experiment and model issues of scalability. Siting concerns, size of reactor, size of wastewater pond, distribution of energy produced, maintenance, handling of by-products and recycling.

3.3 Comparing preconstruction and as-built data

As-built elements will be generated from GIS for outdoor or visible objects as they are being installed. This information will be moved to the data fusion bin, or an IFC-compatible file that features [x,y,z] coordinates in raster version, as well as element properties and sequencing information in 3D Object Exchange format. The coordinate origin [0, 0, 0] is elicited from the BIM model, and the x, y, and z positions of each tag in feet or meters are defined with respect to that origin, thereby constructing an IFC file. This file will be merged to the existing GIS data in the data fusion bin. In the third step, a task allocation protocol activated in the data fusion bin will revise the integrity of the data as to contain [x,y,z] coordinates of the objects’ centroids and exterior corners. Data, formatting integrity and file compatibility will be periodically checked for comparison adequacy with IFC-based as-planned data provided by BIM, as well as to generate progress reports and visualization protocols. A client program using stream communication protocol will be required to retrieve data from the IEEE 802.11 network hub and to establish the parameters by which the data fusion bin will proceed with the combination with GIS-based data, as specified in the task allocation.
protocol. The default data output stream provided from the processing hub to the client program is:
<Data Header>, <tag #>, <X>, <Y>, <Z>, <timestamp>, <unit><LF>

The data output from the hub can be modified to provide increased resolution in the timestamp field. <Data Header> represents the tag dimensional information. Expected values for <Data Header> are nomenclature representing construction elements. For example, C #, which indicates an internal column with an extension number whose value will denote possible configurations as specified in the BIM model; <tag #> is the tag ID; <X>, <Y>, <Z>, are the calculated tag coordinates in feet or meters with respect to a user supplied origin; <timestamp> represents the hub system time. The format for timestamp is UTC time, day, and year the data was computed. An additional data field is also programmed to indicate the quality of data provided from the hub and the type of element being positioned; finally, <LF> is a Line Feed character (with ASCII code = 0x0A), to terminate a location data string. This data structure will enter the data fusion bin, where it will be merged with IFC-based GIS data through a task allocation protocol.

In order to perform the element matching, a software application will be developed to export the data fusion bin, using an IFC Industry Foundation Classes format (based on the ISO-STEP EXPRESS language which is an industry standard for building information exchange and consistent data representations), to a BIM application (i.e., Revit, ArchiCAD, Bentley). The data file will have the following structure:
ISO-11253-32;
HEADER
FILE_DESCRIPTION ($):
FILE_NAME ($,'Sat Jan 25 14:03:45
2014',(AbC),('Shingle'),'ASX_Roof
v.0.1'ASX_Project',$);
FILE_SCHEMA ('ASX');
ENDSEC;
DATA;
#1 = CATALOG ('Argos1 (644)', 'Argos1
CONSTRUCTION LIBRARY');
#2 = GEOMETRY ('Argos1 (101)', 0.58, $,'Corner', 'Real');
#3 = BOUNDARY ('Argos1 (102)', 0.61, $,
'Boundaries', 'Real');
#4 = SCHEDULE ('Argos1 (103)', 0.63, $,
'Timestamp', 'Real');
ENDSEC;
END-ISO-11253-32;

The BIM application generates the as-built model with the data imported, superposing both GIS and BIM models.

4. Conclusions

This paper addresses a conceptual model for integrating input/output information involved in developing an innovative technology for providing reliable and efficient power generation to improve the energy consumption in American homes. This model impacts the design of household and residential communities: by integrating domestic waste stream management with the concurrent production of energy, substantial reductions in energy and significant conservation of water and waste can be achieved. Algae technology has emerged as a promising option due to its closed-loop configuration and the ability to produce biofuel energy while reducing waste stream flow and capturing carbon. This paper provides a framework for integrating information from geographic information systems, building information models, construction schedules, construction cost estimates, and constructability reviews. The integration is aimed at designing an algae-powered residential building environment at the level of urban neighbourhood, in which the algae technology is taken as a design intervention to promote energy performance and carbon reduction within the urban system. This framework couples the design intervention with impact simulations influenced by geographic contexts, construction considerations, and digital building technology. By extending the system boundary from a closed algae cultivation system to an open neighbourhood-scale urban environment, urban renewable resources such as energy, water, material and carbon flows are connected to the algae cultivation process. The framework would further advance the possibilities for sharing information among planners, architects, engineers and construction managers for innovative closed-loop sustainable energy systems in residential construction. This approach will address challenges such as cost, governmental incentives, regulatory barriers, or need of research and development that could overcome limitations for automating predesign, design, construction and facility management.

5. References


Prediction of NO\textsubscript{X} Vehicular Emissions using On-Board Measurement and Chassis Dynamometer Testing

S. D. Oduro\textsuperscript{a}, S. Metia\textsuperscript{a}, H. Duc\textsuperscript{b}, G. Hong\textsuperscript{a} and Q. P. Ha\textsuperscript{a}

\textsuperscript{a}Faculty of Engineering and Information Technology, University of Technology, Sydney, Broadway, NSW 2007, Australia
\textsuperscript{b}Office of Environment and Heritage, Lidcombe, NSW 1825, Australia

E-mail: Daniel.SethOduro@student.uts.edu.au Metia.Santanu@student.uts.edu.au Hiep.Duc@environment.nsw.gov.au Guang.Hong@uts.edu.au Quang.Ha@uts.edu.au

Abstract -
Motor vehicles' rate models for predicting emissions of oxides of nitrogen (NO\textsubscript{X}) are insensitive to their modes of operation such as cruise, acceleration, deceleration and idle, because these models are usually based on the average trip speed. This study demonstrates the feasibility of using other variables such as vehicle speed, acceleration, load, power and ambient temperature to predict NO\textsubscript{X} emissions. The NO\textsubscript{X} emissions need to be accurately estimated to ensure that air quality plans are designed and implemented appropriately. For this, we propose to use the non-parametric multivariate adaptive regression splines (MARS) to model NO\textsubscript{X} emission of vehicle in accordance with on-board measurements and also the chassis dynamometer testing. The MARS methodology is then applied to estimate the NO\textsubscript{X} emissions. The model approach provides more reliable results of the estimation and offers better predictions of NO\textsubscript{X} emissions. The results therefore suggest that the MARS methodology is a useful and fairly accurate tool for predicting NO\textsubscript{X} emission that may be adopted by regulatory agencies in understanding the effect of vehicle operation and NO\textsubscript{X} emissions.

Keywords -
Nitrogen Oxide; On-Board Emission Measurement System; Chassis Dynamometer Testing System; Emission

1 Introduction
Vehicular emissions can bring serious impacts on the air quality, and have thus received increasing research concerns [1]. Outdoor air pollution is estimated to cause 1.3 million annual deaths worldwide [2]. Road transport often appears as the single most important source of urban pollutant emissions in source apportionment studies [3]. In the coming decades, road transport is likely to remain a large contributor to air pollution, especially in urban areas. For this reason, major efforts are being made for the reduction of polluting emissions from road transport. These include new powertrains and vehicle technology improvements, fuel refinements, optimization of urban traffic management and the implementation of tighter emission standards [4]. In recent decades, many emission models have been developed. Afotel \textit{et al.} [5] proposed regression models to estimate light-duty gasoline vehicle emissions of CO\textsubscript{2} based on vehicle velocity, acceleration, deceleration, power demand and time of the day. However, the model did not include NO\textsubscript{X} emissions. Oduro \textit{et al.} [6] proposed multiple regression models with instantaneous speed and acceleration as a predictor variables to estimate vehicular emissions of CO\textsubscript{2} but not NO\textsubscript{X}. Tóth-Nagy \textit{et al.} [7] proposed an artificial neural network-based model for predicting emissions of CO and NO\textsubscript{X} from heavy-duty diesel conventional and hybrid vehicles. The methodology sounds promising, but applied to heavy-duty vehicles only, and the fit function contains many details which make the model difficult to understand. Emission model based on instantaneous vehicle power, which is computed on total resistance force, vehicle mass, acceleration, velocity, and drive-line efficiency, was developed by Rakha \textit{et al.} [8]. However, the model applies for fuel consumption and CO\textsubscript{2} emission factor and does not include the NO\textsubscript{X} emission.

A key gap in our understanding of these emissions is the effect of changes in vehicle speed, power and load on average emission rates for the on-road vehicle fleet. Vehicle power, load and vehicle speed are closely linked to fuel consumption and pollutant emission rates [9]. Improved understanding of the link between operating conditions and emissions could develop accurate models for prediction of vehicle emissions. The quality of the application of any road vehicle emission model largely depends on the representativeness of the emission factor such as carbon dioxide (CO\textsubscript{2}), carbon monoxide (CO), nitrogen oxides (NO\textsubscript{X}), volatile organic compounds (VOCs) and particulate matter (PM). This refers to the accuracy with which the emission factor can describe the actual emission level of a particular vehicle type and driving conditions applied to it.

This work focuses on using the MARS methodology to
improve the prediction accuracy of chassis dynamometer and on-board measurement systems. The dynamometer testing is one of the three typical vehicle tailpipe emission measurements methods, where emissions from vehicles are measured under laboratory conditions during a driving cycle to simulate vehicle road operations [10]. The real-world on-board emissions measurement is widely recognized as a desirable approach for quantifying emissions from vehicles since data are collected under real-world conditions at any location travelled by the vehicle [11]. Variability in vehicle emissions as a result of changes in facility (roadway) characteristics, vehicle location, vehicle operation, driver, or other factors can be represented and analysed more reliably than with the other methods [12]. This is because measurements are obtained during real-world driving, eliminating the concern about non-representativeness that is often an issue with dynamometer testing, and at any location, eliminating the setting restrictions inherent in remote sensing. Though this measuring technique seems to be more promising, the need to improve the prediction accuracy of emission factor especially with NO\(_X\) emissions by using effective statistical techniques is important in any emission inventory.

A number of the models discussed above either do not estimate NO\(_X\) emissions, or are so sophisticated as to require excessive data inputs. There needs to be a balance between the accuracy and detail of a model for its ease of application. Therefore, to enhance the prediction performance for the NO\(_X\) emissions, the MARS modelling approach is proposed in this paper. This, we aim to estimate, with high accuracy, the NO\(_X\) emissions. The effectiveness of the model is then determined by dividing the data into two parts, one for building the model (learning) and the other for validating the model (testing). The results are verified by comparing the real data and the MARS predicted values.

2 Methodology

2.1 Chassis Dynamometer Data Collection

This study uses secondary data corrected by the New South Wales (NSW) Road and Maritime Service (RMS), Department of Vehicle Emission, Compliance Technology Operation. The data were collected on the second by second basis and four vehicles were used for the test. The test vehicles include Toyota, Ford, Holden and Nissan from 2007 and 2008 model year with an engine displacement ranging from 1.8L to 2.0L. A chassis dynamometer set-up in the laboratory simulates the resistive power imposed on the wheels of a vehicle, as shown in Figure 1. It consists of a dynamometer that is coupled to drive lines that are directly connected to the wheel hubs of the vehicle, or to a set of rollers upon which the vehicle is placed, and which can be adjusted to simulate driving resistance. During testing, the vehicle is tied down so that it remains stationary as a driver operates it according to a predetermined time-speed profile and gear change pattern shown on a monitor. A driver operates the vehicle to match the speed required at the different stages of the driving cycle [13]. Experienced drivers are able to closely match the established speed profile.

![Figure 1. Schematic representation of a chassis dynamometer testing.](image1)

2.2 On-Board Data Collection

Data from on-board instruments, can facilitate development of micro-scale emission models [10]. Compared with conventional dynamometer testing under carefully controlled conditions, on-road data reflects real driving situations. Accordingly, second-by-second emissions data were collected using a Horiba On-Board Measurement System (OBS-2000), as shown in Figure 2, with the same testing vehicles as with the dynamometer test cycle. The equipment is composed of two on-board gas analysers, a laptop computer equipped with data logger software, a power supply unit, a tailpipe attachment and other accessories. The OBS-2000 collects second-by-second measurements of nitrogen oxides NO\(_X\), hydrocarbons (HC), carbon monoxide (CO), carbon dioxide (CO\(_2\)), exhaust...
temperature, exhaust pressure, and vehicle position (via a global positioning system, or GPS). Although the instrument measured other pollutants, the focus of this work was to build a model for NO\textsubscript{X} emissions. For the measurement scale used, accuracy for the NO\textsubscript{X} emission measurements, reported in percentage, was ±0.3%. A second lag in NO\textsubscript{X} emission measurement was accounted for in the data spreadsheets. NO\textsubscript{X} sensor calibration was carried out throughout the data collection period. To ensure consistently smooth and good data collection without frequent interruptions due to any possible unit malfunction, inability of batteries to stay charged and calibration issues throughout the period, proper maintenance and diagnostic procedures were strictly followed.

3 Multivariate Adaptive Regression Splines (MARS) Model

MARS was introduced for fitting the relationship between a set of predictors and dependent variables [14]. MARS is a multivariate, piecewise regression technique that can be used to model complex relationship. The space of predictors is divided into multiple knots in order to fit a spline function between these knots ([14], [15]). The basic problem in vehicular emission modelling is how best to determine the fundamental relationship between dependent variables, and vector of predictors, such as speed, acceleration, load, power, ambient temperature including other factors.

The MARS algorithm searches over all possible univariate hinge locations and across interactions among all variables. It does so through the use of combinations of variable called basis functions. The approach is analogous to the use of splines. This study aims at exploring the potential of applying the MARS methodology to model NO\textsubscript{X} emissions using the following set of input parameters: speed, acceleration, load, power and ambient temperature of chassis dynamometer and on-board emission measurements. The problem can be stated as a multivariate regression problem. Suppose that N pairs of input-output parameters are available: \(\{y_i, x_{i1}, \cdots, x_{im}\}_1^N\), where the depend variable \(y_i, i = 1, 2 \cdots, N\), is the \(i\)th measure of NO\textsubscript{X} and the predictor \(x_{i1}, i = 1, 2 \cdots, N, l = 1, 2 \cdots, m\), is the \(l\)th measure of the \(i\)th parameter. We assume that the data \(\{y_i, x_{i1}, \cdots, x_{im}\}_1^N\) are related through the following equation

\[
y = f(x_1, \cdots, x_m), \ (x_1, \cdots, x_m) \in D \subset \mathbb{R}^m, \quad (1)
\]

where \(f(\cdot)\) is an unknown multivariate deterministic function and \(D\) is the domain of inputs. Since the true mapping in (1) is not known, it is desired to have a function \(\hat{f}(x_1, \cdots, x_m)\) that provides a "good" fit approximation of the output data. The good fit between \(\hat{f}(x_1, \cdots, x_m)\) and the output data is using the integrated mean square error (MSE) estimated.

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} [y_i - \hat{f}(x_{i1}, \cdots, x_{im})]^2. \quad (2)
\]

To regularize the problem, that is, make it well-posed, a restriction is imposed for the solution \(\hat{f}(x_1, \cdots, x_m)\) as functions residing in the linear space:

\[
F = f : f(x) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(x), \quad (3)
\]

where \(\{h_m(\cdot)\}_m=1^M\) is a set of basis functions and \(\{\beta_m\}_{m=0}^M\) are coefficients of representation. In this paper, \(h_m(\cdot)\) is the splines basis function defined as:

\[
h_m(x) = \prod_{k=1}^{K_m} [s_{k,m} \cdot (x_{v(k,m)} - l_{k,m})]_+, \quad (4)
\]

where \(s_{k,m}\) are variables that take values ±1, \(v(k,m)\) labels the predictor variables and \(l_{k,m}\) represents estimated values on the corresponding variables. The quantity \(K_m\) is the number of “splits” that give rise to each basis function \(\beta_m\). Here the subscript “+” indicates a value of zero for negative values of the argument. The basis functions involved in (1) are known as “hockey sticks” basis function. MARS searches over the space of all inputs and predictors values (knots) as well as interactions between variables. Now, given the estimated coefficients \(\{\beta_m\}_{m=0}^M\), basis functions \(\{h_m(\cdot)\}_{m=0}^M\) and operation parameters describing a new measurement, the emission of the new measurement can be predicted by taking the following steps:

1. Segregate operation parameters including speed, acceleration, power, load and ambient temperature from the raw data.
2. Predict the emission NO\textsubscript{X} by using the approximate function \(\hat{f}(\cdot)\) with \(\{\beta_m\}_{m=0}^M\) and \(\{h_m(\cdot)\}_{m=0}^M\), that is

\[
\hat{f}(x_1, \cdots, x_m) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(x_1, \cdots, x_m),
\]

\(i = 1 \cdots N\), where \(\{x_1, \cdots, x_m\}_1^N\) are from new measurements. The basis functions, together with the model parameters, are combined to produce the predictions given the inputs. The general MARS model equation is given as:

\[
\hat{f}(X) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(X), \quad (5)
\]

where \(\{\beta\}_m\) are the coefficients of the model that are estimated to yield the best fit to the data, \(M\) is the number of sub-regions or the number of basis functions in the
model, and \( h_m(X) \) is the spline basis function given in (4). This model searches over the space of all inputs and predictor values (referred to as “knots”) as well as the interactions between variables. During this search, an increasingly larger number of basis functions are added to the model to minimize a lack-of-fit criterion. As a result of these operations, MARS automatically determines the most important independent variables as well as the most significant interactions among them. From Put et al. [15], it is noted that the search for the best predictor and knot location is performed in an iterative process. The predictor as well as the knot location, having the most contribution to the model, are selected first. Also, at the end of each iteration, the introduction of an interaction is checked for possible model improvements.

3.1 Model selection and pruning

In general, non-parametric models are adaptive and can exhibit a high degree of flexibility that may ultimately result in over fitting, if no measures are taken to counteract it. The second step is the pruning step, where a “one-at-a-time” backward deletion procedure is applied in which the basis functions with the least contribution to the model are eliminated. This pruning is based on a generalized cross-validation (GCV) criterion. The GCV criterion is used to find the overall best model from a sequence of fitted models, where a larger GCV value tends to produce a smaller model, and vice versa. The GCV criterion is estimated using (7) as the lack-of-fit criterion [17]:

\[
\text{GCV} = \frac{1}{N} \sum_{i=1}^{N} \left( y_i - \hat{f}(X_i) \right)^2 \left[ 1 - \frac{C(M)}{N} \right]^2,
\]

where \( 1 - \frac{C(M)}{N} \) is a complexity function, and \( C(M) \) is defined as \( C(M) = C(M) + d.M \), of which \( C(M) \) is the number of parameters being fit and \( d \) represents a cost for each basis function optimization and is a smoothing parameter of the procedure. The higher the cost \( d \) is, the more basis functions will be eliminated [16].

4 Results and Discussions

Five vehicular emission predictor variables, namely, speed \((m/s)\), acceleration \((m/s^2)\), power \((W)\), temperature \((°C)\) and load \((Nm)\) were used with the response variable of NOX \((g/s)\) in an attempt to identify the relationships that vehicular emission models developers wish to understand. To explore factors affecting vehicular emission models, the present study provides results and some interpretations from the MARS model. Table 1 and 2 summarize the variable selection results using MARS, whose beta factor coefficients \( \beta_m \) are denoted \( BF_m \). In a MARS model, basis functions are used to predict the effects of independent variables on NOX emission factor. The interpretation of MARS results is similar to but not as straightforward as that of classical linear regression models. A positive sign for the estimated beta factors for the basis function indicates increased NOX emission, while a negative sign indicates the opposite. The value of beta factor implies the magnitude of effect of the basis function (i.e., variable effect) on the NOX emission.

Table 1. List of basis functions of the MARS and their coefficients for on-board measurements.

<table>
<thead>
<tr>
<th>Basis function</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF0</td>
<td>0.249827</td>
</tr>
<tr>
<td>BF1</td>
<td>-0.000142123</td>
</tr>
<tr>
<td>BF2</td>
<td>0.000342034</td>
</tr>
<tr>
<td>BF3</td>
<td>0.000442178</td>
</tr>
<tr>
<td>BF4</td>
<td>0.0032363</td>
</tr>
<tr>
<td>BF5</td>
<td>0.011587</td>
</tr>
<tr>
<td>BF6</td>
<td>0.0439038</td>
</tr>
<tr>
<td>BF7</td>
<td>-0.0013075</td>
</tr>
<tr>
<td>BF8</td>
<td>0.0073075</td>
</tr>
<tr>
<td>BF9</td>
<td>0.0113075</td>
</tr>
<tr>
<td>BF10</td>
<td>0.0311017</td>
</tr>
<tr>
<td>BF11</td>
<td>0.00023075</td>
</tr>
<tr>
<td>BF12</td>
<td>0.00313022</td>
</tr>
<tr>
<td>BF13</td>
<td>0.02113075</td>
</tr>
<tr>
<td>BF14</td>
<td>0.01561811</td>
</tr>
<tr>
<td>BF15</td>
<td>0.0179656</td>
</tr>
<tr>
<td>BF16</td>
<td>0.02113075</td>
</tr>
<tr>
<td>BF17</td>
<td>0.0148769</td>
</tr>
<tr>
<td>BF18</td>
<td>0.01567893</td>
</tr>
</tbody>
</table>

For the effect of each basis function, max \((0, x - t)\) is equal to \((x - t)\) when \( x \) is greater than \( t \); otherwise the basis function is equal to zero. As shown in Table 1 and 2, the MARS model contains 19 and 15 basis functions for on-board and dynamometer testing respectively. The on-board measurements and dynamometer testing have similar interpretations. It can be observed that all the five predictor variables play crucial roles in determining NOX vehicle emission. From Table 1, beta factors BF1, BF2, BF3, BF4, BF5 and BF6 account for the nonlinear effect of vehicle speed in the emission model. The effect of speed on NOX emissions can be explained as follows. By using the onboard measurements method, if the speed of the vehicle is less than 8.1127 \( m/s \) or 29.2 \( km/h \), it has negligible effect on the NOX emission (indicated by...
BF0), but from 11.667 m/s or 42 km/h this effect is increased with an increase in speed (indicated by BF2-BF5). The emission rate can reach 0.0439 g/s when the speed is about 24.1667 m/s or 82 km/h (indicated by BF6).

Table 2. List of basis functions of the MARS and their coefficients for dynamometer testing.

<table>
<thead>
<tr>
<th>Beta factor</th>
<th>Basis function</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF0</td>
<td></td>
<td>0.313578</td>
</tr>
<tr>
<td>BF1</td>
<td>Max(0, SPEED-6.428)</td>
<td>-0.00017255</td>
</tr>
<tr>
<td>BF2</td>
<td>Max(0, SPEED-9.356)</td>
<td>0.000625</td>
</tr>
<tr>
<td>BF3</td>
<td>Max(0, SPEED-18.368)</td>
<td>0.00575</td>
</tr>
<tr>
<td>BF4</td>
<td>Max(0, SPEED-25.136)</td>
<td>0.0635</td>
</tr>
<tr>
<td>BF5</td>
<td>Max(0, ACCEL-1.119)</td>
<td>0.00943</td>
</tr>
<tr>
<td>BF6</td>
<td>Max(0, ACCEL-4.235)</td>
<td>0.0567</td>
</tr>
<tr>
<td>BF7</td>
<td>Max(0, ACCEL-6.243)</td>
<td>0.0663</td>
</tr>
<tr>
<td>BF8</td>
<td>Max(0, AMBT -21.54)</td>
<td>0.000321</td>
</tr>
<tr>
<td>BF9</td>
<td>Max(0, AMBT -23.15)</td>
<td>0.00443</td>
</tr>
<tr>
<td>BF10</td>
<td>Max(0, AMBT -24.62)</td>
<td>0.0372</td>
</tr>
<tr>
<td>BF11</td>
<td>Max(0, LOAD -15.67)</td>
<td>0.0132</td>
</tr>
<tr>
<td>BF12</td>
<td>Max(0, LOAD -45.67)</td>
<td>0.053</td>
</tr>
<tr>
<td>BF13</td>
<td>Max(0, Power -13.76)</td>
<td>0.0168</td>
</tr>
<tr>
<td>BF14</td>
<td>Max(0, Power -20.64)</td>
<td>0.0212</td>
</tr>
</tbody>
</table>

This expected finding is consistent with previous findings in literature. From Carslaw et al. [18], it is noted that NO$_X$ emissions rise and fall in a reverse pattern to hydrocarbon emissions (HC). As the speed of the vehicle increases the mixture becomes leaner with more HC’s at high temperatures in the combustion chamber, there appear excess oxygen molecules which combine with the nitrogen to form NO$_X$. From Table 1, as the speed increases (indicated by BF2-BF6) the total NO$_X$ emission emitted from the tail pipe also increases.

Beta factors (BF7-BF10) on Table 1 show the nonlinear effect of vehicle acceleration on the NO$_X$ which can be described as fellows. If the vehicle acceleration is less than 0.95 m/s$^2$, NO$_X$ emission will reduce by 0.0013075 g/s (indicated by BF7), but if the acceleration is increased from 1.25 m/s$^2$, to 5.85 m/s$^2$, the NO$_X$ emission will increase by 0.0113075 g/s (indicated by BF8 and BF9). The NO$_X$ emission can reach more than 0.0311017 g/s when the acceleration exceeds 7.21 m/s$^2$. This result is similar to that of the speed because of depressing the accelerator pedal increase acceleration as well as speed simultaneously. The ambient temperature is also found to influence the NO$_X$ emission as indicated by BF11, BF12 and BF13 of Table 1, the effects of ambient temperature on NO$_X$ emission occurrence include: (1) if the ambient temperature is less than 22.12°C then it has no effect on vehicle NO$_X$ emission (indicated by BF11); (2) if the ambient temperature is greater than 22.12°C but less 23.47°C, NO$_X$ emission will increase by 0.00023075 g/s for 1°C increase of ambient temperature (indicated by BF11 and BF12); (3) if the ambient temperature is greater than 23.47°C but less than 24.76°C, the vehicle NO$_X$ emission will increase by 0.00313022 g/s for 1°C increase in ambient temperature (indicated by BF12 and BF13) and (4) if the ambient temperature is greater than 24.76°C the NO$_X$ emission will increase by 0.02113075 g/s for 1°C increase in ambient temperature (indicated by BF13). The higher ambient temperature resulting in more vehicle NO$_X$ emission is expected, because NO$_X$ is formed in a larger quantity in the cylinder as the combustion temperature exceeds the required limit. This finding is also consistent with previous explanation. In addition, temperatures greater than 24.76°C (B13) will significantly produce NO$_X$ emissions. As indicated by BF14, BF15 and BF16, the MARS results show the effect of load: (1) if the load is less than 10.53 N m, then it has no effect on NO$_X$ emission (indicated by BF14); (2) if
the load is greater than 10.53 Nm but less 52.34 Nm, the NO\textsubscript{X} emission will increase by 0.01561811 g/s for 1 Nm increase of load (indicated by BF14 and BF15); (3) if the load is greater than 52.34 Nm but less than 60.15 Nm, the vehicle NO\textsubscript{X} emission will increase by 0.0179656 g/s for 1 Nm increase in load (indicated by BF15 and BF16) and (4) if the load is greater than 60.15 Nm the NO\textsubscript{X} emission will increase by 0.02324571 g/s for 1 Nm increase in load (indicated by BF16). As far as the effect of power on NO\textsubscript{X} emission, BF17 and BF18 indicate that the occurrence can be described as: (1) if the power is less than 8.98 W, then it has no effect on vehicle NO\textsubscript{X} emission (indicated by BF17); (2) if the power is greater than 8.98 W but less 21.32 W, NO\textsubscript{X} emission will increase by 0.01567893 g/s for 1 W increase of power (indicated by BF17 and BF18); (3) if the power is greater than 21.23 W, the vehicle NO\textsubscript{X} emission will increase by 0.01567893 g/s for 1 W increase in power (indicated by BF18). The NO\textsubscript{X} emission as a result of the increasing load and power is expected, following the remark by Pierson \textit{et al.} [19] that driving a vehicle against a higher resistance will increase the engine load and power which will result in increases of the carbon dioxide (CO\textsubscript{2}) and NO\textsubscript{X} emissions.

To illustrate the NO\textsubscript{X} emission during real-world driving conditions and the dynamometer testing drive cycle, Figures 3 and 4 show the MARS model that has the best performance basis on independent test samples. There were 557 data points used in the analysis, 65% of which for building the model (Learn) and 35% for validation (Test).

### Table 3. Comparison of MARS and Multiple Linear Regression (MLR) model

<table>
<thead>
<tr>
<th>Model</th>
<th>Summary Statistics</th>
<th>On-Board</th>
<th>Dynamometer</th>
</tr>
</thead>
<tbody>
<tr>
<td>MARS</td>
<td>$R^2$</td>
<td>0.63</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Adjusted $R^2$</td>
<td>0.62</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>$3.35 \times 10^{-6}$</td>
<td>$1.33 \times 10^{-5}$</td>
</tr>
<tr>
<td>MLR</td>
<td>$R^2$</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>Adjusted $R^2$</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>MSE</td>
<td>$2.57 \times 10^{-5}$</td>
<td>$3.12 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

The on-board system model has nineteen basis functions with the best model with the least mean square error occurring at 17\textsuperscript{th} basis function, $R^2$ value of 63% while the chassis dynamometer has the $R^2$ of 57% with the best model occurring at BF12. Table 3 compares the MARS and Multiple Linear Regression (MLR) model and presents the model summary statistics. It is clear that the MARS model performs better than the MLR model as the latter gives $R^2$ of 51% and 50% for both the on-board and the dynamometer test. The 12% and 7% differences in contribution achieved by the MARS model confirms its ability in improving the prediction accuracy of the NO\textsubscript{X} emission. Among all the predictor variables the speed appears to have the highest contribution to NO\textsubscript{X} emissions. Figures 5 and 6 provide a detailed plot of the real data and prediction using MARS techniques. Note that the predicted emissions follow the real data with sufficiently good precision although there is a slight deviation in the dynamometer predictions. The MSE of the on-board system was $3.355 \times 10^{-6}$ while that of dynamometer was $1.33 \times 10^{-5}$.

![Figure 5](image5.png)

**Figure 5.** Predicted values of NO\textsubscript{X} and the real data plotted for on-board measurements.

![Figure 6](image6.png)

**Figure 6.** Predicted values of NO\textsubscript{X} and the real data plotted for dynamometer testing.
5 Conclusion

This paper has presented a MARS modelling approach to effectively estimate vehicular NOX emissions. The model approximates the nonlinear relationship between the NOX emission which is a function of speed, acceleration, temperature, power and load as predictor variables. The MARS model is implemented with 19 and 15 effective piecewise-linear BFs. The model predicts the NOX emission by forming a weighted sum of the predictor variables; thus, the predicted emission changes in a smooth and regular fashion with respect to the input variations, offering some performance improvements. The results obtained indicate a promising application of the proposed method in the estimation of NOX emissions with a reasonable accuracy. The proposed method may usefully assist in a decision-making policy regarding urban air pollution.

Acknowledgements

The data used for this study were supplied by the Road and Maritime Service, Department of vehicle emission, Compliance Technology & Compliance Operations, NSW Office of Environment & Heritage, and Horiba Australia. Assistance provided by Paul Walker and Thomas Mahrslie is gratefully acknowledged.

References


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

IT APPLICATIONS
Application of Building Information Modeling in Designing Fire Evacuation – a Case Study

Kun-Chi Wang\textsuperscript{a}, Shih-Yu Shih\textsuperscript{b}, Wen-Shuo Chan\textsuperscript{c}, Wei-Chih Wang\textsuperscript{d}, Shih-Hsu Wang\textsuperscript{e}, Abdoul-Aziz Gansonre\textsuperscript{b}, Jang-Jeng Liu\textsuperscript{c}, Ming-Tsung Lee\textsuperscript{c}, Yuan-Yuan Cheng\textsuperscript{f}, and Ming-Feng Yeh\textsuperscript{f}

\textsuperscript{a}PhD Student, Department of Civil Engineering, National Chiao Tung University, Taiwan \\
\textsuperscript{b}MS Student, Department of Civil Engineering, National Chiao Tung University, Taiwan \\
\textsuperscript{c}Engineering Assistant, National Synchrotron Radiation Research Center, Taiwan \\
\textsuperscript{d}Professor, Department of Civil Engineering, National Chiao Tung University, Taiwan \\
\textsuperscript{e}Assistant Professor, Department of Civil Engineering, ROC Military Academy, Taiwan \\
\textsuperscript{f}Assistant Engineer, National Synchrotron Research Center, Taiwan \\
E-mail: chi780118.cv00g@g2.nctu.edu.tw

Abstract -

In recent years, the construction industry has been actively seeking for possible applications of Building Information Modeling (BIM), including the domain of building disaster-prevention management. BIM enables us to present a facility layout in three-dimension (3D) and carry the disaster-prevention objects and information whereas the traditional approach can only be presented in two-dimension. Therefore, it could be possible to improve the traditional two-dimensional (2D) plane of building disaster-prevention management. By focusing on fire disaster, this study develops a BIM-based model to support disaster-prevention management. This model includes four modules: personnel safety evaluation, escape route planning, education and training, and equipment maintenance modules. The personnel safety evaluation module integrates BIM with a type of fire simulation software (Fire Dynamics Simulator), while the escape route planning module and safety education and training module help generate escape routes and emergency exits according to the 3D functions of BIM. The firefighting equipment maintenance module utilizes the information management ability provided by the BIM to demonstrate 3D location information and condition (maintenance records) of all fire equipment, to increase the quality of the safety education, and ensuring the all fire equipment are maintained in great condition. Above all, this proposed model is applied to a high-tech facility for testing its feasibility.

Keywords -

Building information modeling; disaster-prevention management; evacuation routes; disaster-prevention training

1 Introduction

In order to ensure the safety of personnel, the disaster prevention management of buildings is what the construction industry has been focused on. The tasks includes the design of escape routes in the building, escape routes simulation, disaster prevention briefs, educational training etc. However, the conventional escape routes were usually demonstrated in 2D views, and educational trainings often take places with 2D visual aids to explain hazardous areas and escape routes. Average people are not able to interpret the drawings in a short period of time, to understand their exact position within the building and to select the appropriate escape exit.

In recent years, the construction industry has been actively developing the Building Information Modeling (BIM) technology under all its possible applications, including the domain of building disaster management (Eastman et al., 2011). BIM enables us to present a facility layout in three-dimension (3D) and carry the disaster-prevention objects and information whereas the traditional approach can only be presented in two-dimension (2D). Therefore, it could be possible to improve the traditional two-dimensional (2D) plane of building disaster-prevention management.

This study uses fire disaster as an example to propose a disaster prevention management module based on BIM, and connect the module with FDS (Fire Dynamics Simulator, NIST, 2010). This way, the module can predict the time to evacuation, evaluate the safety of personnel, and assist on planning the escape routes using BIM. Finally, the module would present the escape routes in films to enhance the quality of educational training.
### 2 Archive Review

#### 2.1 Fire Emergency Exit Indication

Escape to safety can be defined as the emotional reaction of a personnel inside a building complex, trying to escape to safe zone via escape facilities or equipment within allowable time when a disaster happens (Lee, 2005). A British scholar, Marchant (1980), proposed that escape time can be classified into necessary escape time and allowable escape time. The necessary escape time should be shorter than what is allowable to ensure the safety of personnel. The main factors affecting the necessary escape time are escape facilities, the psychological spirit of those who escape, their physiology, behavioral change, and fire dynamic change, etc. To design escape exits, aisles, stairs, all these factors above should be carefully considered (Marchant, 1980).

According to the national standards, the recommended formula from the evaluation manual of building fire prevention and escape ability is used to estimate the escape time \( t_{escape} \) required for the personnel inside the building. The formula is as followed:

\[
t_{escape} = t_{start} + t_{travel} + t_{queue}
\]

where, \( t_{start} \) = evacuation starting time, \( t_{travel} \) = walking time, \( t_{queue} \) = time of blocked exit.

Also, \( t_{start} \) is calculated as \( \left( \frac{\sqrt{\Sigma A}}{30} \right) + 3 \); \( t_{travel} \) is calculated as \( \max \left( \frac{l_i}{v} \right) \); \( t_{queue} \) is calculated as \( \frac{\Sigma \rho A}{\Sigma N_{eff} B_{st}} \). Therefore, formula (1) can be modified as:

\[
t_{escape} = \left( \frac{\sqrt{\Sigma A}}{30} + 3 \right) + \max \left( \frac{l_i}{v} \right) + \frac{\Sigma \rho A}{\Sigma N_{eff} B_{st}}
\]

Where, \( \Sigma A \) = sum of area; \( l_i \) = distance of any point to the exit; \( v \) = speed of walking; \( \rho \) = personnel density; \( Neff \) = effective flow coefficients; \( B_{st} \) = width of exit.

In terms of allowable escape time \( t_{allowable} \), the most common hazards in a fire scene are low oxygen level, flame radiation, and smoke (Xiao and Yen, 2001). Therefore, allowable escape time can use the hazard level index from SFPE (Safe Fire Protective Engineering) manual in Table 1 as a guideline (National Fire Protection Association, 2008).

#### Table 1. definition of personnel hazard classification of SFPE manual

<table>
<thead>
<tr>
<th>Category</th>
<th>Limits (hazardous situation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat convection</td>
<td>Temperature &gt; 60 OC, risks of hazard</td>
</tr>
<tr>
<td>Toxicity</td>
<td>Concentration of CO &gt; 1400ppm, risks of hazard</td>
</tr>
<tr>
<td>Smoke covering</td>
<td>Visibility less than 2 meters, risks of hazard</td>
</tr>
</tbody>
</table>

Note: The limit of hazardous situation means people exposure to hazardous for 30 minutes.

Furthermore, according to the evaluation manual of building fire prevention and escape ability, evacuation can be difficult while smoke layer height down to 1.8 meters from the ground. Therefore, the measurement points about smoke covering rate in Table 1 should be set at 1.8 meters from the ground (Wang, 2007).

#### 2.2 BIM application in Disaster Prevention Management

By the information provided by BIM, users can make response when encountering disaster or emergency situation. For instance, the spatial information in BIM can help make judgment of the relationship between escape routes location and environmental hazards and thus reduce the uncertainty of emergency response (Zou and Wang, 2009; Rüppel and Schatz, 2011). Many national scholars classified the application of BIM on disaster prevention management into two categories: single building and multiple buildings (urban disaster). Some examples are as follows.

Chang (2011) combined BIM and online version of building fire surveillance system such as sensor, monitor, and spatial related information. When sensor is triggered, the management staff can identify the authenticity of the alarm through the system, find out the location of the fire so that the fire rescue would be effective and efficient in a large building.

Chen (2012) uses BIM for the maintenance work for fire equipment in a large building. Utilizing the information and history records of fire safety equipment in BIM and the 3D visualization for the layout, the records are updated in the model, organized automatically and printed out as spreadsheets to simplify all excess works.

Chung (2012) put all the conventional investigation records of fire causes into spreadsheets, connected with BIM, and created a fire cause analysis system. Tablets, smartphones, and other kinds of portable devices are
used to record the damage level and fire scene, and through wireless, can upload them to a database in order to help fire scene investigation crew to examine and analyze in the lab. This system not only increases the efficiency of process but also helps the foundation of future database and classification.

Lin et al. (2009) use Taipei MRT (Mass Rapid Transit) as study subject and build a model of the train station using BIM software and buildingEXODUS escape software. Using urban streets as subject, Huang et al. (2011) combines 3D visualized environment and time line to create a 4D dynamic simulation system to analyze the evacuation routes in the city. Considering human behaviors and environmental factors, the safety of the street area is evaluated by the crowdedness and by assigning a code of colors to the results, so that planners can based on the simulation results to select the appropriate evacuation routes.

3 Disaster-Prevention Management model based on BIM

The disaster prevention management model proposed by the present study focus on the evacuation safety of the indoor personnel and evacuation equipment performance to do an overall evaluation. This model includes four prevention module: personnel safety evaluation, escape route planning, education and training, and equipment maintenance modules. The model structure is shown in Fig. 1. The overall is explained below and the following subsections, respectively give us the details of the contents of each module:

1. The personnel safety evaluation module evaluates the allowable escape time and required escape time. By comparing the results of the two values, we can determine the adequacy of the time for the staff to escape.

2. The escape routes planning modules utilizes the 3D information provided by the BIM model to determine the validity of the escape distance from each room, then select the exits.

3. The safety education and training module uses the distribution maps of hazardous areas to suggest that the disaster management personnel strengthens the inspection of the area. Then, the personnel is called to make a film of evacuation with the 3D escape direction map in order to better understand the escape routes.

4. The fire equipment maintenance module provides equipment information to the maintenance staff to enable them to complete more quickly firefighting equipment testing and replacement.

3.1 Personnel safety evaluation module

The personnel safety evaluation module combines fire dynamic simulation software (FDS) (NIST, 2010; Chung, 2012; Shih, 2013), in order to help evaluating the safety of personnel, which consists in ensuring that the required escape time is within the allowable escape time. Based on the results of the FDS simulation software and according to the personnel hazard classification of SFPE manual (Table 1), this module calculates the allowable escape time. On the other hand, it uses the formula (2) to calculate the required escape time (ABRI, 2004).

This module combines BIM with FDS to calculate the allowable escape time as described below:

1. 3D model of BIM: BIM models must necessary contain components such as columns, beams, walls, slabs, windows, doors, etc. The model level of development is the LOD 200 (Eastman et al., 2011) which is a 3D building massing including area, volume, height, and other basic information. The BIM models are converted into DWG file format and imported to the FDS software.

2. FDS fire simulation: set various combustion parameters, including: combustion reactants, heat release rate, fire source location, visibility and temperature measurement points, the simulation time, ambient temperature, etc. Among them, the visibility of the measuring points set at 1.8 m height from the ground floor, to view whether the visibility reaches the evacuation barrier height making escape conditions more difficult.

3. Allowable escape time: the FDS simulation results are interpreted according to the personnel hazard classification of SFPE manual (Table 1). The ambient temperature (limited to 60 °C), carbon monoxide (limited to 1400ppm ) and the visibility (limited to 2 meters) are taken as indicators. Then, we respectively measure the time required to reach the limit indicator and choose the smallest value as the allowable escape time (Chung, 2003).

The required escape time is calculated as followed:

1. Spatial information parameters: retrieve the spatial parameters provided by the BIM model, including the total area of the room, the distance of any point within the room to the exit, the exit width etc.

2. Verification of the building fire prevention safety: the Formula (2) is the formula of the required escape time.

3. Required escape time: with the spatial information provided by the BIM, we can use equation (2) to calculate the required escape time.

Finally, we compare the allowable escape time and the required escape time, if is less than, then people should have sufficient time to escape the disaster site.
3.2 Escape route planning module

In accordance with the "Construction Technology Building Design and Construction Rules" (Ministry of the Interior, 2013), there is a maximum walking distance limit from any point within a room to the escape stairs. That is to say that the maximum distance for the personnel to reach to the evacuation exits depends on the different types of buildings.

The escape information provided by the BIM (such as: the name and level of each room, the nearby escape stairs and emergency exits) can help calculate the escape distance. In the BIM model, this module first selects the farthest room away from the stairs as the starting point, measure the distance to the stairs, determine whether the escape route is in accordance with the rules and plan the corresponding exits to each room within the building.

3.3 Safety education and training module

Safety education and training modules can be divided into two parts. The first part is about the safety training of the disaster prevention management staff. Building disaster prevention management staff must first sort out the areas that contain hazardous materials or areas likely to cause disasters. Then, through human judgment, they point out the areas with high disaster probabilities which are inserted into the BIM 3D model. The above procedure can help the management staff to ensure a good inspection and control of disasters.

The second part concerns the safety education and training of ordinary citizens. The complete planning of escape routes in 3D model are turned into 3D escape direction map in which the escape routes guidance are clearly shown. Finally, a film is made to improve the quality of the training and make citizens more aware on how to escape in case of a disaster.

3.4 Maintain fire equipment module

Except the fact that BIM models can provide disaster-related necessary evacuation and spatial information, it can also provide fire equipment maintenance information that can be effectively used by the maintenance personnel, such as: device name, type, last repair time, maintenance staff, exterior features, maintain records, equipment warranty, specifications, manuals, contracts regulations, equipment manufacturers, unit prices information and equipment components brands, rules, warranty manufacturers, location of fire-fighting equipment, equipment status and other information.

In the past, the basic information about these devices were often dispersed between the device manager, the device manufacturer and other parties, not integrated in a single platform, what could easily lead to data loss problems. Through all the information within a single BIM model, all the information will be passed to the fire-fighting equipment manager with the BIM model. When the fire-fighting equipment is damaged, fire equipment maintenance personnel can obtain such information in order to efficiently test the device and change if necessary.
3.5 Building the system

For the implementation of BIM models as a strong base for the management of disaster, during its establishment, the BIM model shall be added to specific parameters or components in order to carry out the evaluation of each module. The model operations are described below:

- First, establish the BIM model into the LOD 200, at this point, the model contains columns, beams, walls, slabs and other basic structural elements, but it also contains the volume of buildings, height, opening dimensions and other information.
- Meanwhile, while doing the BIM model, draw the route from the escape stairs to the farthest room, then define each space within the building.
- In addition, in the model, add a variety of disaster prevention equipment and the need for equipment information fields, increasing the device name, type, time of last maintenance, maintenance staff, exterior features, maintain records, locations, specifications and other important information.

After completion of the above system, you can use the BIM model to the modules. Through the interaction of the BIM model and the information, we can assist the analysis of the four modules.

4 Case study

The present research case is about a high-tech factory facility in the northern part of Taiwan. The following sections describe the facility configuration status of the factory concerned and the application process of the proposed model in this research. Finally, through interviews with experts, we can evaluate the advantages and disadvantages of this model.

4.1 Case facility configuration

The circular building (refer to the study case building) of this public high-tech factory is shown in Fig. 2. The building facility configuration can be divided into inner ring (A zone), the shield tunnels (B zone) and outer ring (C zone). The inner diameter and the outer diameter are approximately 65 meters and 110 meters while shielding tunnel is located between the inner and outer rings. The BIM model of the case study building has already been built during the construction phase, the building software is Autodesk Revit (2013).
4.2 Personnel safety evaluation

First, this research will use the personnel safety evaluation module to facilitate the calculation of allowable escape time. In fact, the case building of the BIM model combined with FDS fire simulation software will simulate the circumstances of a fire in the building. To set the appropriate parameters in the FDS, it is important to first explore the building areas that are more likely to catch fire and the flammable items. It is also important to consider the most difficult evacuation areas of the building in case of fire. The following three points illustrate the ways of calculating allowable escape time:

1. BIM 3D model: convert the BIM model building into DWG format, then import into the FDS to transform it into FDS model, as shown in Fig. 5.
2. FDS fire simulation: during the simulation, the settings of each parameter begin with a more conservative value. In case of possible fire, the building has 58 engine rooms on the second floor: ventilation room, substations, computer control room and other engine rooms. In these Engine rooms, the connectivity between high voltage cables and electrical equipment, on a long term, is likely to lead to current overload caused by wires. So, after discussing with building fire prevention personnel, we reach to the conclusion that the engine rooms’ area of the second floor have the higher probability of a disaster. Also, in this case, we consider that the high-voltage transformer substation is the fire ignition place and that the cables can be taken as fire items.

The visibility and temperature measurement points are set at the entrance of the escape stairs in order to estimate the personnel escape time.

We also use the double of the general electrical fires maximum heat release rate (290KW) which is equivalent to a safety factor of 2; the simulation time is 10 minutes and the ambient temperature is normal (25 °C). And in the simulation, each space is set to be opened (no doors and windows) and fire compartments are not considered, the objective being to realize a fire simulation under conservative situations.

3. Allowable escape time: According to the results of the FDS simulation and corresponding to SFPE indicator level of personnel exposed to hazards (Table 1), we respectively identify the time required for temperatures to reach 60 °C, the time for the carbon monoxide concentration to be higher than 1400ppm and the time required for a visibility less than for 2 meters. We then proceed to a comparison of the three values and we consider the minimum as the allowable escape time.

Next, we must calculate the required escape time; the following descriptions will show the ways to calculate the time required for the evacuation:

1. Spatial information parameters: to calculate the required escape time (t_escape), we must provide the total area of the room, the distance between any point within the room to the exit, walking speed, occupant density, effective flow coefficient, the exit width and other information. Among them, the total area of the room, the distance between any point within the room to the exit, the exit width can be achieved in the BIM model, According to the BIM model of the case building, the total area of the room (ΣA) is 372 (M²); maximum
walking distance (li) is 46 (M); exit width (Bst) is 1.64 (M). For the other parameters, we use the suggestions of the Architecture and Building Research Institute (Ho and Chien, 2008) including: walking speed (v) is 72 (M / min), personnel density (p) 0.07 (persons / M²), the effective flow coefficient (Neff) of 90 (persons / min / M).

2. Verification of the fire safety code of a building: in this case we use the formula (2) to calculate the time required for personnel to escape.

3. Calculate the required escape time: referring to the previous parameters collected, substituting into equation (2) to calculate.

Figure 5. Schematic of the Revit model imported into FDS

4.3 Planning escape routes

The regulations specify only escape route walking distance. According to the Building Technical Regulations, article 93 of the rules of Building Design and Construction, the case building belongs to the industrial sector with a storage classified in the C type, the distance between the farthest room to the stairs must not be more than 70 meters (Ministry of the Interior, 2013). During the design stage of this case, the architect used manual calculations to check whether it meets the requirements (shih, 2013). In this study, we use the BIM model to determine the distance of the farthest room from the stairs and measure the distance between each stair (exits) and the farthest room. The objective is to test the distance between each space of the building and the exit in order to confirm the previous results of the manual calculations.

4.4 Safety education and training

In the safety education and training, the first part is about the training of the building prevention management personnel. The second part concerns the training of the working personnel.

1. Training of the disaster prevention management personnel: the case building contains many engine rooms. In order to help identify the engine rooms that are likely to lead to disaster, they should be marked out in the BIM model with characters. Then, with the help of the disaster prevention management personnel, we determine the hazards of each room according to its characteristics and identify the most hazardous room. On the second floor of the present building, there are all the different types of engine rooms. Many of these engine rooms contain multiple wire and combustible materials. Among them, the high pressure room is the most hazardous; so, we mark it in red in the BIM model in order to alert the maintenance staff about the importance to strengthen the inspection in this area.

2. Training of the working personnel: Referring to the escape routes planning, use the BIM model to create a 3D escape direction map including the guidance. Because the case building is a ring structure, the interior is very similar and it can easily cause confusion for the personnel. Therefore, in order to enable the personnel to quickly understand their location within the building and follow the correct escape routes in case of emergency, the column-numbers are clearly marked on the map (Fig. 6).

4.5 Fire-fighting Equipment Maintenance

In the case building, there are many people working and there is no doubt that their safety is very important. Therefore, in order to ensure the safety of the personnel, the indoor fire-fighting equipment must be maintained in good condition. The fire-fighting equipment used in the present case building are: fire extinguishers, fire hydrants, fire detectors, smoke extraction equipment, emergency lights, etc. Through the BIM model components, we can have all the information relative to the equipment maintenance to help conducting the fire-fighting equipment maintenance work.

4.6 Discussions

After the testing of the present case building in the model, the results are presented to the disaster prevention management staff for evaluation and their correlated suggestions are listed below:
1. The personnel safety evaluation module:
   a. In the past, the 3D model was required to be set up before the FDS simulation, including structural components such as columns, beam, wall and slab. Besides, the procedure of establishing the model was complex, time-consuming and low accurate. The parameter of objects (length, wide, and high) always needed to be entered item by item. Through BIM model (especially if the BIM model has been included early in the construction phase), it is possible to save a lot of time during the modeling.
   b. Due to the reliability of the space and size information provided by the BIM model, the results from the FDS simulation are more qualified for the actual situation.

2. Escape route planning module:
   a. When planning escape routes, it is very important to take into consideration a variety of information, such as the evacuation floor, the exit location, the nearby escape stairs etc. All these different kinds of information don’t have the same origin. Through the data access of the unique BIM model, we can significantly reduce the time required to collect data and increase the accuracy of the information.
   b. The BIM can help measuring distance between the farthest room and the escape stairs and can provide a 3D visualization of the escape routes in order to avoid the design of impractical escape routes.

3. Safety education and training module:
   By identifying the hazardous areas in the BIM model, the safety personnel can be aware about the necessity to strengthen the inspection of the high fire risks areas. (but, currently it can only be displayed in 2D; the 3D features could be an interesting option to be developed in the future).

4. Fire-fighting equipment maintenance module:
   Using the parameterization function of the BIM model, the information and records of the disaster prevention equipment maintenance are collected in a unique model to assist the maintenance staff in the testing of the relative equipment. It can reduces the information collection time and ensure that each fire-fighting equipment is in the most favorable condition.

5 Conclusions and Recommendations

This study uses the BIM 3D architectural space configurations and the possibility to mount the information about disaster prevention in the model, to propose a significant improvement of the traditional use of 2D documents on the disaster prevention management. The study has been done through four main modules that are the personnel safety evaluation, the escape routes planning, the safety education and training, and the fire-fighting equipment maintenance. Through the analysis of the present case building, the proposed model should be feasible. The personnel safety evaluation combines the BIM and FDS to provide the allowable escape time in case of emergency and uses the BIM parameters to estimate the required escape time for the personnel, which should be a great academic innovation.

Following are the subsequent research directions.

1. In the domain of the personnel safety evaluation, the fire simulation considers only the geometry of the interior space, it did not consider wall material.

2. Concerning the escape routes planning, the BIM model only consider the restrictions of the Construction Technology Building Design and Construction Rules, not taking into account other regulatory norms, nor personnel psychological characteristics.

3. The current operation mode of the model has not been automated yet. So, establishing an automated system should enhance the effectiveness of its use.
6 Acknowledgements

We would like to highlight the National Synchrotron Radiation Research Center (Project No.: SRPO-1020091) and the National Science Council (Project No.: NSC101-2622-E-009-013-CC3) who partly funded the present study. We thank the Concord Tech Company for their support on the design technology, Autodesk that provided the BIM design software and the National Chiao Tung University research team for their help in the Building design work. We are also indebted to all the engineering experts who participated in one way or another to provide valuable information and practical experience.

References

General Information Model and Its Application for Rock Excavation in Underground Work Sites

Esa Viljamaa*, Irina Peltomaa† & Annemari Kaaranka‡

* VTT Technical Research Centre of Finland, P.O. Box 1100 90571 Oulu Finland
† VTT Technical Research Centre of Finland, P.O. Box 3, 92101 Raahe Finland
‡ University of Oulu, Department of Mechanical Engineering, Erkki Koiso-Kantilankatu 3, FI-90570 Oulu Finland

E-mail: esa.viljamaa@vtt.fi, irina.peltomaa@vtt.fi, annemari.kaaranka@oulu.fi

Abstract

In the underground mining and tunnelling rock excavation processes, there is plethora of different machine, tool and information system providers involved and practically no general information model covering the complete process. The competition in the field is getting more intense and all the actors should intensify their operations. The information management and integration of different information systems’ data is among the most interesting enablers for such a better operational efficiency. The main goal of the research was to develop flexible methods for better application interoperability and information management for the underground rock excavation processes. The research used a literature survey, a commercial systems survey, professional interviews and worksite observations in order to develop a new general information model for rock excavation and define a toolchain for the exploitation of the developed model. In this research, the information model for underground rock excavation actions was developed applying applicable parts of existing standards and software tools. The main information model can be extended by using separately designed sub-models. In addition to information model also a toolchain of how to practically do the model development and information integration, has been described. These actions will enhance information management and application interoperability in the rock excavation processes and enable more efficient information exploitation leading potentially to better controlled processes, more efficient work phases and cost savings.

Keywords - Information Model, Information Management, Data Integration, IFC, IREDES, Rock Excavation

1 Introduction

Like in many other industries, the underground mining and tunnelling operations are facing more global and more severe competition than earlier. This means that all the operations must be intensified and made more cost-effective to meet new challenges. Also the common demand for less polluting and safer processes is pressing system and process developers and managers globally.

In order to intensify processes, more accurate and more extensive information about the processes is needed. The days when a company could thrive with stand-alone systems for different enterprise functions are over. Enterprise systems focused on different functions in a company may never become one, but they must share information to assure success. [1] In the underground mining and tunnelling operations, there may be tens or hundreds of different sources of information that are potentially useful from the process control and management point of view. This information should be brought to the disposal of the process management either as an individual data sources or rather as integrated information. Better methods for converting the raw data into integrated and useful information for both a product and a process control through the complete value chain are essential for development of mines [2]. There are few commercial solutions in the market to integrate information in the tunnelling operations, but those solutions remain to be vendor specific, closed and covering only some narrow sub-fields of needed total view for information. Information management solution must ensure that the right people and the right processes have the right information and the right resources at the right time [1].

In this research, the general information flow and information system integration were improved by means of the developed application interoperability concept.
consisting of a new information model and a defined toolchain for information integration concerning the model. The model was based on industry standard from the building construction field that was combined with de-facto rock excavation communication method along with other necessary parts. The model is mainly targeting the improved process management and decision making support among other benefits. This paper discusses mainly how the model was developed and its applicability will be tested later on in a real-life in real environment with a developed pilot system. The results of the research are potentially applicable within advanced mining and tunnelling worksites with proper process data collection systems and production control software. A practical functionality of results of the research will be validated in subsequent comparative research phase.

2 Methods

The information model and the toolchain developed in the research were based on information collected from a literature survey, worksite visits and personnel interviews involved to work in tunnelling and underground mining operations.

2.1 State-of-the-Art

For the underground mining and tunnelling operations there are several commercial solutions, however these solutions are concentrated on certain functions or are tightly vendor specific. Interoperable and overall information management solution for whole chain of underground mining and tunnelling operations doesn’t exist.

Commercial solutions for underground mine planning include Dassault Systèmes GéOVIÀ’s Surpac [3], CAE’s CAE Studio 5D Planner [4], and Mintec Inc.’s MineSight [5]. Commercial solutions for tunnel construction site planning include Trimble’s Business Center HCE [6], Tekla’s Tekla Civil [7], and Vianova Systems’s Novapoint [8]. Integration and accessing to planning information is usually possible through common file formats.

In academia, the application interoperability research in underground mining and tunnelling fields is not executed widely. However the need for improved information integration and management in mines is noted in the research [2][9]. Compared to underground mining operations the interoperability research is more advanced in the field of AEC/FM (architecture, engineering, construction and facilities management). [10] have reviewed wireless web services enabling technologies, [11] have made an extensive review from data and application integration research in construction domain, and [12] have conducted a comprehensive review on system integration technologies in AEC/FM.

In the following, few information management studies from AEC/FM field are introduced.

[13] present domain taxonomy for highway construction. The taxonomy is based on Foundation Classes (IFC) and several other classification systems. It uses seven major domains to classify construction concepts: Process, Product, Project, Actor, Resource, Technical Topics, and Systems. The taxonomy was used to develop an ontology for the construction domain including semantic relationships and axioms. The major ontological model is process-oriented and can be summarized as follows: Construction knowledge is encapsulated in several overlapping Systems, where a set of Actors use a set of Resources to produce a set of Products following certain Processes that are part of a Project according to boundary conditions and within the confines of the work environment (Technical Topics). [13] The operation of the developed domain ontology was evaluated during a e-COGNOS project [14] as a part of a web based knowledge management software, which connected various systems using a Web Service technology. The development of the ontology architecture was continued by adding more knowledge levels (application knowledge, user knowledge) to domain ontology [15].

[16] present more general level domain ontology called Infrastructure and Construction PROcess Ontology (IC-PRO-Onto) for infrastructure and construction domain. IC-PRO-Onto captures the most fundamental concepts in the domain in a structured, extendable, and flexible format. It uses five concepts to represent things: Entity, Constraint, Attribute, Modality, and Family. An Entity is a Project, Action, Actor, Product, Resource, or Mechanisms. Concepts were structured into hierarchal taxonomies using subsumption and aggregation relationships. The taxonomical classification of concepts supports multi-perspective viewing of the same process through Process Modality Views including Core Processes, Management Processes, Knowledge Integration Processes, and Support Processes. [16]

[17] present AEC-domain specific ontology merger Onto-Integrator. Existing ontology mergers could not fully address requirements set by AEC domain, which is multidisciplinary and has several stakeholders. Onto-Integrator uses semantic similarity comparison and relational concept analysis (RCA) for merging ontology taxonomies and relations, and knowledge base integration for axiom merging. The merging approach uses a heuristic. [17]

[18] have developed an IFC based ontology, which purpose is information retrieval from an IFC model. The
ontology source is IFC2x3 version and the ontology format is Web Ontology Language (OWL) 2 DL. The ontology consists of two parts Basic and Extended Ontology. The Basic Ontology includes those ontology components that are derived from the IFC specification directly. Extended Ontology is the ontology components that are not originally included in the IFC specifications but are added according to the requirement of the system. These components facilitate the information retrieval when the inputted query includes terminologies that are not readily available in the IFC specifications. [18]

[19] introduce semantic web services based approach to enable interoperability between Computer-Aided Design (CAD) and Geographical Information System (GIS). The approach consists of three modules: Task Interpretation, Web Service Matching, and Web Service Composition. The implementation of these modules requires development in for example algorithms and Quality of Service parameters. [19]

[20] propose a method for intensifying road construction process management. Method utilizes recent advances in communication and machine control systems, and the development of information management technologies. The prototype implementation called Dynamic Site Control Center (DSCC) enhances the information acquisition during road construction process. DSCC integrates information from different companies’ information systems participating to road construction process. In ontology-based integration process DSCC combines and refines the information gained and visualizes it for users. DSCC ontology forms from sub-ontologies, corresponding sub-processes of infrastructure building process. DSCC improves the process management’s situation awareness and provides platform independent access and visualization tool for process data. [21][22][20]

In the following two central data exchange formats are introduced more closely as they are central background formats in our information model development.

2.1.1 IFC

Industry Foundation Classes (IFC) is a schema developed to define an extensible set of consistent data representations of building information for exchange between AEC software applications. It is a neutral exchange format, which was designed as an extensible “framework model”. That is, its developers intended it to provide broad, general definitions of objects and data from which more detailed and task-specific models supporting particular exchanges could be defined. It covers all building information supporting full lifecycle from planning and design to maintenance. [23] However, [24] state that the formal standards on Building Information Modelling (BIM), such as the IFCs are complex and have not had the resources for rapid development and promotion that their potential deserved. In addition IFC has been weakened by varied non-consistent implementations [23]. Therefore it may take some time for this approach to be widely adopted [25].

Nevertheless IFC has become an international standard for data exchange and integration within the building construction industries. While IFC is able to represent a wide range of building design, engineering, and production information, the range of possible information to be exchanged in the AEC industry is huge. The IFC coverage increases with every release and addresses limitations, in response to user and developer needs. The latest version IFC 4 has about 800 entities (data objects), 358 property sets, and 121 data types. While these numbers indicate the complexity of IFC, they also reflect the semantic richness of building information, addressing multiple different systems, reflecting the needs of different applications, ranging from energy analysis and cost estimation to material tracking and scheduling. [23]

From system architecture perspective the IFC consists of four conceptual layers: Resource, Core, Interoperability, and Domain Layers. [26] The bottom layers define base reusable constructs which are generic for all types of products. The base entities are then composed to define commonly used objects in AEC, termed Shared Objects in interoperability layer. At the top level are the domain-specific extensions, which deal with different specific entities needed for a particular use. [23].

All IFC models provide a common general building spatial structure for the layout and accessing of building elements. It organizes all object information into the hierarchy of Project-Site-Building-Storey-Space. Each higher-level spatial structure is an aggregation of lower-level ones, plus any elements that span the lower-level classes. All application-defined objects, when translated to an IFC model, are composed of the relevant object type and associated geometry, relations, and properties. In addition to objects that make up a building, IFC also includes process objects for representing the activities used to construct a building, analysis geometry that is often abstracted from the building geometry, and analysis input and result properties. [23]

Model views are another level of specification, above the IFC schema. Model View Definitions (MVDs) identify what should be expected for an exchange to be effective. MVDs specify exactly what information should be exchanged, and in what form and structure the IFC entities are to be used. Defining MVDs requires principle decisions and workarounds because the IFC
itself does not address a number of semantic issues comprehensively. Two sets of semantics are at the core of any successful model exchange. One of which is the user or application functional semantics defining the information that must be exchanged and the other being the representational semantics available in IFC or other data modelling schema representing the user intentions. [27]

2.1.2 IREDES

The need for effective exchange of process information was driven force for standardization initiative called International Rock Excavation Data Exchange Standard (IREDES) [28]. The objective of IREDES is to define a "common electronic language" for easy and standardized data exchange between mining machines and corporate IT systems. [29] IREDES is based on eXtensible Markup Language (XML) technology and uses XML schema to define the formats of the machine generated files. Thus, the produced data report is understandable for both humans and machines. [30] XML structure supports object oriented software design and data encapsulation. Furthermore all major databases provide standard XML import and export features. [29]

IREDES architecture consists of three different levels: IREDES Base, Application Profile, and Equipment Profile. The IREDES Base level covers all data which objects in other levels require. The Application Profile level contains all information specific for one application purpose (e.g. Planning Data, or Production Quality Log). This is information which may be used independently from a specific type of equipment. The Equipment Profile level adds detailed information specific to each Equipment Profile applicable for the specific equipment type. The first available Equipment Profiles cover the Drill Rigs and LHD’s/Trucks as transport. [29]

[30] presents the idea of IREDES On-line which purpose is to setup an industry standard to enable volatile data to be transferred via network reliably such that the machine current status is visible to authenticated users. Volatile data includes real time data and current status information as these types of data vary rapidly over time. IREDES On-line suggests two alternative ways for communication: direct or infrastructure communication. In direct or peer-to-peer communication end user connects to machine using IP address, and machine acts as a server responding to client requests. In infrastructure communication the communication between machine and end user is done via server. This is more efficient method as machine’s calculation resources are limited and software is purpose specific. In infrastructure communication there is possibility to send more complex requests to server, for example ask for history information of the machine. [30]

2.2 Personnel Surveys and Observations

Part of the information needed for the developed information model and its application were gathered from personnel surveys from two different Finnish worksites. Other worksite was a tunnel construction site from a Helsinki subway expansion project and other was from an underground hard rock mining site. Although a final product of those separate work sites are different (ore and a subway tunnel), the production processes in the rock excavation phases are quite similar. Due to the worksite similarity, the replies to inquiry were processed as they were all from a single site. Part of the information was collected by interviewing personnel from different levels of responsibility from production planning staff and work supervisors to the machine operators as well as maintenance and measurements personnel. Rest of the individual data was collected by paper inquiries.

Questions in the interviews and the inquiries were about the current good and bad practises, bottlenecks and improvement suggestions related to information flow in general and also special ones related to assignment of the specific staff member in the question. Most of the interviews were done while a current interviewee was actually doing his duty making it also possible to do actual worksite observation and notes based on communication between different actors in the process.

The technology level of those two work sites involved varied between modern production control systems and all-covering production Wi-Fi to the paper based reporting method and analogue walkie-talkie communication system. There were 23 persons interviewed and 76 replies to inquiry.

Table 1 depicts a job type share of employees that answered the inquiry.

<table>
<thead>
<tr>
<th>Job Type</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Production planning system</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>Production system engineers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurements staff</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Work supervision</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Maintenance and equipping staff</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Machine operators etc.</td>
<td>28</td>
<td>37</td>
</tr>
<tr>
<td>3rd party contract supervision</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Safety personnel</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>76</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Table 1. Share of different job types in the inquiry.
3 Results

The resulting model and the data integration toolchain were based on the literature survey of the academia and the commercial systems as well as professional interviews and inquiries.

3.1 Information Model

Based on the literature survey and the professional interviews and inquiries, there were multiple requirements gathered related to developed information model. The requirements included topics from model general requirements to single data fragments from process to be included to the model structure.

First, there were multiple requirement suggestions received from interviewees related to detailed information fragments of different sub-topics in underground tunnelling process. Depending the position of an interviewee there were many different perspectives brought up. Different work planning and process follow-up terms like project, task and assignment with related performance parameters like cost, time, a sequence and such were considered important for the model. Also sub-process and machine specific topics like drilling, blasting and hauling machine use, process and maintenance related parameters were supposed to be included to the model.

The system designers and State-of-the-Art research pointed out quite obvious requirements for the model; the model should be modular and extensible in order to guarantee its updateability. Also it was supposed to be semantic in nature for application design that can distribute intelligence of the system also to the model side.

The core of the developed information model consists of selected components from IFC4 and IREDES definitions. Since the main goal of the research was to develop tools for an improved information flow to the tunnelling operations, only applicable parts were selected. From the IFC4 definition, Core schemas with Resource schemas were applied with Construction Management Domain schema. These schemas were then extended using the tunnelling process specific information gathered from the IREDES definition, professional interviews and the SoA research. The extension was named as TunnelBuildingDomain. Since IREDES is mainly an information exchange format that may be getting more common in the future, the activity definition in the developed information model is mapped to IREDES working sequence for later compatibility. The developed model does not cover the physical tunnel model, but physical models can be added to the main model using mechanisms provided by IFC relationship connectors. The model structure and the basis are depicted in Figure 1.

![Figure 1. The basis and the structure of the developed model.](image)

![Figure 2. Snippet from the model class hierarchy.](image)
The information model and its components were composed to a class hierarchy following the example of the original IFC4 model. An example snippet from the model can be seen in Error! Reference source not found., where TunnelBuildingDomain sub-classes are inherited from the ifcResource class. The model was modelled following the principles of Unified Modelling Language (UML) Domain Model extension. This model type is so called Domain Model that can be later on grounded to a physical model e.g. Structured Query Language (SQL) database scripts etc.

3.2 The Model Application

A prerequisite for the application interoperability is usage of a common information model. In addition, there are also multiple other actions that may be done to enable successful multi-actor data share. This chapter explains briefly how the model can be developed and deployed in practice.

First, the main information model must be modelled using some appropriate enterprise architect software tool. It may be done using UML Domain Model tool that also supports different relationships between entities. The domain model may then be derived to a physical model that includes database tables and their relations.

Secondly, an essential data may be then brought to database tables. E.g. any ETL (Extract, Translate, Load) tool may be used in order to upload and convert data from data sources to the database and to the model. The important data is then easily accessed from the database over enterprise service bus (ESB) solution or other communication solution. Figure 3 shows an explanatory development and application of the presented information model.

4 Discussion

In this research, the application interoperability of underground rock excavation processes intensification was enabled by developing a common information model and by defining a practical way to exploit the resulted information model.

The research developed an information model and its application description that were based on literature survey, professional inquiry and interviews and worksite observation periods.

The developed model was based in IFC4 and IREDES standard definitions. Currently there is no holistic standard information model available for underground rock excavation processes, but IFC4 have been used as a basis in few other information model researches [13]. However, in this research, the developed model was realized without utilization of semantic web standard languages like Resource Description Framework (RDF) or OWL. In addition to the model development, also the short description how the model could be applied was presented.

Figure 3. Model development and application practises.
There are several benefits while using the developed model; it is derived from standards or de-facto technologies meaning that a part of it has been reviewed in long run by experts from the different fields. IFC is still developing and maybe in some day there will also be an extension of a tunnelling physical model on it.

The developed model has been realized using UML which enables that the domain model can be read in and updated in multiple software. Also it is easy to derive a physical model automatically out of the domain model, in order to quicken the deployment. Due to the nature of IFC4, the model should be very easily extended and added to upper models.

Of course, there are also disadvantages in the developed model. Information models are always difficult to get standardized. The developed model is not modelled using “pure” semantic technologies meaning that some of the advantages of semantic web cannot be used. E.g. the model cannot be automatically aligned or mapped to other models using those highly advanced semantic linking technologies. Also it should be noted that the inquiry was only concerning two Finnish companies. There may be different requirements for models used in other work cultures with different technology readiness levels and practises.

There are many potential steps for further research work in the field of the application interoperability of underground rock excavation. First, the model and its applicability in practise should be tested in comparative tests in real-life environment including model realization and actual use as a basis for the application interoperability. In addition to the applicability, also its potential financial benefits should be researched. Also the information integration features of the model should be increased by modelling the wholeness with real semantic methods in order to enable automated mapping or aligning of other models to it. Also, the possibility to add real IFC-like tunnel physical model to main model should be developed.

5 Acknowledgements

This research was conducted in a research project called “Flexible Optimization of Underground Mine Production” that was mainly funded by Finnish Funding Agency for Innovation, TEKES. Also the contribution of co-funding companies is highly appreciated.

References


Performance and Impacts of Web-based Project Management Systems in Construction Projects

H. Doloia

a Faculty of Architecture, Building and Planning, The University of Melbourne, Australia
E-mail: hdoloi@unimelb.edu.au

Abstract

While web-based project management systems are widely being implemented in most construction projects, there is a clear lack of understanding among the industry professional in relation to any quantified benefit and achieving success in projects. The aim of this research is to unfold the understanding of the key factors affecting the implementation decisions and aided benefits of such web-based systems in projects empirically. Adopting the partial least square structural equation modelling (PLS-SEM) approach, the relational links between the six latent factors associated with the functionalities of web-bases PM systems have been assessed. The effect of six latent variables on the performance of web-based project management systems has been analysed using the data from a questionnaire survey of 77 respondents comprising the stakeholders within construction industry. The results of the data analysis suggest that project complexity (PC) and information streamline (IS) are the two key factors supporting the implementation of web-based systems (PWB) in projects. The general perception of increased users’ satisfaction (US) and transparency and accountability (TA) has no strong empirical basis for rationalising the use of web-based systems across projects. Effective monitoring and control (EMC) during project delivery can only be marginally enhanced by implementing web-based project management systems. With a clear understanding of the significance of these factors in the context of implementation of web-based project management systems, these findings could potentially contribute to the development of company’s procedures or to enhance existing knowledge within the construction industry.

Keywords: IT application, web-based systems, project management, project performance, structural model

1 Introduction

Increased complexity of modern construction projects with involvement of scattered stakeholders present enormous challenges for achieving seamless collaboration within the organisations. While the implementation of web-based project management systems has been significant over past years, the current literature lacks any clear consensus on rationalizing such decisions across the industry. The research proposes a PLS-SEM based model for investigating the impacts of the key factors on the performance of the web-based systems being implemented in construction projects. The resulting structural model is used to test ten underlying hypotheses as postulated in the study.

The basic rationale behind the implementation of web-based PM systems is communication and information handling in most projects. Increased practices of concurrent design, frequent change of scopes, budget constraints, design novation etc. demand centralized access with reliable means of transmitting and storing the project information. As such, the web-based centralized system offers a level of access to project information that supersedes the traditional means of communication using telephones, faxes, hard copy ring binders or emails (O’Brien, 2000). In the advent of technological advancements and rapid launches of competitive products in the marketplace, the idea that web-based project management is the only way to achieve success in modern projects among the industry professionals remains obscure. In fact, research shows that hastily introduction of web-based project management systems across organizations are showing the sign of dissatisfactions and reduced productivity in projects (Alshawi and Ingirige, 2003).

2 Literature Review

The implementation of web-based PM systems generally helps to achieve better communication practices, speed of construction and greater collaboration amongst the project team members (Alshawi and Ingirige, 2003). The need for increasing the efficiency of these processes via exchanging massive volumes of information at high speed and relative at a low cost has been long recognised by the industry (Dong et al, 2001). As suggested by Baldwin et al. (1999), the key reasons for implementing the web-based PM systems are combination of poor-coordination, inaccuracy and inconsistency in information sharing in most projects. The study by Thomas et al. (2004) asserted that project size was the single most important factor in determining the degree of IT use on a project. One of the primary reasons for implementation of web-based PM systems in most projects is the enhance communication between the project participants above and beyond traditional paper-
base poor communication which is believed to impact projects significantly by causing delays and inefficiencies (Williams, 2007). However, Thorpe and Mead (2001) suggested that web-based systems have the ability for generating information overload, which potentially limits its effectiveness (Thorpe and Mead, 2001). Mitropoulos and Tatum (2004) argue that the four main forces that drive the adoption of new technology by construction companies are competitive advantage, technological opportunity, improving construction processes and external requirements. However, these findings do not provide any understanding on the main forces contributing to the overall success on implementing a web-based PM system in the project. As asserted by Waye et al. (2002), successful implementation of IT requires not only strategic decision making at a top management level, but also technical managers input. Anecdotally, there have been numerous examples where the implementation of IT has been unsuccessful or did not meet the expectations of the end users in the organization. For this reason, it is important to understand the key drivers associated with the decision making of such implementations and communicate the tangible benefits among the project teams (Williams, 2007).

2.1 Hypothetical model

The aforementioned studies provide the theoretical basis to construct the model for this study. It is assumed that information streamline (IS), Transparency and accountability (TA), effective monitoring and control (EMC), users’ satisfaction (US), increased competition (IC) and project complexity (PC) collectively determine benefits of web-based project management system (PWB) being implemented in projects. All these technical factors are important in terms of rationalising the decisions for implementation of web-based PM systems and thereby exhausting the benefits in the project. Thus, a series of hypotheses can be drawn to verify the relational linkages and underlying effects between these technical factors and the performance of web-based systems being implemented in projects.

In order to analyze the PLS-SEM model, the initial path diagram was developed based on a total of 10 hypotheses as follows:

- The benefit of web-based project management system in projects is exhausted by the six key factors namely high Users’ Satisfaction (hypothesis 1), high degree of Transparency and Accountability (hypothesis 2), Increased Competition (hypothesis 3), Information Streamlining (hypothesis 4), Project Complexity (hypothesis 5) and Effective Monitoring and Control (hypothesis 6);
- Web-based systems implementation is driven by high degree of Transparency and Accountability (hypothesis 7) in projects;
- The high degree of Transparency and Accountability is due to high degree of Information Streamline (hypothesis 8), which enhances Effectiveness in Monitoring and Controlling (hypothesis 9) of projects;
- Project Complexity is easily managed with increased Effectiveness in Monitoring and Controlling in projects (hypothesis 10).

2.2 Research Method

As mentioned earlier, the survey method was adopted to test the hypotheses proposed in this study. A questionnaire survey was designed for respondents to assess the performance of web-based systems implemented in relation to the perceived influences of users’ satisfaction, transparency and accountability, information streamline, increased competition, project complexity and effective monitoring and control on those projects. The questions were phrased to ask the respondents an affirmative response on the relevant variables or indicators impacting the overall performance of the web-based systems in their projects. Respondents’ profile and the project information were also collected in the survey.

Before undertaking an industry-wide survey, a pilot study was conducted among a six member focus group explaining the research intents and the questions in order to validate the contents for accurate translation of the overall model construct. Based on the feedback received, the questionnaire was refined and the ethics clearance was obtained from the University Ethics Committee for conducting the industry-wide survey. The preliminary data was collected from a total of 15 medium to large construction firms in Australia. The target population of the survey in this study was clients, contractors, designers or consultants, project managers, site managers and site coordinators involved mostly in residential and commercials projects. In total 150 survey questionnaires were distributed amongst the identified organizations with an expectation that there would be a rate of return somewhere between 25-50% of the total number of questionnaires issued. This range of rate of return would be considered applicable and consistent with other reported mail surveys in the literature reviewed (Fellows & Liu, 1997; Stevens, 2002). The overall aim was to distribute the survey questionnaire to as many applicable organizations and to receive at minimum five questionnaires back from any one firm. Studies have shown that as low as two people completing a research questionnaire can provide a general consensus on how the remainder of the organizations employees may also perceive the data content under question (Cheung et al 2004). In total 77 completed questionnaires were returned, constituting a 48% response rate. Eight additional questionnaires were received incomplete and thus excluded from the sample in the analysis. Such a response rate was primarily due to the selection of the sample and the interaction between the researchers and the respondents in confirming willingness and participation in the study.

While 23% of the respondents have been in the industry for less than 5 years, 51% have been in the industry with the
experience bracket of 6-15 years. The remaining 26% respondents had the experience over 15 years and considered fairly senior among the participants. As uptake of the web-based systems in most construction firms are spanned over past 5-7 years, over 73% of the respondents were able to provide with informed judgements based on their experiences from pre- and post-web based systems implementation periods. Of the 77 respondents, over half of them (55.6%) noted that their current project was worth between 300-599 million dollars which is considered quite significant in the context of deriving benefits from the implementation of the web-based PM systems. The remaining projects were worth between 100-300 million in the sample.

The valid dataset was then analysed on Structural Equation Methodology (SEM) using the PLS-Graph 2.0 software.

3 PLS-SEM Model

Partial least square structural equation modelling (PLS-SEM) is selected over the covariance-based SEM (CB-SEM) due to the fact that PLS-SEM is based on a series of ordinary least square (OLS) regressions which does not demand a high sample size, yet without compromising the high levels of statistical power (Reinartz et al. 2009). Conversely, CB-SEM involves constraints regarding the number of observations and small sample sizes, often leading to biased outcomes with reduced statistical power and inadmissible solutions especially in complex model set-up. Thus, PLS-SEM, which is a distribution-free soft modelling approach, is highly suitable for applications without having to make any strong assumptions which often cannot be fully met in most models (Hair et al., 2011).

The outputs from the PLS-SEM models require two step interpretations, validity and reliability testing and assessment of the relationships based on the path coefficients. The validity and reliability of the measurement model is first evaluated by examining the individual loadings of latent factors for internal composite reliability and discriminant validity (Chin, 1998). After making adjustments of the items in the model and acceptance of the final model, the relationships between the independent latent variables and dependent variables are assessed based on “standardised beta estimates” as the path coefficients. These path coefficients are then used to prove or disprove the hypotheses in the research (Aibinu and Al-lawati, 2010).

4 Model analysis and validation

The initial structural model hypothesised (Figure 1) was analysed using PLS-Graph 2.0 software and the results of the measurement model are discussed in the following sections.

Before accepting the final model for testing the hypotheses, the adequacy of the structural model was tested using individual variable reliability analysis, convergent validity measures of the latent factors and discriminant validity of the measurement model. Individual variable reliability measures are mostly the standardized loading or simple correlations of the measured variables with their respective latent variables relative to the errors (Hulland, 1999). As a rule of thumb, the standardized loadings of the measured variables should be 0.7 or more which implies that about 50% of the variance in the observed variables (i.e. the square of the loadings) is due to the latent variables. There are numerous suggestions on the threshold values of the standardised loading below which the measured variables should be dropped in the final model. While Hulland (1999) suggests the threshold as 0.4, the recommended values by Fornell and Larcker (1981) and Chin (1998) are 0.7 and 0.707 respectively. With the view of these suggestions and adopting a threshold of 0.7 in the outputs from PLS-Graph 2.0, six measured variables were dropped before accepting the final structural model. The standardized loadings of the measurement variables as shown in Figure 1 demonstrate satisfactory level of individual variable reliability in the final structural model.

5 Model Development and Validation

The paths of the structural model that corresponds to the underlying hypotheses have assessed using the PLS-Graph 2.0. Table 1 depicts the summary of the path results and corresponding significance (p-values), standard errors (SE) and the t-values associated with each structural path. As seen, all these indicators can be used to prove or disprove the hypotheses by careful interpretations of their sizes, signs and significance levels (significant at p<0.05).

Figure 1: Structural Model of the factors affecting the implementation of Web-Based project management systems
The final PLS-SEM results supported only two primary hypotheses (H3 and H5) with an acceptable level of p being less than 0.05 with reasonably strong path coefficients. These results suggest that project complexity (PC) and information streamline (IS) are the two most influencing factors with highest positive correlations (with standardise coefficient = 0.589 and 0.498 respectively) with the implementation and performance of web-based PM systems in projects. The path between increased competition (IC) and performance of web-based PM systems (PWB) with statistical insignificant (p=0.073) coefficient of 0.198 does not support the hypothesis H4. Similarly the path between information streamline (IS) and effective monitoring and control (EMC) with a statistically insignificant (p=0.089) coefficient of -0.195 does not support any direct and meaningful linkages as postulated in hypothesis H9. The reverse influence depicting the influence of effective monitoring and control on information streamline resulted from the implementation and hence hypothesis H9 can be dismissed.

The influence of the factors, users satisfaction (US) and transparency and accuracy (TA) on performance of web-based PM systems (PWB) with part coefficients of 0.231 and 0.252 at their respective acceptable borderline significances (p=0.052 and p=0.057) depict that Hypotheses 1 are 2 are marginally supported. The hypothetical paths between transparency and accuracy (TA) and users satisfaction (US) and information streamline (IS) and transparency and accuracy (TA) with relatively smaller path coefficients of 0.187 and 0.191 and with acceptable significance levels (p=0.009 and p=0.038 respectively) support hypotheses 7 and 8 partly. The structural paths linking effective monitoring and control (EMC) to performance of web-based PM systems (PWB) and project complexity (PC) with their respective acceptable borderline significances (p=0.050 and p=0.049) depict that hypotheses 6 are 10 are marginally supported but in reverse orders with their negative path coefficients of -0.085 and -0.212 respectively.

### Table 1: Path Coefficient Estimates of the PLS-Graph

<table>
<thead>
<tr>
<th>Path</th>
<th>Standardised Coefficient estimate</th>
<th>Sig. (p)</th>
<th>S.E.</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: PWB ---&gt; US</td>
<td>0.231</td>
<td>0.052</td>
<td>0.102</td>
<td>1.98</td>
</tr>
<tr>
<td>H2: PWB ---&gt; TA</td>
<td>0.252</td>
<td>0.057</td>
<td>0.117</td>
<td>0.66</td>
</tr>
<tr>
<td>H3: PWB ---&gt; IS</td>
<td>0.498</td>
<td>0.008</td>
<td>0.098</td>
<td>0.56</td>
</tr>
<tr>
<td>H4: PWB ---&gt; IC</td>
<td>0.198</td>
<td>0.073</td>
<td>0.104</td>
<td>1.08</td>
</tr>
<tr>
<td>H5: PWB ---&gt; PC</td>
<td>0.589</td>
<td>0.004</td>
<td>0.097</td>
<td>1.17</td>
</tr>
<tr>
<td>H6: PWB ---&gt; EMC</td>
<td>-0.085</td>
<td>0.050</td>
<td>0.142</td>
<td>0.51</td>
</tr>
<tr>
<td>H7: US &lt;-- TA</td>
<td>0.187</td>
<td>0.009</td>
<td>0.091</td>
<td>0.43</td>
</tr>
<tr>
<td>H8: TA &lt;-- IS</td>
<td>0.191</td>
<td>0.038</td>
<td>0.192</td>
<td>2.65</td>
</tr>
<tr>
<td>H9: EMC &lt;-- IS</td>
<td>-0.195</td>
<td>0.089</td>
<td>0.131</td>
<td>0.55</td>
</tr>
<tr>
<td>H10: PC &lt;-- EMC</td>
<td>-0.112</td>
<td>0.049</td>
<td>0.271</td>
<td>1.22</td>
</tr>
</tbody>
</table>

*All standardised coefficient estimates are significant at p=0.05*

### 6 Results and Discussions

As depicted in Table 1, the test results generally support the relationships between project complexity, information streamline and implementation of web-based project management systems which validate both hypotheses 3 and 5 reasonably well. A significant and positive relationship (with standardise coefficient = 0.187) was found between transparency and accuracy and users satisfaction. Similarly, a significant and positive relationship (with standardise coefficient = 0.191) suggests that information streamline potentially influence in achieving the transparency and accuracy in projects. Due to the relatively smaller path coefficients between the respective two factors (with standardise coefficients of 0.187 and 0.191 respectively), these findings support hypotheses 7 and 8 partially. The path coefficients between users satisfaction and performance of web-based PM systems (with standardise coefficient = 0.231) and transparency and accuracy with performance of web-based PM systems (with standardise coefficient = 0.252) with borderline significances suggest that hypotheses 1 and 2 are marginally supported respectively. The existence of marginal negative standardised coefficients of -0.085 between the factors EMC and PWB and -0.112 between the EMC and PC with borderline significances clearly make the hypotheses 6 and 10 only marginally supported in reverse orders. The influence of effective monitoring and controlling of projects on the implementation of web-based PM systems is found to be marginal which contradicts the findings reported by Lee et al. (2005). The assertion that the demand for effective monitoring and controlling increases with the increase in project complexity is also found to be marginally supported in this research. While the increased competition is believed to have a direct influence in implementation of web-based PM systems within construction organisations, the finding of an insignificant standardised coefficient (= 0.198) between the two factors does not support this hypothesis (H4). Similarly, information streamline resulted from the implementation of web-based PM systems does not contribute in effective monitoring of the controlling of project due to an insignificant standardised coefficient (-0.195) between the factors and hence hypothesis 9 is rejected.

This study has proved the hypothesis that the project complexity and information streamlining are the two most influencing factors for exhausting full benefits of the web-based PM systems implementation in projects. The assertions by El-Gohary and El-Diraby (2010) that increased necessity of the seamless integration of knowledge-carrying processes required in complex projects is significantly facilitated by the implementation of online collaborative platform are clearly validated in this research. It has been revealed that transparency and accuracy in information communication and satisfaction of users with added confidence significantly contribute to effectiveness of the web-based PM systems in projects and hence improving the rate of success. Therefore the
blanket rule that web-based PM systems is an absolute necessity and key driver for such implementation is due to increased market competition and technological advancement eventually undermine the objective decision making in projects (Nithithamyong and Skibniewski, 2004). This finding is in the line of the previous findings that the necessity of web-based PM systems increases with the increased complexity of modern projects. Implementation of web-based PM systems should be based on objective analysis of functionalities, discussions among project team members and clear support of the senior management in projects (Ryoo et al, 2010; Chan and Leung, 2004; Forcada et al., 2010). Furthermore, while the PLS-SEM model shows some contrast to the to the results published by Lee et al. (2004) and Cheung et al. (2004) contending that onsite decision making, information sharing in the context of planning and control and enhanced are some of the key factors impacting implementation and performance of web-based PM systems in projects, the findings however support the research by Mohamed and Stewart (2003) asserting that enhanced coordination and reduced response time supplemented by web-based systems implementation cannot be considered as a clear industry practice.

7 Conclusions

The model suggested that overall success of the implementation of web-based PM system significantly relies on degree of project complexity and need for information streamlining across the contracting parties in project. Underestimation of project scope and lack of in-depth understanding of information sharing and data exchange requirements across the diverse project stakeholders potentially undermines the true performance of web-based systems implementation. However, there is no any significant evidence of effective monitoring and controlling being supported by the web-based systems implementation in projects. The findings of the PLS-SEM suggest that the factors such as users’ satisfactions, market competitions, increased transparency and accuracy, real-time communication barely have any impact on objective benefits of web-based systems being implemented and overall project success. However, the implementation of the web-based PM systems for exhausting the necessary benefit have been supported by the factors associated with the complexity and necessity of the seamless integration of knowledge among the users in the project.

References


Integration of Augmented Reality and Indoor Positioning Technologies for On-site Viewing of BIM Information

Hung-Ming Chen and Ting-Yu Chang

Abstract -
This study applied integrated indoor positioning and augmented reality (AR) technologies to develop an on-site viewing model of BIM. In the proposed model, BIM information is displayed and superimposed on real building elements in 3D component models or surrounding texts using AR technology. The building must have a wireless network environment inside, and each room uses the same set of markers to mark up the building elements in the room. The system uses wireless access points for indoor positioning to identify the user’s location in the room, and then uses markers to identify the building elements in the room. Once a building element is identified, the system retrieves the 3D model or property information from a Cloud-based BIM database and then superimposes them on a view of real building elements. A prototype system was developed based on the proposed model using the WorldViz Vizard platform. The proposed system achieved the integrated display of virtual BIM and real on-site scenes on a head-mounted display (HMD) without manual navigation for model viewing.

Keywords -
IT Applications; Augmented Reality; Indoor positioning; BIM

1 Introduction

Building Information Modeling (BIM) is used to integrate a three-dimensional visual model of construction projects and digital data of various fields with properties into a file or a database with a specific format, allowing project designers, project managers, construction units, owners, and clients to view the design through a three-dimensional visual model and to obtain relevant digital data of the project in a three-dimensional visual model [1]. There are several commercial software systems that are BIM-based, such as Autodesk Revit, Tekla Structures, Bentley Architecture, and Bentley Structural. These commercial BIM software products can integrate 3D models with architectural, structural, electromechanical, air-conditioning, firefighting, and other various field data and store them in a single common BIM project. A feature of these commercial BIM software products is that they primarily support the CAD system of BIM design for a studio work environment. Currently, BIM has been applied successfully during the planning and design phases of projects. However, when one wants to use BIM to support tasks during the construction and maintenance phases, one must bring the BIM to the site and use it. If a user needs to view and query while on the site, the existing approach is still to export the model as 2D floor plans and to bring them on-site to be used with conventional methods, or to use a tablet, notebook, or other mobile devices on the site to run the BIM Viewer software. When a user wants to query the property data of a certain actual building element through BIM Viewer, he/she must first manually search for the 3D scenes in the BIM, find them, and then select the corresponding BIM component to see its relevant components and parameters. Such an approach requires the manual operation of model viewing to return the necessary BIM information to the user. The model is unable to perform automatic interception or to provide immediate feedback regarding the information, and the BIM presented through the computer screen and the actual on-site scene are also provided in two separate views. In this viewing method, the information between the two is less likely to be synchronized and combined. Therefore, BIM applications in engineering are currently mostly used for the planning and design phases. Its application in the construction or maintenance phase is relatively rare.

Augmented Reality (AR) is a technology which utilizes computer visualization to superimpose virtual objects on real world images for user interaction [2]. Its mode of operation is to calculate the position and angle of a video device through real-time positioning or image processing, and then superimpose a virtual model and information on the image of the real world. Users can
IT APPLICATIONS

instantly and automatically obtain the necessary information based on their location. Its objective is to present virtual objects in the scene of the real world via a video device and to allow users to interact with it. Some studies have attempted to apply AR technology in the field of construction engineering. Park et al. [3] proposed the application of a construction defect management system using BIM and AR. In this system, site executives make use of the BIM data obtained from the database, convert it into a 3D model through smart mobile devices, and demonstrate it on-site. Not only can the on-site workers confirm their work, reducing the errors made, it can also allow an on-site inspector to check if their work is proceeding correctly, thereby reducing defects throughout the construction process. Behzadan and Kamat [4] proposed a teaching system that combines combines AR and real world views of a construction jobsite. In their system, real-time imaging of the construction site can be obtained via an IP-accessible camera and transmitted through the internet to a projection screen for display in a classroom; the positioning system can receive the location of a user and interact with the AR teaching books, designed by the author to achieve the purpose of teaching. Bae et al. [5] adopted a hybrid 4-dimensional AR (HD4AR) system. It is a high-precision mobile AR system that allows on-site workers using mobile devices to capture an image of the building element, enquire, and receive messages in AR as well as overlapping 3D network information that superimposes the real-world. HD4AR can contain important projects, such as the position of the element, progress, specifications, and cost (which have been traditionally difficult to render on the construction site), and act out these projects by superimposing the real world. Users can use a mobile device with a video camera for wireless network positioning, uploading, or downloading the related models, and reading the relevant information. Feng and Kamat [6] employed AR technology to develop a set of indoor Architecture, Engineering, Construction, and Facilities Management (AECFM) systems, with which users can identify AR markers using a smart mobile device. The system calculates the user’s pose, the marker’s pose, and the relative pose of the destination, then reads the data on the database, returns it to the system, and displays it to the user for navigation.

However, most of these studies applied the AR technology on an outdoor on-site construction environment; relatively few studies have applied the AR technology in an indoor environment to support construction and maintenance operations. Therefore, based on the application concept of an outdoor environment AR system, the motivation of this study is to integrate AR and indoor positioning technologies to superimpose virtual reality images of BIM components and properties inside an indoor work environment, so that user do not require a complicated operation manual for model viewing. The virtual reality images can be superimposed on a single screen, provide an immediate view of BIM, and apply it to support construction and maintenance operations.

2 Objective

The purpose of this study is to propose an on-site view mode of BIM information by integrating AR and indoor positioning technologies. This mode was applied to superimpose the BIM information directly upon a real-time video of the building interior. By overlaying on the architectural elements via an AR device, it displays the 3D model, established using BIM components, and the property parameters of the elements. This is directly presented on-site, providing an immediate feedback to the user.

For this purpose, the indoor positioning technology is first employed to locate the room where the user is. The location of the user is then coordinated with an AR marker to superimpose the created BIM component, including the building elements and the property parameters. With a 3D model to render the exterior of this component and present the properties in text, this information is then superimposed onto an actual building element of the on-site room, affixed with a corresponding marker.

Using Wi-Fi positioning [7] as the indoor positioning technology, viewing of the component through a display requires the corresponding marker. Hence, before using the system, on-site construction and maintenance staff need to set up wireless access point (WAP) for Wi-Fi and configure the markers on building elements such as walls, doors, windows, floors, etc. WAP carries out indoor positioning through the signals emitted by the notebook of this system. Based on the positioning space, the system downloads all BIM components of this space and superimposes virtual reality onto the building elements using the marker.

Users can view the BIM components through wearable video devices and carry out shifting, rotation, scaling, and transparency changes of BIM components.

Before using the in-site viewing mode of BIM information with the integrated AR and indoor positioning technologies, the proposed system must first be pre-set on-site, as described in Figure 1. Its mode of operation is illustrated in Figure 2.
The building interior requires the establishment of a wireless network environment. All rooms used a group of identical markers, according to the database of its BIM components, affixed onto each building element inside the space. When a user enters the room on-site, the system starts downloading a reference table from the BIM database. Its contents consist of the rooms of the indoor site and the corresponding BIM component codes. To carry out indoor positioning via WAP and to identify the space where the user is, the system automatically downloads all the BIM components of this space, then coordinates with the marker on the building elements for identification. It can also access the 3D models and related property parameters of the elements from the BIM database based on the identified building elements. Users can instantly view the BIM model and parameters of the target component via a wearable display device as well as using a keyboard or a mouse to rotate, move, and scale the BIM components. They can even change the transparency to facilitate viewing of the BIM components and the hidden components. This integrated model can be conducive to the work of the on-site construction and maintenance staff. They no longer need to view the BIM through notebooks and tablets. By using the model proposed in this study, users will be able to view the BIM in a way that is more convenient, faster, and more efficient.

3 System Requirements and Mechanisms

In order to achieve the objective mentioned, we investigated the system requirements for each function and proposed the mechanism plan for system operations. Namely, these functions include types and properties of BIM components, database schema for BIM storage, BIM component lookup mechanism, workflow of Wi-Fi positioning technology, connection with BIM database, BIM component information retrieval, marker design and deployment rule, BIM component display mechanism, as well as BIM component model manipulation. The following subsections focus on the analysis and planning description of these 9 directions.

3.1 Types and properties of BIM components

In the field of civil and construction engineering, there are many types of building elements, such as walls, columns, beams, and floors. As the system is a prototype for verifying the feasibility of the model, only 5 building element BIM component types were used. The models were walls, doors, windows, floors, and pipelines which were sourced from Google Sketchup. The models were altered in accordance with the properties of their BIM components in order to match properties of the on-site building elements, such as the size, and species. As noted in Table 1, the properties of BIM components are presented in text format on the system's screen.

<table>
<thead>
<tr>
<th>Component type</th>
<th>3D model</th>
<th>Component properties</th>
<th>Property presentation format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall</td>
<td><img src="image" alt="Wall" /></td>
<td>position, length, width, height</td>
<td>Text or number</td>
</tr>
<tr>
<td>Door</td>
<td><img src="image" alt="Door" /></td>
<td>area, volume, material</td>
<td></td>
</tr>
<tr>
<td>Window</td>
<td><img src="image" alt="Window" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor</td>
<td><img src="image" alt="Floor" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pipe</td>
<td><img src="image" alt="Pipe" /></td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

3.2 Database schema for BIM storage
In this study, the BIM data sources were obtained from an established cloud database of Apache Hbase [8]. Apache HBase is a structured database system, developed in the Java programming language and set up on distributed systems of the computing cloud. The data is unstructured, untyped, has no functions of JOIN or QUERY, and is not supported by the SQL syntax. The storage method provided can be regarded as an infinite table, in that the information is stored into table, rows, and columns in a logical and organized way. In this study, we took a building as a unit. Each floor was distinguished using rows; information of the BIM components on a floor was stored in the same row. Each BIM component and the BIM components attached on it were taken as a unit of a Column Family; a group of BIM components is stored in a column of a Column Family, as shown in Figure 3. In Figure 3, the name of the table is BIM. Its row is 4F, implying that this row contains information of the fourth floor. The Column Family of B1 contains BIM component data of a wall and hidden pipelines; B2 contains the BIM component of a door; B3 contains the BIM component of a floor and so on.

### 3.3 BIM component lookup mechanism

The Apache HBase adopted in this system was a BigTable type of NoSQL database system. Because of its non-relational database, data storage and creation do not rely on the primary key, but rely on the structural characteristics of the data. Thus, the form and concept of this database is comparatively different from a traditional relational database.

There are two tables in the storage of the BIM database in this system. One is for storing the properties of BIM components, and is called the BIM Table; the other is for storing BIM components as well as the corresponding room and marker, and is called the reference table. When the system is first initiated, the system will download the reference table, then locate the current position with the indoor positioning system by looking up the reference table for all the corresponding BIM components in this position and download them into the system. Finally, marker identification is captured by a video device via AR technology to present the corresponding BIM component. This mechanism is shown in Figure 4. Within the Family Column of the Reference Table, we will look at 417 with C1 which has a Value of B1. This represents the marker encoded as C1 in the room 417 is corresponding to the BIM component encoded as B1. B1 corresponds to the Family in the BIM table, which stores the BIM components in the column B1. Loading the value of column B1 lets us obtain the properties of the B1 component. Throughout this lookup mechanism, data replication of BIM components can be avoided by having Family in the BIM database as the primary key for query. One BIM component is only stored once, which is in line with the principles of BIM.

### 3.4 Wi-Fi-based indoor positioning

The indoor positioning technology of this integrated system is a Wi-Fi positioning technology which adopts the triangulation method. Hence, prior to operation, the indoor construction and maintenance staff on-site need to configure at least 3 available WAPs to carry out Wi-Fi positioning for the system, as shown in Figure 5. Each separation distance should not exceed 50 meters, enabling the formation of a local area network to transmit information to one another.

A building contains many architectural elements. To use ARToolKit [9] to render the BIM components, many markers are required. Taking a ten-story apartment building as an example, if each floor has ten rooms and each room has ten architectural elements, 1000 different markers will be needed, resulting in less efficient implementation. Therefore, the system integrated 2 sets of Wi-Fi positioning software, AeroScout MobileView [10] and AeroScout System Manager [11], in order to reduce the number of markers to a minimum. The same group of markers presents
different BIM components in different rooms by using Wi-Fi positioning technology to locate the current position of the user. Then, this system should be able to determine the BIM components for display.

Figure 5. Deployment of WAP.

When installing AeroScout MobileView and AeroScout System Manager, a database for Wi-Fi positioning will be created automatically in the Microsoft SQL Server. Before use, it requires system settings. First, AeroScout System Manager loads each floor plan of the room on-site. The floor plan must contain the set scale and WAP information, such as location, IP and MAC. Then, AeroScout MobileView is needed to import the set floor plan and the designated room range. Since the system is developed in the programming language of Python, it cannot directly access the information in the Microsoft SQL Server. After setup, the system loads the room information on the Wi-Fi positioning database by having the necessary Pyodbc kit execute SQL queries such as SELECT, WHERE, IN, etc.

3.5 Connection with BIM database

The BIM database contains the necessary information of the BIM components. Since this system was developed and integrated using the Python programming language, whereas Apache Hbase was developed using the Java programming language, the system cannot directly access the data in Apache Hbase. As a communication bridge, Apache Thrift is required to connect the two different programming languages. Apache Thrift acts like an interpreter of the human world who is able to translate two different programming languages and enable them to communicate.

When using Apache Thrift to connect to Apache Hbase, we must first specify the IP address and port number of the computer with this database. After the system sends a connection request for the database, Apache Hbase will then respond to the request by creating a connection. At this point in time, the built-in commands of Apache Hbase are used to read the tables for downloading. The data is saved into a pre-established array and the connection with Apache Hbase is finally terminated.

3.6 BIM component information retrieval

After the BIM component information has been downloaded from the BIM database to the system, it cannot be used directly because of the data format and pattern. There is a need to perform BIM data analysis and processing before being used. After the system loads the BIM database via Apache Thrift, the BIM data is in a TCell format, containing the two properties of Timestamp and Value. Timestamp is the corresponding time when the data manipulation occurred, and represents the storage time of the BIM data; Value contains the properties of BIM components, such as the size, types, and materials. The system must first retrieve the data within the Value because its contents are a sequence of characters. Hence, it is required to carry out analysis on the sequence of characters for the required content, such as the size in numerical values, type, and materials of BIM components. Subsequently, these sequences of characters are stored into an array in the format of strings. As the system does not recognize the numerical value, the numerical value portion is converted from the string format to a number format, as listed in Figure 6.

3.7 Marker deployment rule

In this study, the AR system used for integration was ARToolKit. BIM components and properties need to be presented on the marker. Before usage, the indoor construction and maintenance staff on-site must first carry out configuration for the marker. The connection between a marker and a BIM component is in the form of 1-on-1. According to the BIM database, one BIM...
component corresponds to one marker. Thus, all the building elements to be viewed must be configured.

In the system, the design and number of markers can be determined by the user. All markers must comply with the recognizable style of ARToolKit. Users can decide to increase or decrease the number of markers in accordance with the number of BIM components. A marker with a complete design requires the marker capturing program, provided by ARToolKit, to be sent back to the system to be used, as illustrated in Figure 7.

![Figure 7. Marker creation module of ARToolKit.](image)

According to the spatial form and BIM components, this study set a group of 9 markers. Each room had the same group of markers with the identification of A, B, C, D, E, F, G, H, and I. Each marker in a different space had corresponding BIM components and possibly displayed a different BIM component, as shown in Figure 8.

![Figure 8. The system shows different BIM components on the same marker in different rooms.](image)

3.8 BIM component display mechanism

The BIM component of this system is rendered via Wi-Fi positioning technology and by superimposing a marker on a target component which is chosen to be viewed. When a user first initiates the system, the system will start downloading the Reference Table from the BIM database. This is a comparison table which contains the BIM components and the corresponding room and marker. Subsequently, the system will drive the notebook of the wearable device to send out a wireless signal to the WAP. After receiving the wireless signal from the notebook, the indoor positioning system will carry out indoor localization on the area where the user is. Having the Reference Table with the Wi-Fi positioning software to locate the room from the database, the system will load all BIM components of the room, then capture and identify the marker inside the room using a video device with AR technology. Finally, the system can connect the marker with the corresponding BIM components and superimpose the virtual BIM components on the actual building element.

Every minute, the system will load the user’s current location. Based on the user’s current location, it will determine the need to re-download the BIM components. If the user enters another room and wants to view its BIM components, this system will re-download all the BIM components of the new room; if user is in the same room, there is no re-downloading required.

The BIM components of this system need to be rendered on a marker. For each room, the connection between the BIM component located in a room and the marker laid inside this room is 1-on-1. One unique BIM component corresponds to one unique marker. Markers need to be placed on the building element to be viewed. All markers and BIM components have their own codes. Markers and BIM components with the same code corresponding to each other to form a connection. When a user views any of the markers inside a room via a wearable AR device, the system will read the BIM component with the code that is consistent with the code of the marker. For instance, a user is currently in the room No. 417; if the user wants to view a door and the door has been affixed with the C1 marker, the system will read all the BIM components of room No. 417 from the downloaded Reference table; when the user views the marker affixed onto this door via a wearable AR device, the system will superimpose the BIM component with the room No. 417 and marker code C1 onto this door. Figure 9 shows snapshots of the superimposed scenes displayed in the prototype system.

![Figure 9. Snapshots of the superimposed scenes displayed in the prototype system.](image)
3.9 BIM component model manipulation

In order to let a user view the on-site BIM components more easily, the system provides operating functions to view the model from different angles. Due to the many different kinds of buildings, the format and configuration of the interior may also be different. Viewing of the BIM components may occur under such influences and restrictions. Even during the setup of the system hardware, if the construction or maintenance staff on-site do not affix a marker onto the correct position, the virtual BIM components cannot be superimposed precisely on the building elements, resulting in uneasy viewing. Therefore, the system provides functions of model operation, allowing users to control the model view more conveniently using a keyboard and mouse.

The system includes 6 functions to operate the BIM model. These are model position shift, model angle rotation, model size scaling, model transparency change, on/off property display, and restoring original state, as demonstrated in Figure 10.

4 System Framework

The proposed on-site viewing model of BIM information integrates the indoor positioning, AR, and BIM indoor positioning. The source of BIM components is from a well-established Apache Hbase cloud database, and contains two tables. One is the BIM information for the components, designed based on the building interior, and includes 3D models, properties data, geometric data, and its relationship. Another is a comparative table of the building room and the BIM components and allows the downloading of all BIM components of the current room. The AR technology superimposes the virtual reality of BIM components on the on-site building elements by using a video device as well as providing immediate feedback to the user. Figure 11 shows a user equipped with the devices of the proposed system.

The functional framework of the system is described in Figure 12. IndoorAR4BIM represents the whole system. ReferenceLoader is the loading function on the database for uploading/downloading the Reference Table. ReferenceReader is the query function on the Reference Table; it is responsible for inquiring about information on the Reference table for system operation. IndoorPositioner is the query function of indoor positioning data; it enquires about the information of the indoor positioning database and locates the current position. BIMLoader is the loading function of the BIM data; it loads the BIM information from the BIM database. BIMReader is the reading function of the BIM data, and it is responsible for reading information of downloaded BIM components for system operation. BIMDisplayer is the rendering feature of BIM components; it connects up the BIM components and the marker, rendering the model and properties of BIM components. BIMManipulator contains the operating functions of BIM components; it is used to change the model position, size, angle, transparency, on/off properties display, and to restore the original state of BIM components.
5 Conclusions

This study proposed the on-site view of BIM information by integrating AR and indoor positioning technologies. The purpose is to have a system that uses indoor positioning technology to locate the current indoor position of an on-site user, then downloads BIM-related information from the database. Finally, a 3D model of BIM components and the corresponding properties is superimposed on the building elements via augmented technology, providing an immediate feedback to the user. This system provides (1) Immediacy: It can instantly locate the user's current location and provide immediate feedback to the user. Moreover, users can use a keyboard and mouse to change the position, angle, size, transparency, and on/off property display of BIM components, achieving the purpose of immediate feedback. (2) Synchronization: The models of the BIM components look as if they really exist on the site, achieving synchronization with the user. (3) Convenience: This approach can be more convenient for on-site viewing of BIM. The information is stored in a cloud-based database, and is able to simultaneously support multiple users. This greatly improves the efficiency of on-site construction or maintenance staffs. Summarizing the above, the proposed on-site view of BIM information by integrating AR and indoor positioning technologies can improve the existing view model of BIM to obtain information efficiently and effectively reduce the required operations. Moreover, the immediate feedback of BIM components provided to users for viewing as well as superimposing it on the actual scene in a single synchronized screen should be able to support the on-site operations of construction and maintenance. The limitations of the proposed system are as follows. If the markers are partially blocked, or if the room does not have enough lighting so the camera can detect the markers. Also, the upfront time and energy needed to install markers and then maintain them may be a prohibitive factor.

References


Towards 3-D Shape Restructuring for Rapid Prototyping of Joining Interface System

S.-K. Lee, C. Georgoulas, and T. Bock

Chair of Building Realization and Robotics, Technical University Munich, Germany
E-mail: {lee.seong-ki, christos.georgoulas, thomas.bock}@br2.ar.tum.de

Abstract -
Ready-made building components and joining systems are generally unfit for existing building structures with irregular shapes. Therefore, unplanned and immediate manual work is dominantly practiced as measuring and manufacturing of the customized joining system is not supported. In this respect, by obtaining 3-D shapes of the existing building components and using them for the customization of the joining parts, construction productivity can be improved while minimizing the construction errors. The aim of this research is to develop a joining interface production system that can be adaptively reconfigured for general renovation of building facades. The proposed system can precisely capture the built geometry of the existing infrastructure and in order to manufacture tailored joint components. This approach can not only enhance the assembly process and quality in renovation process, but also minimize the expenditure by preventing additional manual labor. As a pilot test before introducing 3-D Laser scanning for precise measurement, depth images using Kinect infrared (IR) sensors are utilized in order to register the 3-D shape of irregular building components. This geometry is modified and updated by considering the geometry of the joints, which becomes the interface 3-D model between a prefabricated joint part and an irregular pre-existing building component. For the customized fabrication of the interface 3-D model, a mock-up prototyping is tested using a 3-D printer. A 3-D rapid prototyping system using 3-D depth reconstruction is proposed and developed, for the manufacturing of 3-D joint interface components. Interoperability with BIM tools needs to be considered to integrate the design and production steps. At last, this research contributes towards robotic-oriented lean construction methodologies [1], by eliminating drawbacks by providing a CAD/CAM framework that incorporates modularity, design for assembly (DFA) and design for variety (DFV) principles.

Keywords -
IT Applications; 3-D Shape Registration; Joining System; Rapid Prototyping

1 Introduction

As the needs of owners and occupants change, buildings must be frequently renovated, resulting in a significant amount of recurring costs [2]. Among others, design, fabrication and assembly of new building components in building construction processes are the main procedures that greatly influence on the productivity, and therefore also on the feasibility. Thus, developing an efficient method for the rapid manufacturing of new building components for pre-existing buildings is of great importance in order to cope with increasing number of renovation projects worldwide [3]. Ready-made building components and joining systems are generally not compatible with existing building structures that present irregular shapes. Therefore, unplanned and immediate manual work is dominantly practiced as measuring and manufacturing of the customized joining system is not supported. In this respect, by obtaining the 3-D shape of the existing building components and using them for manufacturing customized joining parts, construction productivity can be improved while minimizing the construction errors.

The research purpose is to develop an adaptive joining interface production system that can be reconfigured for general building refurbishment, particularly for building façades. When fully developed, this system can greatly enhance the building assembly process and the building construction quality in building renovation processes.

This paper is structured in the following manner: in section 2, literature review is presented regarding issues of building refurbishment, additive manufacturing, 3-D rapid prototyping, Building Information Modeling (BIM) and reverse engineering. In section 3 the 3-D shape restructuring system architecture is depicted. Section 4 presents the ongoing development of application tool for the proposed system and finally, in section 5 conclusions and further development steps are provided.

2 Literature Review

2.1 Trend of Building Refurbishment

It is anticipated that most of the buildings in EU will be refurbished until 2050. For example, buildings from 1960s are being renovated several times in Germany and the newly built are expected to go through a same step.
According to EU report [24], refurbishment will be the largest portion of building construction sector in many developed countries. In addition, according to the Korean housing market trends, many building renovation processes will be active within several years by the new motivated building law that allows the horizontal and vertical expansion, in order to increase the housing value and to reduce the building owners’ financial burden. As the Korean National Assembly will introduce the law which allows vertical expansion remodeling, construction industry shows great interests in the relevant market. According to industry sources, the vertical extension remodeled apartments that were built the last 15 years in the country, amounts to about 430 million households [25].

When considering that the annual supply of housings is around 30 to 40 million households, remodeling with vertical extension is up to 11 times the size of the arithmetic formation of new markets. According to the Housing Act Amendments which was submitted to the Korean National Assembly last June in 2013, vertical extension of apartments that are older than 15 years are allowed up to three floors. However, in the case of apartments lower than 15 floors, expansion up two floors is possible for stability. Allowed number of units through the expansion of households is within a 15% range. Allowed expansion of each household is: area 85 m² (only area) is less than the maximum 40%, area over 85 m² is 30%. As of the end of September, according to the Korean real-estate portal service, apartments older than 15 years that are allowed to be renovated with vertical expansion, reaches a total of 4,285,130 households. Among them, 46.5% (1,992,626 households) are concentrated in Seoul, Gyeonggi, Incheon, etc. in the National Capital Region. Most buildings that were built the last 30 years, have maintained a relatively good structural soundness, so by replacing old facility with new infrastructures such as Mechanical, Electrical, and Plumbing services (MEP), combined with new structural expansion, the economy of building construction can be revitalized even during the harshness of the economic crisis. As the productivity of the construction area has been decreased, the foremost goal in construction is to maximize the productivity for the Return on Investment (ROI). About 10% of the global population work at the construction area but, the productivity has been decreased continuously. As the birth rate in EU and developed countries such as Japan, Korea is decreasing, it is expected that a shortage of workers will be encountered, and safety issues in construction will comprise one of the major concerns that must be tackled, but without admitting construction innovation, this might end up with being very difficult task. From this viewpoint, it is asserted that new technology can minimize the experienced pitfalls.

2.2 Technical Outlook

To reach the goal in case of building refurbishment, interdisciplinary efforts need to be considered. There are several key technologies that can influence on current building refurbishment projects. First is the surveying system which can capture the existing physical as-built configuration. By using the hardware and the supporting software for dealing with huge data (e.g. Big Data), it is possible to acquire huge data within short time. Conventional imprecise measuring that relies on manual methods will be replaced by new computational methods. Moreover, according to current researches on precise measuring system development that targets at forestry growth tracking, maximum error between actual model and captured data is less than 6 mm within 1 meter distance using consumer level stereovision systems [4] (e.g. Microsoft Kinect or Asus Xtion Live Pro). That can enable early design fixation during the whole design stage and as a result, more flexible decision making, such as new production method adaptation. Second is the advancement of Additive Manufacturing (AM) technology which is commonly known as 3D rapid prototyping. Still, 1:1 scale building production is limited as can be seen from the examples of 3D-Shape and Contour Crafting by the USC [5]. By inventing new construction materials and optimization algorithms for production, production for building component can be quite flexible enough to resolve building constraints, and satisfy dwellers’ unique and diverse needs. Third is the development of BIM applications in AEC/FM. Decade year’s long effort by IAI alliances to make information sharing tools among stakeholders promotes interoperable BIM software development, which will transform construction industry to that of production industries. Open API that is supported by BIM software vendors also can help third party developers program advanced software modules for the specialized information process service (e.g. MagiCAD™ supports MEP design in Revit).
Figure 2. Building refurbishment research process framework

Originally actively developed in Japan for construction automation, specialized assembly robots have been developed worldwide to enhance the constructability by alternating dangerous manual construction works onsite. Effects from this application in construction are not only the improvement of productivity, but also fatal construction accidents can be significantly reduced. For new construction project (e.g. Dubai' Burj Khalifa), for example, factory-like construction equipment is mounted on top of the building to function as onsite construction bed. This kind of construction factory is advantageous in that the working condition is seldom influenced by natural environment, and automated hoisting and moving devices can speed up the whole construction process. However, still the robotics-assisted assembly strategy for the building refurbishment for better assemble-ability did not receive much attention and hoisting vehicles have been planned and operated by human workers nearby.

In addition, as the building façade system has developed to integrate energy control system by combining diverse systems, which change the static building façade system into the active system that interacts with environmental entities, onsite joining and interfacing works will be a challenging issue that otherwise hinders the productivity improvement. As a preliminary research for combining those technics that are mentioned above for building refurbishment, a preliminary research on rapid manufacturing method for joining interface will be introduced in this article as conceptually represented in Figure 2.

2.3 Reverse Engineering

Scanning technology using Lasers beam leads to the development of reverse engineering in the construction engineering industry [6]. Reverse engineering generally utters the process of discovering the technical principal of devices and systems, as opposed to the concept of designing. Originally used for the analysis of military and commercial hardware, reverse engineering aims at reasoning design decision making process without prior knowledge of the production process of the original product [7] or only with partial knowledge of the product. While the designing of target objects precedes the production of the objects in general engineering, reverse engineering extracts shape feature information from existing objects. In particular, reverse engineering in the construction industry means the process of extracting shape information from constructed building objects, and utilizes it for additional design changes and production-related activities, etc. [8]. There are active researches to gather information of pre-constructed buildings, bridges, and tunnels for various purposes during their life cycle. Point-cloud data from a 3-D Laser scanner (3-D point cloud) is converted to various CAD representational models, such as B-rep, Constructive Solid Geometry, tessellated surface model, and so on; and the models are utilized for various engineering tasks such as benchmarking, quality control, and advanced drawing information management [9]. But reverse engineering is relatively in its inchoate stage in the construction industry, and is generally restricted to the acquisition of drawings of as-built building structures when there is no drawing information such as in case of historic buildings. Nowadays, for quality control in construction processes, decision making based on numeric analysis by comparing CAD data and 3-D scanned data is necessary. In these applications of reverse engineering, it is challenging to enhance the level of precision of the obtained 3-D shape information as shown in Figure 3.

Figure 3. Real-time 3D shape restricting demo (Photo: S.-K. Lee)
2.4 Applications of BIM

The preparation of 3-D parametric building models, which is based on the survey data that serves the design, prefabrication and installation processes, has been the focus of a few studies [10]. As mentioned before, the analysis of the survey data and its conversion into a BIM model is still a major challenge. An additional challenge is how to include non-geometric information in the model as object attributes. Such information can be of importance for the construction processes (information on structural elements onto which the new elements can be mounted, thermal bridges that require additional insulation, damaged elements that need repairs, etc.). In [11, 12], methods for the documenting data on damaged elements in the IFC format are suggested. Basis for all refurbishment activities is always an exact knowledge of the quality and condition of the building fabric. The aim should be to capture all necessary information during the survey in a uniform Building Information Model (BIM) as basis for all further services and activities.

3 3D Shape Restructuring System

3.1 System Architecture

As a basic study for the proposed joining system for building refurbishment, the authors suggested stepwise research models:

1. Scan-planning the work process for building refurbishment
2. Scanning the target building objects using depth images
3. Post processing the acquired data
4. Merging the files using ICP (Iterative Closest Points)
5. Exporting CAD file (.stl file)
6. CAD Modeling and modifying in BIM authoring tool
7. Automatic joint modeling using Add-on Design module
8. Converting the modeled file to STL file for Additive Manufacturing
9. 3D printing
10. Assembly test for verification

3.2 3D Shape Restructuring Algorithm

Infrared sensors such as MS Kinect™ and Asus Xtion Live™ [4] brought low-cost and commercial application quality for depth sensing. These 3-D depth sensors extract a depth image by emitting a structured pattern of light and by measuring the pattern deformation that is transformed by objects in the optical scene. Experimental results show that the random error of depth measurement increases with increasing distance to the sensor, and ranges from a few millimeters up to about 4 cm at the maximum range of the sensor [4]. OpenNI library is used to get the depth map images from the Asus Xtion Live device.

Patterned IR rays are emitted from the IR emitter and the reflected images are sequentially gathered by IR receiver. Maximum spatial resolution of each captured frame is 640x480 pixels, and each pixel value of the IR image represents the distance between the device lens and the reflected object. To convert the depth map image to the actual feature dimension, simple triangulation equation is used to project to the real world. The translated point clouds are converted to STL files from the developed software module and each file is merged using Meshlab software. In [13], a system for accurate real-time mapping of indoor scenes in variable lighting conditions is proposed. The idea is to use a single moving depth camera. They fused all of the depth data that are streamed from the sensor into a single global implicit surface model of the observed scene in real-time. The current sensor pose is obtained by tracking the live depth frame relative to the global model using a coarse-to-fine iterative closest point algorithm (See: http://msdn.microsoft.com/en-us/library/dn188670.aspx). This method is full automation method which use tracking outliers, but in case for building shape registration, only using camera movement can generate a certain level of preciseness for construction.

3.3 Interface design

One of the most important aspects to lead to the automated construction system is developing automatic assembly system using multi-purpose connectors. Currently common fixation technology (bolting, welding), which is time consuming, even when automated, is similar to the automotive industry in terms of adhesive bonding technology. The development of a connector system that connects complex components in a robust way to each other is a key point in complex products such as cars, aircrafts and buildings in particular. In order to support efficient assembly, connector systems need to, for example, be compliant or plug-and-play like.

In construction, the most promising potential for reducing cost is based on reaching a maximum degree of prefabrication by integrating different crafts. And this could considerably reduce the time necessary for interior finishing and completion. Connecting elements have to be examined with regard to positioning, adjustment, and fixation. Transition to automated manufacturing processes should be implemented in current processes of predominant manual installation and assembly, meaning in the fields of mechanical services and interior finishing [14]. In [15], several assembly connectors for the assembly of the modules were developed, considering
the structural connection, and electrical and service pipes connections. These connectors ensure automatic performance of complete assembly between modules.

Connector systems can also be designed to support efficient disassembly, re-manufacturing or recycling. In current construction practice, neither real modularity nor applications of connectors that allow for simplified and fast connection are common practice. To structure the on-site environment and reduce complexity on-site (e.g. through the number of assembly operations, or kinematic complexity) an OEM-like industry structure, shifting the creation of components, modules and units to internal and external company suppliers and integrating the principles of Robot Oriented Design (ROD) has to accompany the introduction of Automated/Robotic On-site Factories [16]. Adjusting and fixation processes during on-site assembly have to be reduced to a few minutes or to near zero for each module installation process by connectors in order to approach the high operational speeds of the automated positioning equipment. The developed connectors enable functionality of assembling prefabricated modules themselves, connecting structurally and electrically, and connecting diverse services such as plumbing, ventilation, etc.

The state of research includes prefabricated panels are mainly used as façade cladding or lost formwork [17]. Several concrete molding technologies exist; The two most widely used ones are a densely packed array of pins that are leveled by a flexible top layer [18] and the combination of wider spaced actuators with an interpolating sheet material layer that is bent into shape [19]. Recently, completely waste-free fabrication method of formwork for free-form, cast-on-site, non-repetitive concrete structures using wax formwork elements is introduced in [20], but the research is far from a fully automated robotic molding process.

From the micro-level viewpoint, a casting method using RP (e.g. Fused Deposition Modeling) was tested, and proven efficient in case of intricate and complicated pattern designs which cannot be made by a pattern maker in wood or metal [21]. So the development directions are as follows: a) to develop connector systems that robustly connect not only information or simple elements (such as cables, storage systems, etc. in computers), but also physically complex shaped, and sometimes heavy or bulky components to each other, b) to develop standardized interface and connector systems in construction in order to modularize the infill of buildings (different types of connectors should also be used - standardization of connectors will make disassembly quicker and require fewer types of tools), and c) to develop open component system kits which have the ability to be further developed and transformed into new product models or product lines.

### 3.4 Rapid Prototyping using 3D printer

Laborious formwork can be removed by direct application of additive manufacturing (AM), which can save about 50% of the total concrete work cost [22]. Additional mold parts for metal connectors and elements are needed in order to assemble prefabricated building components to the existing building façade. Additive Manufacturing (AM) can maximize the reuse of the materials by recycling the raw material up to 95%. In principle, in powder-based rapid prototyping techniques, powder compaction is used to create thin layers of fine powder that are locally bonded. An object is made by stacking these layers of locally bonded material, but stiffness for structural building components from powder-based rapid prototyping techniques needs to be tested in a hybrid way such as rapid molding for casting.

The objective of the proposed work is to introduce a revolution in on-site construction industry, from mass-production manufacturing to tailor-made manufacturing in which products are manufactured adaptively. For that, Rapid Prototyping (RP) method was used for making molding forms for casting a product or building components. The development directions are as follows: a) a new type of a digitally controllable flexible mold for building element casting, b) minimization of manual work and therefore, errors for high production and assembly quality, c) an economically viable variable formwork process owing to rapid molding time scheduling, low manual work, and low formwork material cost, d) a sustainable onsite building component production process cycle via maximized raw material reuse, e) Optimized onsite building component process for enhanced workability by considering work breakdown structure (WBS).

### 4 Tool Development

Figure 4 depicts the process diagram of building refurbishment project. The old building is scanned using both 3D Laser scanner, for high quality building shape registration, and low-end Kinect depth map sensor for acquiring shapes for interface and joining parts. This CAD information is post-processed via the BIM tool and is quickly measured and verified via 3D rapid prototyping before final module fabrication. Remaining processes such as robotic assembly, logistics via tracking technology, will be combined.

Figure 5 shows an example of multi-service integrated joining system, which function can be alternated flexibly and by introducing structurally sound material such as steel, which can perform as a structural interlocking system between building elements. Figure 6 is an image of connector parts shape registration using Asus Xtion Pro Live.
This process will be upgraded by attaching precise motion actuator that can rotate and translate regularly in order to register point clouds for better performance. Figure 7 is an independently developed software interface that can control image
streaming software by histogram distribution calculation. The closest point clouds can be sorted based on density and proximity levelness. The point clouds data can be exported as an excel style sheet data (e.g. excel) or Stereo lithography file (.stl) for other post processing such as data merging for better quality control using third party software.

Users can manually modify the focus area if necessary by adding a bounding box in “Object Map Window” as shown in Figure 8.

5 Conclusions and further development

This article’s purpose is to describe the preliminary research direction for the building refurbishment project using advanced technology such as reverse engineering, BIM, customized component production and robotics assembly. Judging from each technology level, added value can significantly influence on the productivity of building construction, safety improvement. Next research will be how to integrate those ideas for better performance and optimization. At the end, this study can contribute to the tool development for building refurbishment by combing design for manufacturing and design for assembly. When fully matured, this approach can allow cost-effective principle to work while permitting individual adaptation without increasing cost.

Acknowledgment

The research of the first author is partially supported by Erasmus Mundus AUSMIP+ program.

References

[18] Larsen, K. E., Lattke, F., Ott, S., and Winter, S. Surveying and digital workflow in energy performance retrofit projects using prefabricated


Reactive Adaptation of Construction Schedules by Applying Simulation-based Optimization

Kamil Szczesny and Markus König

*Computing in Engineering, Ruhr-Universität Bochum, Germany
E-mail: kamil.szczesny@ruhr-uni-bochum.de, koenig@inf.bi.rub.de

Abstract -

This paper deals with the adaption of construction schedule due to real-time data. This real-time data needs to be evaluated regarding the effects on the schedule. If significant delays or other problems are identified, the schedule should be adapted. That means that under the existing condition a new schedule needs to be generated. Thereby, different constraints of the target schedule like contracted delivery dates, milestones or resource allocation should be considered. In the presented approach these additional constraints are modeled and integrated into a simulation model which represents the planned target schedule. By applying simulation-based optimization a new efficient schedule is generated considering existing and additional so-called target constraints. If the identified delays are that large, such that an adaption considering all significant constraints is not possible, some constraints, like resource capacity or shifts, can be relaxed. The proposed concept for reactive adaptation of construction schedules by applying simulation-based optimization is verified by a standard construction schedule example.

Keywords -
Reactive Construction Scheduling; Schedule Adaptation, Simulation-based Optimization

1 Introduction

An efficient execution and control of construction projects depends highly on the accuracy of the underlying construction schedules. In turn, the quality of construction schedules and their being up-to-date depends on the availability of real-time data. In general, up-to-date schedules should include actual information about the overall construction progress, modified planning documents or available equipment. One crucial point is the gathering and evaluation of real-time data regarding their impact on the overall project progress. In the next step the updated or so-called actual schedule can be compared with the target schedule with the aim of detecting significant deviations. In the case that crucial delays are identified, the existing target schedule needs to be adapted. This procedure is known as reactive scheduling. Thereby, additional constraints need to be considered such as fixed material delivery dates, on-site resources, contracted milestones and established sequences. Currently, the adaptation of construction schedules is a manual, time consuming, error prone, and poorly supported process [1]. The aim of the presented research work is to enable an efficient and transparent adaptation of construction schedules considering general project objectives, actual progress information and underlying target schedules. In consequence, new methods need to be developed to model these additional constraints and to modify schedules in an optimal way.

2 Related Work

In practice, a lot of significant real-time data are logistics-related. Real-time logistics data contain information about available material, equipment and personnel as well as updated delivery dates and site conditions. In this context, Auto-ID techniques are appropriate methods to collect logistics data automatically. Biometrics and RFID are typical Auto-ID techniques that can be readily applied on construction sites and the transportation to the construction site. These data imply different types of uncertainties due to infrequent collections, varying transport times, or manual assumptions. By integrating uncertainties into the existing schedule their consequences on construction works can be evaluated. In [2] the authors developed a goods inward inspection on construction sites for prefabricated components and steelwork for bridge construction. However, their research does not consider evaluation or integration of the collected real-time data for scheduling purposes. In [3] the authors propose an RFID-based controlling management for goods inward inspection and installation of tubes. In order to assess the progress of construction activities an automated progress control by applying laser scanning technology is presented in [4]. Laser scanning technology is utilized to overcome not satisfactory results by image processing and other tech-
techniques. Their system is able to assess progress control with minimum human input. In [5] a methodology for indoor location identification and material tracking based on RFID is presented. This automated approach enables acquisition of data about construction project status in almost real-time and detects the locations of worker and material with 100% accuracy. However, none of these researches considers uncertainties and their impact on schedules.

Only few research works exist, regarding the adaptation of target construction schedules in case of infeasibility or non-optimality due to disturbances. In [6] a formal identification and re-sequencing process is presented. This approach supports fast development of sequencing alternatives in construction schedules. The approach is based on CPM, which is a dated method not considering resources or working shifts. In [7] another rescheduling approach is presented, which is based on Constraint Programming. Two rescheduling methods are presented, the complete regeneration and the partial rescheduling. However, this rescheduling approach is based on a mathematical model. Additionally, general research for job-shop scheduling exists. A review of scheduling concepts under uncertainties is given in [8]. The authors consider also various rescheduling concepts. However, none of these concepts applies simulation, but all are based on mathematical models.

Regarding the optimization of real construction schedules several research approaches exist. For detailed literature reviews we refer to [9, 10]. In [11] a crossover operator for genetic algorithms is proposed to tackle resource allocation. A time-cost trade-off mechanism to the resource allocation consideration is presented in [12]. Application of genetic algorithms to solve scheduling problems with respect to resource leveling is presented in [13]. In [14] a utilization of the Non-dominated Sorting Genetic Algorithm (NSGA II) [15] to tackle multi-mode resource constraint project scheduling problems (MRCPSP) is presented. Furthermore, the approach also considers resource allocation and resource leveling simultaneously.

In order to apply soft constraints, researchers have investigated how to formalize and model them in an appropriate way. Semiring Constraint Satisfaction Programming as a formalization methodology for soft constraints is purposed in [16]. The authors compare their methodology with Valued Constraint Satisfaction Programming. Furthermore, an exhaustive description of soft constraints and a semiring based framework is given in [17].

However, the utilization of discrete-event simulation models is an established methodology for analysis and planning of construction activities. Domain specific construction activities, resource requirements, and technological dependencies can be described by applying different modeling concepts. For example, a special-purpose simulation-modeling tool for planning and estimating earth-moving operations is developed in [18]. The authors of [19] propose a special purpose tunneling simulation tool. Nevertheless, the effort to model realistic simulation models is very high. Because of this, simulation is not applied often in practice. Therefore, recent research is investigating model driven simulation modeling. With the help of building information models (BIM) and knowledge-based methods, semi-automatic model generation and adaption can be utilized [20, 21]. Moreover, by combining the Constraint Satisfaction Approach with discrete-event simulations, it is possible to guarantee that valid construction schedules will be generated [22].

To our best knowledge, there is no research work available that tackles the integration of uncertain real-time data into simulation models for the purpose of performing a simulation-based optimization with respect to reactive construction scheduling. Therefore, we purpose a concept for reactive construction scheduling in which real-time logistics data are considered for controlling and updating construction schedules.

3 Adaptation of Construction Schedules

In this paper the adaptation of construction schedules is based on actual logistics data. The adaption is one crucial part in the context of reactive scheduling. Reactive scheduling means that real-time data, in this case actual logistics data, are considered for automated controlling and updating of construction schedules. A schematic overview of this approach is illustrated in Figure 1.

The concept consists of four steps. First, the acquisition, preparation and adaption of real-time logistics data is performed. The accuracy and inherent uncertainty depends on the location where the real-time data were collected. In addition, manual assumptions must be taken into account. In the next step, the prepared data is integrated into the construction schedule. For that purpose a simulation model is created which represents the target schedule including all activities, resources and restrictions. Thereby, a fundamental assumption is that construction activities are modeled in a highly detailed manner. For example, to construct a concrete wall-section, single activities for formwork and reinforcement installation, concreting, curing as well as removing the formwork must be specified and scheduled. In consequence, typical schedules consist of several thousands of activities. The real-time logistics data are defined as additional constraints for the involved activities. A sensitivity analysis is performed to analyze how the real-time logistics data affects the schedule and
the project duration. Based on this target-actual comparison a decision needs to be made, if an adaption of the target schedule is necessary. In the last step, the planned schedule should be updated, if crucial delays or other significant deviations were detected. Thereby, the adaption should be as much as necessary and as less as possible. That means that execution sequences, resource allocations or the planning of the upcoming weeks should be retained unchanged. The focus of this paper lays on an optimal adaption of the target schedule considering uncertain real-time logistics data.

3.1 Uncertain Real-time Logistics Data

In construction practice, a lot of significant real-time data are logistics-related. Real-time logistics data contains information about available labor resources, material, and equipment as well as updated delivery dates and actual site conditions. All these logistics data have some effects on the execution of the construction activities. For example, if material is not available the execution of related activities cannot be started. In this research real-time logistic data always results in so-called availability constraints for certain objects, like worker, material, equipment or spaces. Availability constraints are implemented as time windows.

Auto-ID techniques, e.g. biometrics or RFID, are appropriate methods to collect logistics data on the construction site or during transportation. Due to infrequent collections, varying transportation times, or manual assumptions the data features varying types of uncertainties.

Usually, such uncertainties are modeled by using probabilistic distributions or Fuzzy sets. In both cases sampling strategies, like random sampling, Latin hypercube sampling, or alpha-cut analysis, are frequently applied to generate discrete values. Discrete values, in this case discrete availability constraints, are required for evaluating the impacts on the schedule. Further information how to model and integrate real-time logistics data can be found in preceding research by the authors in [24]. For the target-actual comparison the results of the single discrete experiments are statistically interpreted. In many cases, an adaption of the target construction schedule is required if the comparison shows a significant deviation and important project goals cannot longer be maintained. In this paper it is assumed that an adaption of the target schedule is necessary and the discrete availability constraints representing the uncertain real-time logistics data are available.

3.2 Constraint-based Simulation

The basis for the reactive adaptation of construction schedules is the application of discrete-event simulation. Simulation is used to generate detailed constructions schedules considering several types of constraints, like precedence relationships, varying resources, shift calendars, and required material. To enable a flexible definition and integration of different constraints, constraint-based simulation is utilized for generating and updating construction schedules [22]. Constraint-based simulation is an extended discrete-event simulation approach. Each time an event occurs all constraints are checked to identify which activity can be started next (cf. Figure 2).
restrictions, resource requirements, earliest starting dates or fixed time windows. Meaningful conditions, like preferred staring times and resource allocations, delivery dates, or established execution sequences, can be modeled as soft constraints. Soft constraints do not have to be satisfied completely. Solely, a fulfillment degree is calculated to evaluate the satisfaction. In this paper all soft constraints are defined based on weighted or k-weighted constraints. If weighted or k-weighted constraints cannot be completely fulfilled, the weights can be used to calculate so-called violation cost factors. For k-weighted constraints the threshold k specifies a lower or upper bound for the calculation of the cost factor. Costs factors can represent monetary costs or abstract costs for an evaluation. Consequently, different schedules can be analyzed regarding their fulfillment by calculating the schedule’s total cost factor due to constraint violations.

3.3 Target Schedule Constraints

Due to the fact that the adaptation should be as much as necessary and as less as possible, additional scheduling restrictions based on general project goals and the target schedule should be defined. In the following some typical restriction of construction projects are highlighted. Obviously, this listing is not exhaustive. In some cases additional project specific constraints should be considered. However, additional target schedule constraints can be defined in the same way.

3.3.1 Delivery Dates

Frequently, short-term modifications of some delivery dates are not possible or very costly. This is particularly true, if the delivery is scheduled within a few days. For example if the delivery and the just-in-time placing of pre-casted elements is scheduled in five days, no storage areas are available on the construction site, and the supplier can only guarantee available means of transport in the following five to seven days, the activities for placing the pre-casted elements should not be postponed substantially. Time windows in which certain activities should be scheduled can model delivery date constraints. These time windows are represented by k-weighted constraints. That means, if an activity is scheduled outside the bounds of the time window, additional costs will occur.

3.3.2 Temporal Equipment

Another restriction is the assignment of temporal equipment, like mobile cranes or piling machines. The costs for hiring these machines are often very high and they are usually scheduled within different projects. In consequence, significant postponing of operation times is sometimes not possible or additional supplier need to be contracted. These restrictions can also be modeled by time windows. Activities that require special temporal equipment should be scheduled within the bounds of the time windows.

3.3.3 Established Sequences

Construction activities are often scheduled several times in the same order, in particular, in the context of high-rise buildings with similar structure. That means for each level the same types of activities are defined in the target schedule. One aim is to enable learning effects to increase the performance and the resulting quality. In consequence, the established sequences should not be changed in the scope of the adaptation. Established sequences can be considered by using weighted constraints. Thereby, the violation costs are calculated based on the distance between the activities of a certain sequence compared to the target sequence.

3.3.4 Milestones

Milestones are important events for clients and contractors. A milestone is often put at the end of a stage to mark the completion of a work package or phase. Furthermore, payments are often associated with reaching certain milestones. Therefore, another goal is to keep the defined milestones or not to exceed them significantly. Milestones are modelled as k-weighted constraints. Violation costs occur if the milestone date is exceeded by a certain time.

3.4 Additional Resources

If the current delays are significant or the target schedule is very ambitious, it is sometimes not possible to fulfill the additional target schedule constraints and the general project goals, respectively. In this case, two strategies can be pursued. One possibility is to release certain constraints. Another way is to define additional or reallocate specific resources. Additional resources can be integrated by increasing the amount or extend shifts of critical resources for a certain period. For example, the daily shift can be extended by two hours or the contractors can order weekend shifts. Reallocation means in this context that more resources are scheduled to perform a certain activity to reduce the execution time. However, this is only useful in certain cases.

3.5 Status of Construction

It is obvious that the actual status of construction
needs to be taken into account during the adaption of the schedule. Activities which are already completed are no longer considered. Furthermore, the activities which are already started will be continued if possible. However, the allocated resources can be increased to complete the activity in a shorter time. The presented paper does not cover how to collect the actual status of construction. Different techniques from manual inspection to automated approaches using laser scanning, Auto-ID systems or images are possible. In the end, information about the progress of each activity must be defined. Within the constraint-based simulation the actual status of construction is considered as the following. Completed activities are removed. Already started activities get a high priority and all precedence relationships are removed. Furthermore, the duration needs to be adapted based on the actual progress. If actual information about the resources assigned to started activities is available, the same resources should be used again in the scope of the schedule adaptation. Of course, this is only possible, if the resources are still available.

3.6 Simulation-based Adaption

The adaption of a target schedule is based on searching a solution under the consideration of all additional constraints and the uncertain real-time logistics data. In general, this problem is called a constraint satisfaction problem. Due to the fact that not only hard constraints, but also soft constraints and general project goals like time, costs and quality need to be considered, it is advisable to apply some kind of optimization strategy to find a good solution for the constraint satisfaction problem. In this case a good solution is a schedule, which fulfills all hard constraints and minimizes the constraint violation costs. Additional objectives, e.g. minimal project duration, minimal costs and maximal quality, are neglected. As mentioned before, the construction scheduling problem is represented by a complex constraint-based simulation model. Metaheuristic optimization approaches are often applied to orchestrate the simulation model in such a way that the resulting schedule is an optimal or near optimal solution of the constraint satisfaction problem.

In this paper the NSGA-II algorithm is applied. This evolutionary algorithm is chosen due to its advantageous ability to conduct multi objective optimization and because it outperforms other existing multi objective evolutionary algorithms [15]. The chromosomes are defined as an activity list to generate construction schedules. The length of each chromosome corresponds to the number of activities that have to be scheduled. Each activity is modeled by a unique key value. The first key value entry in each chromosome represents the modeled activity that in turn should be scheduled first. The last chromosome key value entry represents another activity that should be scheduled as the last activity. Thus, the order of key values within a chromosome represents the execution sequence of the activities modeled by the key values.

4 Implementation

As it has been mentioned before, one main benefit of the constraint-based construction simulation is that additional constraints can be easily integrated into an existing simulation model. However, defining a realistic and highly detailed simulation model for construction scheduling can be very time-consuming. Due to this fact, a user-friendly software tool has been developed to speed-up the simulation definition process in the last years. The so-called SiteSimEditor is a BIM-based tool to prepare input data for construction simulation (cf. Figure 3).
tivities and various constraints can be specified interactively. To keep the expenditure of time for data preparation as low as possible, reusable templates have been developed. Currently, several templates for activities, resources, technology dependencies, strategic execution sequences, and working shifts are available.

Based on the prepared and generated input data simulation experiments can be directly executed by using the SiteSimEditor. These simulation experiments are performed by a constraint-based discrete-event simulation engine, which is implemented as a component of the SiteSimEditor. Afterwards the results can be imported and visualized as Gantt charts, 4D animations or in form of different diagrams (cf. Figure 4).

4.1 Simulation Model Adaption

The SiteSimEditor has been extended to integrate target schedule constraints. For defining delivery dates and temporal equipment the existing availability and shift management plugin can be used. In general, for each resource different partially overlapping shifts, associated with duration limits, can be specified and considered within the simulation model.

New components and user-interfaces have been developed to define constraint like milestones or established sequences. However, information about upper bounds or sequences between activities, which should be observed, must be specified manually.

Other important aspects are the definition of additional resources and modified shifts to relax the existing constraints. In many cases this has to be performed to find a realistic adaption of the target schedule and to satisfy more important constraints like milestones or project duration.

5 Demonstration

In order to demonstrate the purposed approach a case study is adapted from literature. The case study is based on a simplified warehouse construction project introduced in [12]. It contains 37 construction activities, and considers precedence and resource constraints. The project is realized with 12 labor resources. Additionally, an adaption has been made to support a mobile crane for casting activities. The mobile crane is available only for a limited time period. Furthermore, a milestone has been defined that marks the point in time when all casting activities should be finished.

Discrete-event simulation is applied to generate a target schedule. The generated target schedule has a make-span of 190 days (cf. Figure 5), which is equal to the results presented in [23]. An additional adaption introduces the availability of real-time logistics data, such as delivery status of needed material. It is assumed that on day 37 of project execution, new real-time data are available. This data reports a delay of material delivery required by activity 7 and 10. In this example the

![Figure 5. Target schedule excerpt with activity delays, mobile crane requirements, and project milestone](image-url)
delay is set to seven days to enable traceability. In general, probabilistic or Fuzzy-based delays can be considered and a stochastic simulation and optimization have to be performed.

This data is prepared and integrated into the existing simulation model and a sensitivity analysis is executed. The results of this analysis are depicted in Figure 5 as calculated mean delay. The milestone is delayed by roughly 10 days and the mobile crane is not available when required. Thus, a simulation-based optimization is performed to adapt the schedule. During the reactive rescheduling some constraints are relaxed. The daily shift of the workers was extended by two hours. Furthermore, the required mobile crane is also available on the subsequent days. The relaxation of the mobile crane’s availability is modeled as a k-weighted constraint. If the mobile crane is assigned on another, not contracted, day additional costs occur. During the contracted time frame the mobile crane costs 162$ per hour. On additional days, the costs increase to 200$ per hour.

The optimization compromises the variation of activity sequences and resource allocation considering two main objectives: observance of the milestone and minimizing the total resource costs of worker and mobile crane. Four simulations experiments were performed by combining shift and mobile crane extension (cf. Table 1).

Table 1. Simulation experiments for schedule adaption

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Shift extension</th>
<th>Crane extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Exp2</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Exp3</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Exp4</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

The optimization results of the experiments Exp3 and Exp4 are shown in Table 2. The observance of the milestone cannot be fulfilled for the first two experiments. The experiments Exp3 and Exp4 differ regarding the costs due to shift and mobile crane extension. However, both schedule adaptions are possible. In the next step the project manager needs to decide which adaption is most suitable.

Table 2. Simulation experiments for schedule adaption

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Exp3</th>
<th>Exp4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs worker</td>
<td>29,900</td>
<td>43,400</td>
</tr>
<tr>
<td>Costs crane</td>
<td>63,146</td>
<td>52,746</td>
</tr>
<tr>
<td>Total costs</td>
<td>93,046</td>
<td>96,146</td>
</tr>
</tbody>
</table>

6 Conclusion

The adaption of construction schedules based on real-time data is an important aspect in the context of efficient project execution. However, the planned target schedule should be adapted as less as possible but as much as required. In this paper an adaption approach using simulation-based optimization is introduced. Applying constraint-based simulation generates the construction schedule. The information about the target schedule can be modeled by additional soft constraints. Different target schedule constraints are classified and modeled. By considering these additional constraints in the simulation model, an efficient, adapted schedule can be calculated. The optimization is based on the NSGA-II algorithm. The concept was verified by using an existing use case. It has been shown that solutions can be found which fulfill all hard constraints and additional target schedule restrictions.

Future work will compromise the definition of additional target schedule constraints. Furthermore, stochastic optimization will be integrated to find a robust and reliable adaption. Another important aspect is the availability of real-time data. Additional research work is required to improve the availability and accuracy of real-time data for construction management.

Acknowledgments

The authors gratefully acknowledge funding from the German Research Foundation DFG under grant number KO-3473/4-1.

References


Detection of Internal Defects in As-Built Pipelines for Structural Health Monitoring: A Sensor Fusion Approach Using Infrared Thermography and 3D Laser-Scanned Data

H. Son, C. Kim, and C. Kim

Department of Architectural Engineering, Chung-Ang University, Seoul, Korea
E-mail: hjson0908@cau.ac.kr, changmin1101@cau.ac.kr, and changwan@cau.ac.kr (corresponding author)

Abstract -

Internal defects of pipelines are among the main factors causing accidents in the production phase of industrial plants. Periodic monitoring of a pipeline’s inner surface condition is of great importance for minimizing the risk of failure of industrial plants. This study proposes a sensor fusion approach to detect internal defects automatically in as-built pipelines during their service lives to ensure structural safety. The proposed approach uses infrared thermography combined with three-dimensional (3D) laser-scanned data. For this purpose, a multi-sensor system equipped with a thermal infrared camera and a 3D laser scanner was internally and externally calibrated. From the combined data set, 3D points corresponding to the as-built pipelines are extracted from laser-scanned data. Then, thermographic analysis of the corresponding thermal data of those pipelines is performed. In this step, the local thermal gradients on the pipeline’s surface are calculated to detect areas having different thermal values. In addition, the global thermal gradients along the longitudinal or radial axes of the pipeline are calculated to determine the consistency of its internal thickness. The field experiment was performed at an operating petrochemical plant to validate the proposed approach. The experimental results revealed that the proposed approach has potential for detecting internal defects in as-built pipelines from infrared thermography combined with 3D laser-scanned data.

Keywords -

As-built pipeline; Infrared thermography; Internal defects detection; Sensor fusion; Structural health monitoring; 3D laser scanning technique

1 Introduction

Internal defects of pipelines are among the main factors causing accidents in the production phase of industrial plants. After a time period of operation, internal defects can develop in an as-built pipeline for a number of reasons during the pipeline’s service life [1–3]. Especially for high temperature and pressure pipelines, internal defects can be the main reasons for thinning of the pipeline’s wall, which may result in heat loss that degrades energy efficiency in the production line. In the end, such internal defects may lead to the whole pipeline’s breakdown [4]. Unplanned downtime is costly and can spell disaster for an industrial plant, and pipeline failures may lead to injuries to personnel. Therefore, periodic monitoring of the pipeline’s inner surface condition is of great importance for minimizing the risk of failure of an industrial plant and creating a safer working environment.

Monitoring a pipeline’s inner surface condition requires inspection at regular intervals to prevent breakdown. For many years, infrared thermography has been deployed to survey and inspect the integrity and performance of pipelines, including their protective coatings and insulation in industrial plants. Although thermal energy is invisible to the naked eye, it can be measured by using thermal infrared cameras [5]. Infrared thermography has been proven as a cost-effective way to survey and monitor large areas of as-built pipelines because of its non-contact and non-destructive temperature measurement properties, so that it requires no shutdown or preparation [4]. By detecting hidden defects in process pipelines, such as internal corrosion damage, infrared thermography allows for planning to repair or replace critically damaged areas before they disrupt the production process.

However, manual interpretation of infrared thermography, to locate internal defects in large areas of process pipelines and decide which pipeline is causing the problem, is not a trivial task. One needs to analyze the temperatures at various points along each pipeline. There are many factors to take into consideration, such as identifying the areas of pipelines in two-dimensional (2D) infrared thermography and the determination of internal defects in pipelines: the thermal contrasts of most internal defects in faulty pipelines are undetectable to the human eye. For these reasons, specialized
engineers conduct such interpretation to detect internal faults in pipelines from infrared thermography. Nevertheless, since analysts must interpret a large number of images (typically, several hundreds to thousands of infrared thermography images for an industrial plant, depending on its size), manual interpretation becomes extremely inefficient, sometimes inconsistent, and impractical.

The aim of this study is to propose a sensor fusion approach to detect internal defects automatically in as-built pipelines to ensure structural safety. The proposed approach uses infrared thermography combined with 3D laser-scanned data. The rest of the paper is organized as follows: Section 2 provides an overview and details of the proposed approach for detecting internal defects in as-built pipelines. Section 3 provides experimental results. Finally, Section 4 provides conclusions and recommendations for future research.

2 Detection of Internal Defects in Pipeline

This study proposes a sensor fusion approach to detect internal defects automatically in as-built pipelines during their service lives to ensure structural safety. For this purpose, a multi-sensor system equipped with a P640 thermal infrared camera developed by FLIR Systems (Notting Hill Victoria, Australia) and a FARO Focus3D laser scanner by FARO Technologies UK Ltd. (Coventry, United Kingdom) was internally and externally calibrated. Using the results of intrinsic and extrinsic calibration, each pixel of the infrared thermography is mapped to the corresponding 3D point. Once registered, the as-built 3D pipeline extraction method proposed in our previous work [6] is applied to the combined data set to identify 3D points corresponding to the as-built pipelines. Then thermographic analysis is performed on the corresponding thermal data of those pipelines to determine the existence of internal defects in the pipelines by identifying the differences in the pipeline thickness caused by corrosion. The next subsections describe each of the steps in more detail.

2.1 Data Acquisition

Especially with large facilities, single infrared thermography can cover only a limited area. For this reason, it is insufficient to infer and obtain the 3D structure of the scene given 2D infrared thermography. Therefore, identifying and locating internal defects in as-built pipelines requires infrared thermography of 3D surfaces [7, 8]. In this study, a P640 thermal infrared camera with the FARO Focus3D laser scanner is used to acquire 3D information of as-built pipelines with infrared thermography of industrial plants. Figure 1 shows a photographic image taken from a petrochemical plant.
The P640 thermal infrared camera used herein has a resolution of 640×480 pixels in a 25°×19° field of view. The advances in thermal infrared cameras have enabled accurate temperature measurement from the scene, as shown in Figure 2(a). However, infrared thermography obtained from industrial plants contains a variety of objects including not only the pipelines but pieces of equipment, structures, and many other types of objects, and these objects are tangled in a complicated way, making them visually confusing even for practiced professionals. Therefore, it is challenging to identify the areas of pipelines solely on visual analysis of infrared thermography. In addition, infrared thermography has limited utility for identifying and locating internal defects in pipelines due to the lack of quantitative information describing the locations and sizes of the defects. Therefore, the proposed approach uses infrared thermography combined with laser-scanned data. Figure 2(b) shows the colored laser-scanned data acquired from the petrochemical plant illustrated in Figure 1.

The multi-sensor system included the thermal infrared camera and the laser scanner. Both sensors were internally and externally calibrated. The calibration methods presented in Lerma et al. [9] and Borrmann et al. [10] were adopted and modified for this purpose. Using the results of intrinsic and extrinsic calibration, infrared thermography is registered with the laser-scanned data.

As a result, each pixel of infrared thermography is mapped to the corresponding 3D point in the laser-scanned data. This allows us to combine both infrared thermography and laser-scanned data, to superimpose them onto each other, and to use them together in the subsequent processing steps. Figure 3 shows the resulting infrared thermography overlay that is superimposed on the color image, for the purpose of visualization.

**2.2 As-Built 3D Pipeline Extraction**

To detect internal defects in pipelines, especially as-built pipelines, it is necessary to extract semantic information from the data. From the combined data set, the detection of internal defects in as-built pipelines begins with as-built 3D pipeline extraction. The segmentation of the laser-scanned data is performed at the intersections of the pipelines and other industrial parts to first extract 3D points corresponding to the as-built pipelines. The segmentation step uses a criterion based on a combination of surface normal similarity and spatial connectivity, which was defined by Rabbani et al. [11] as a smoothness constraint. In this step, a set of the laser-scanned data is partitioned into meaningful segments identified as belonging to pipelines or not. The next step is the extraction of sets of 3D point clouds that constitute the pipelines.

For this purpose, the as-built 3D pipeline extraction method proposed in our previous work [6] was adopted to separate sets of 3D point clouds that constitute pipelines from laser-scanned data. Pipelines typically have cylindrical surfaces. Therefore, the pipeline extraction step is based on computing the curvature at certain points on an object’s surface to decide if the segment has a cylindrical surface with the pipeline’s geometric information (radii) drawn from the P&ID.

This method requires only one-third of a pipeline’s surface to compute its radius. Then, based on the results of the curvature computation, the objects belonging to pipelines are identified and all others are discarded. For details regarding the as-built 3D pipeline extraction step, please see Son et al. [6]. By giving such semantic labels to the 3D points corresponding to the as-built pipelines, it is possible to diagnose the presence of internal defects in the pipelines’ inner surfaces with subsequent thermographic analysis.

**2.3 Thermographic Analysis**

Using infrared thermography, we can measure the different thickness of materials’ surfaces between the thermal infrared camera and the source of heat [4, 5, 7, 12–17]. In other words, internal material thickness can be extrapolated from the external temperature exhibited in infrared thermography. In this study, thermographic analysis consists of two parts. First, the local thermal gradients are calculated on the pipeline’s surface to detect areas having different thermal values. Figure 4 shows an example of a local internal defect analyzed from the combined data set visualized in Figure 3. As a result, this local defect was reported in the pipeline.
The area shown in blue has temperature values more than 10° lower than the surrounding areas in the pipeline, where the pipeline has become thin walled. Such areas can be seen in welded joints or other locations where minor defects exist, and which may result in local fracture of pipelines. Second, the global thermal gradients along the longitudinal or radial axes of the pipeline are calculated to determine the consistency of the internal pipeline’s thickness. Such areas can be seen in pipelines comprised of parts having different operating lives. Some parts of the pipelines may be replaced after being taken out of operation for various reasons. In this way, thermographic analysis allows inspection for hidden defects that are hard to find by the naked eye, such as thickness assessments of materials’ surfaces.

3 Experimental Results and Discussion

Field experiments were performed to validate the proposed approach at an operating petrochemical plant located in Yeosu, South Korea. Internal defects of pipelines in such petrochemical plants are of the greatest concern, because highly dangerous liquids and gases are processed throughout the pipelines. For this reason, in many cases, petrochemical plant pipelines are replaced long before they should be, to be on the safe side.

Figures 5 and 6 show the experimental results of two scenes, respectively. Figure 5(a) shows some portion of the colored laser-scanned data, and Figure 5(b) shows infrared thermography mapped onto the corresponding 3D points in Figure 5(a). Figure 5(c) shows an example of a recognized pipeline’s thermal data having a local defect. Here, thermal values constituting the pipeline ranged from 15° to 39°.
Figure 6(a) shows some portion of the colored laser-scanned data, and Figure 6(b) shows infrared thermography mapped onto the corresponding 3D points in Figure 6(a). Figure 6(c) shows an example of a recognized pipeline’s thermal data having different thermal values along the longitudinal axis of the pipeline. Here, thermal values constituting the pipeline ranged from 5º to 38º.

The proposed approach reported heat losses revealed in the bright parts in Figures 5(c) and 6(c), but these defects were otherwise hard to identify using the naked eye. The experimental results demonstrated that the proposed approach has the potential to automatically detect internal defects in as-built pipelines from infrared thermography combined with 3D laser-scanned data.

4 Conclusions

In industrial plants, to minimize the risk of failure and to ensure a safer working environment, one must inspect as-built pipelines with regular frequency to assure changes in conditions are recognized. This paper proposes a sensor fusion approach to detect internal defects automatically in as-built pipelines during their service lives to ensure structural safety, using infrared thermography combined with 3D laser-scanned data. The feasibility of the proposed approach was validated in experiments using real data obtained from an operating petrochemical plant. Preliminary experimental results demonstrated that the proposed approach could be successfully utilized to detect internal defects automatically in as-built pipelines during the operation and maintenance phases of industrial plants.

By performing periodic monitoring using the proposed method, facility managers can determine whether a pipeline needs to be replaced or not. This would be highly beneficial, preserving the continuity of production without unnecessary pipeline replacement and also leading to substantial cost savings. The proposed approach can contribute to preventing pipeline accidents, thereby ensuring the in-situ structural safety of industrial plants. Future work will be devoted to developing automatic orientation of infrared thermography with 3D laser-scanned data to increase the efficiency of data acquisition. Moreover, this approach will be extended to address process equipment condition monitoring and predictive maintenance.

Acknowledgements

This research was supported by a grant (12 High-tech Urban C02) from High-tech Urban Development Program funded by Ministry of Land, Transport and Maritime Affairs of Korean government.
References


Roadmap to Guide Construction Safety Research Commercialization

A.M. Costin a, J. Felkl b, O. Golovina c, and J. Teizer d

a,d School of Civil and Environmental Engineering, Georgia Institute of Technology, USA
b Cockrell School of Engineering, The University of Texas at Austin, USA
c Moscow State University of Civil Engineering, Russia
E-mail: aaron.costin@gatech.edu, j.felkl@utexas.edu, teizer@gatech.edu

Abstract -
There is a great difficulty for students or faculty to commercialize their novel idea or invention from the university laboratory. Traditionally, university research has not been intended for commercializing while still within the university. However, recent trends create a paradigm shift for researchers to include a business mind-set to establish technology commercialization objectives. Still, there are major gaps to the widespread adoption of the practice of commercializing university research. There has been great success in research and technologies for construction safety research, including hazard detection and monitoring, resource monitoring, preventative practices, and Occupational Safety and Health Administration (OSHA) rule checking. However, there is an overall lack of commercialization of construction safety research which prevents the widespread adoption of such safety inventions in industry. The purpose of this paper is to develop a roadmap, integrating research design and known business techniques, for the commercialization of construction safety technology research. Significantly, this paper will help guide students and faculty on a path to apply leading safety research to the construction industry, potentially saving money and lives.

Keywords: Commercialization; Construction Safety; Entrepreneurship; Innovation; Research to Product; Technology Transfer.

1 Introduction

University research advances scientific discovery and generates a large amount of technological knowledge. Since the Bayh-Dole Act (1980), university patenting and licensing has significantly increased as a result of several trends. Yet, because university technologies tend to be very-early stage, only a relatively small fraction of university-generated knowledge results in commercially successful products. Unlike in industry where specific research is in the core expertise of the business, in case of universities, most new technology application will lie outside of university’s core expertise (teaching and research).

From the academic perspective, universities fulfill three missions: research, teaching, and service. Some argue that supporting entrepreneurship and new technology commercialization falls under the public-service mission and possible return on public research and development (R&D) money invested in fundamental research. However, a large challenge to this endeavour and fulfilment is the well-defined Innovation Gap [1, 2] at the university between the technical knowledge creation and a successful product in the marketplace.

Figure 1. The Innovation Gap between scientific knowledge and product development (after [1,2]).

In the area of construction health and safety research, there are many new exciting developments, including personal fall protection guardrail equipment [3], a Building Information Modelling (BIM) Occupational Safety and Health Administration (OSHA) rule-based safety checker [4], real-time pro-active personal protective equipment [5], and other advanced personal protection equipment or computational methods for construction health and safety. Historically, however, very few have been successfully commercialized. This can be in part attributed to lack of knowledge among researchers and the challenges of licensing and/or creating a startup. The goal of this work is to help shed
light and provide a structure and a guide on the commercialization process. It includes resources available, and several investigated case studies to provide a practical reference for Construction students, faculty, and researchers interested in bringing their construction health and safety inventions to market.

2 Background

2.1 Innovation and Commercialization

Innovation can be defined as proposed by Garcia [6] “Innovation is an iterative process initiated by the perception of a new market and/or new service opportunity for a technology based invention which leads to development, production, and marketing tasks striving for the commercial success of the invention.” The definition shows that innovation combines technical invention with market introduction and that the process is iterative, with subsequent innovation improving upon the prior, often as a result of market feedback. Technology commercialization overlaps with this definition as it is the process and a series of steps that move an invention from the early (idea) stages into a product or service. Similar definition is presented in CERF [7].

There are several models of technology commercialization: The Jolly Model with its sub-processes and bridging steps [8] the “Stage-Gate Process” [9] and other similar step-type processes used in technology management [10]. Overarching theme is that these models all begin with the basic scientific theory or newly discovered concept, continue through prototype and further development all the way to market introduction (and beyond). These general technology commercialization models also recognize several distinct phases in the process. Each approach may be more applicable to a different level of invention, technology area, or a technology development status.

“Successful” product in the marketplace should be self-sustaining (i.e. profitable enough to cover related manufacturing, distribution, and overhead) and also able to recover its commercial development costs.

The Jolly model also emphasizes four bridging steps between these phases that complete the process. Bridging steps have two conditions to allow the commercialization process to advance to the next phase: 1) demonstrating enough potential in the previous stage and 2) mobilizing sufficient (higher and higher) resources to proceed to the next stage, as the development costs increases significantly with each subsequent phase from Image to Sustain.

Literature recognizes two approaches to technology commercialization – Market Pull and Technology Push [10]. In Market Pull, a market problem/need exists and usually several competitors are trying to solve the problem to satisfy this need. Contrasting is Technology Push, where a capability given by a technology exists, but a solution, or a problem this capability can solve in the marketplace is yet unknown. A typical market pull can be often addressed by incremental (also continuous or sustaining) innovation, whereas a radical (discontinuous and disruptive) innovation will tend to be in the Technology Push framework. Despite this distinction, ultimately a successful commercialization will result in a market pull. As the authors argue further in this work, for sole university research, the early pull or existence of some customer needs is often first understood or discovered by the researchers in that area. Thus researchers in particular may be aware of opportunities on one hand, but also “blinded” by one area while dismissing other potentially more suited application area(s) for their technology.

Interesting for the case of health and safety in construction is the scarcely discussed type of “Regulatory Push” commercialization [10] common in some eco-innovations. In Regulatory Push, policymaker (or some governing body) decisions help effectively “push” certain technologies, standards, and solutions onto the market, often via subsidies or required/mandated standards.

2.2 The “Lean Startup”

There has been a new paradigm shift of how to create successful startup companies and launch new products to the market. Known as the “Lean Startup”, the process focuses on searching for a business model (and sometimes new customers or channels) by creating, testing, and revising hypotheses made about the products and customer problems [11]. Contrary to the traditional practice of writing a business plan, creating market forecasts, and assuming customer needs before launching the product, the lean startup focuses on discovering the customers’ problems right from the beginning, using feedback from potential users, purchasers, and partners to drive the path of the startup and product engineering. The purpose is to discover the minimum viable product (MVP), which only has the
minimum features that allow the product to be deployed to solve the problem. Blank and Dorf [12] explain how “customer discovery” is one of the most important aspects of the lean startup, which uses a “get out of the building” approach by physically going out to acquire feedback on all elements of the business model. “Lean Startup” acquiring empirical feedback from the marketplace has analogies in industry, where large firms have previously performed “Probe & Learn” or “Expeditionary Marketing” efforts [13]. For a startup, this “Probe & Learn” process means nimble budget and very strict time constraints, or eliminating waste as in “Lean Manufacturing”.

An important lean startup aspect is that startups are expected fail, and that quick and inexpensive failure accompanied by learning and potential pivoting moment as a result is seen as a success. Startups do not know the future, and therefore need to hypothesize. Lean Startups move from failure to failure, all the while adapting, iterating on, and improving their hypotheses as they continually learn from the customers [11, 12, 14]. This strategy greatly reduces the chances that startups will spend a lot of time and money launching products that no one actually will pay for [14] – recall that every later stage of product development becomes more and more expensive. Because they can move faster than established companies, startups can manage uncertainty, certain problems and ambiguities in the new product development process more effectively.

2.3 Research Design and Scientific Method

University researchers “live and breathe” research design and the scientific method. Creswell [15] outlines the necessary tools for proper research design. For the majority of research projects, the first step is to discover a problem or a question within the respected fields. Next is to conduct an extensive literature review of the problem, including what has been done, what needs to be done, and what could be done to solve the problem. During this review, the researcher narrows the scope of the problem and envisions possible solution(s) that can be applied to fix, mitigate, or eliminate the problem in question. The researcher may develop a technology or a new concept, or utilize/modify existing technology(ies) to apply towards the solution of the problem. Next, the researchers develop hypotheses, create experiments to test the hypotheses, and analyse the experimental data to support or refute the hypotheses.

In summary, proper and successful research design first starts with finding a problem, and then invents a technology or concept to provide a solution for that problem. This is the contrary to the idea of “Technology Push” – the practice of inventing or starting with a technology first and then trying to find a market to apply it to, which can often be difficult. Nevertheless, given the initial research problem, one could argue that there is no pure “Technology push” since the researches, even before embarking on a research project, already have some questions and problems in mind they are trying to resolve. Importantly, this early problem definition based on research problems, informs future market search. Significantly, these characteristics give researchers a greater advantage to commercialization than simply choosing a technology from an available list and pursuing an unknown area.

2.4 University Commercialization Funding, Support, and Infrastructure

There are various governmental and private programs that aim to aid the commercialization process at different stages, from early development all the way through capital investment. These programs can be beneficial for the commercialization of university research. Researchers should be aware of the many entrepreneurship and commercialization programs, supports, services, and even (non-dilutive) grant funding available on their university campus. The quantity and scope of these programs have grown dramatically in recent years. The National Science Foundation (NSF) Innovation Corps (I-Corps) was established to advance and commercialize NSF funded research. Some research universities may have established incubators, such as the Georgia Institute of Technology’s VentureLab and Advanced Technology Development Center (ATDC), which are programs that support the successful development of entrepreneurial companies through various resources and services. The Georgia Research Alliance (GRA) coordinates research efforts between Georgia's public and private sectors in order to expand research and commercialization capacity in Georgia’s universities.

At The University of Texas at Austin (UT-A) several programs exist in the engineering and business schools: The Idea to Product Program (I2P®) [1], Texas Venture Labs, and Master of Science in Technology Commercialization (MSTC) Program. Additionally Austin Technology Incubator (ATI), affiliated with UT-Austin (but somewhat separate) has a broad outreach and influence onto campus and to early-stage, student or faculty-run startups. ATI provides mentoring and guidance as well as expertise and access to outside entrepreneurship community. Such assistance programs are extremely beneficial, especially for university researchers, since they provide (free or affordable) critical resources, such as funding network or strategic guidance, which may be difficult for researchers to obtain initially.
3 Purpose, Scope, and Objectives

The purpose of this paper is to develop a general roadmap with guidelines for the commercialization of construction safety technology research. For this paper, the term “technology research” is a catch all, which may include an idea/solution, metric/rule, methodology, software program, or a tangible device. Because not every technology is the same, each may require different care, strategies, and decision making. Although there is no “perfect recipe”, this work highlights general stages and steps. The scope is aimed at construction safety technology research and will draw from case studies in this field. This paper is focusing on the incubation period of commercialization, from the start of an idea through taking it out of the university’s door into industry. Outlining details about taking the startup through major expansion, execution of a business plan, and growing to a successful business will not be fully covered in this paper. Additionally, legal aspects such as intellectual property (IP) are covered elsewhere in more detail [16]. Patents do underline successful technology commercialization and help maintain competitive advantage.

Academics and researchers may be uncomfortable interacting with people outside their fields. However, the real world and successful business requires frequent interaction, especially with potential partners, financiers, and - most importantly - customers. This research paper aims to encourage and show scientific researchers that it is possible to commercialize technology and to create startups even without an extensive business background on their part. This work presents the integration of research design into current business commercialization techniques, giving researchers a guide for the commercialization process. However, some basic knowledge of business and IP is recommended.

4 Methodology

The methodology of this research is to first conduct an extensive literature review of current and past practices of commercialization to review the similarities between the options and highlight the benefits. Since researches excel at the research design and scientific method, it will be advantage to incorporate and apply these skills to the commercialization process. Case studies within the field of construction safety and health will be investigated. Additionally, a six month case study of the process of creating a startup company will also be examined.

Commercialization covers a broad scope of activities with varying definitions. For this paper, we are defining commercialization as the passing of university research from academia into the industry/marketplace. Three pathways that are identified for commercialization or dissemination of research are:

1. Inventors and founders form a startup
2. Company/industry buys/leases the technology rights (exclusive or non-exclusive) or IP
3. Open source (free-sharing and disclosure).

The first channel is when a startup is formed within the university. There are university programs that may assist (i.e., NSF I-Corps) in the formation of a startup to get the technology into the industry. The university may retain ownership stake of the startup, give/sell the ownership of the startup or another company outside of the university, or may claim royalties based on product revenues. These arrangements are usually negotiated with the Technology Transfer Office separately for each deal. The second path is straight forward, in which a company may invest or purchase research from a university to obtain technology IP rights to continue use or further expand. The third channel, often a strategic choice, is to make the research invention open source, which allows universal access to the research, invention, or other IP. However, there are still ways to make a profitable business model on an open source invention.

5 Case Study and Preliminary Results

A research project at the Georgia Institute of Technology has been selected to determine the feasibility of it to be commercialized. The research topic involves the integration of Building Information Modelling (BIM) for construction safety [4]. The process of creating a startup will be documented and the lessons learned will be discussed. In November 2013, the project went through the Gauntlet at the Georgia Institute of Technology, which is a smaller version of the I-Corps. During an intensive four week period, the group had to 1) identify the core of the business model; 2) create hypotheses about market problems, 3) interact with potential consumers to discover the market, 4) analyse results and update the business model. The purpose of the Gauntlet was to determine whether or not there is a product-market (techno-market) fit.

The first step of the project was to determine the marketability of the BIM safety rule checker, by which the group went “out of the building” to interact with potential users and customers in the construction industry. The purpose was to discover current problems companies have with BIM and safety. A total of 35 interviews were conducted within a 4 week period. The interviews were meant to gain insight, and not necessarily to be scientifically sound. The research ultimately applies to construction workers, and thus they were interviewed first, followed by construction safety researchers (due to their knowledge of construction safety). Afterwards, subcontractors and general
The following are the hypotheses that were tested with the conclusions made based on this primary research.

**Hypothesis #1:** Workers adhere to safety rules and standards for their protection. Workers will not always follow safety rules if it impedes with the task. Some admitted to not know exactly what the safe procedures are for the task. Getting the work done is the worker’s priority, and not necessarily safety (might be secondary). However, for larger companies, workers tend to look out for one another’s safety. Safety is important, but sometimes the reward is smaller than the risk.

**Hypothesis #2:** Preventative safety guidelines and leading indicators will help increase worker safety awareness. Research has shown a decrease of reported incidents with the increase of safety programs. Larger companies incorporate mandatory safety training programs and have reported great success in the increase of worker safety awareness.

**Hypothesis #3:** Construction companies want to increase worker safety and safe practices. Worker safety is the number one goal for construction companies. They use any resource available to them to improve safety standards.

**Hypothesis #4:** Construction companies have knowledge about BIM and are (or plan) using it. Most companies have knowledge of and use BIM. However, in many cases the companies neither use its full potential nor use it for the entire project duration. They are mostly using it for clash detections and inspections of installation. There is no record of using BIM for automatic safety code compliance.

**Hypothesis #5:** Companies will automate safety inspection. As safety is the main priority for companies, keeping workers safe while reducing cost is even better. Manual inspection takes much time, and thus companies are willing to invest in new technology to help automate it. Currently, some companies use software programs that let them leave inspection notes right into the BIM model, notifying personnel to fix the unsafe situation.

**Hypothesis #6:** Companies will invest in safety technology. As companies want to maximize safety, they also want to minimize cost. Companies are willing to invest in new technology that will increase technology and decrease costs. Safety technology has been proven to be successful in the construction industry. However, only larger companies are considerably willing to invest in new safety technology since they have more capital and can afford to recuperate losses. More companies are willing to invest if the cost of the technology is low and will reduce costs. Additional lessons learned include:

- Large companies would rather pay a third party for a technology and invest in a proven technology (or assist university research) than developing it in-house.
- Companies would pay for a technology based on the project rather than license or purchase. The price of the technology for a project would be included in that project’s budget.
- There are many unsolved problems with BIM that companies come across. Some are in early development with BIM and some relate to all aspects of the project (incl. logistics, laydown yards, storage, cranes, safety).
- The size of the project affects the safety procedures. Companies typically have a single safety inspector that inspects multiple sites. However, one company stated that large projects over $25 million would have a full time safety inspector. This fact inquires for additional research to investigate the differences between safety and incident rates on projects having a full time safety inspector versus a part time safety inspector.

Note that the interview population was not large enough for a full scientific validation or extrapolation for the entire industry. However, the results shed great light and provide direction for the startup. Again, the purpose of the Gauntlet was not to validate the research, but to give a binary “go/no go” status for the startup. In conclusion, it was determined that the research does have a product-market fit, and will pursue the next stages of the startup process.

### 6 Proposed Roadmap

The following is a proposed roadmap, which is a set of guidelines based on literature, empirical data, experience, and case studies. These guidelines are intended for university researches with a background in proper research design. Technology commercialization and reviewing an early-stage technology does not always result in a binary “yes/no” decision. Often, the recommendation based on market feedback may be that of “not right now”, “not here”, or “not at this price”. Moving to the next phase will be also nuanced (i.e., yes if XYZ, no if ABC, yes in X months, no if this experiment fails). These steps are intended to be iterative, and thus they will be routinely repeated, each time resulting in refined business model to discover the features of minimal viable product (MVP).
6.1 Step I: Identifying the Problem

6.1.1 Utilize Research Design and Background

It is important to have a working knowledge of proper research design. The first step is to have a basic understanding of the problem being researched in the current field. The further along the research is to solving the problem with the technology the better. Here, the key is to not only understand how technology performs the function, but be able to describe what the technology does and the benefits it provides (often those benefits may be outside of first envisioned area). Similarly, it is important to know the problem that exists in the industry that a technology can (at least potentially) solve. When performing background research, one should also consider searching patent databases. Create testable hypotheses about the problems, including the root causes and possible solutions.

6.1.2 Learn How the World Works

Test your hypotheses. Testing the research hypotheses allows researcher to learn how the world works. Experiments, case studies, and analysis do validate research hypotheses, but learning how the world around the technology works determines a greater need for the technology. If the technology is proven to work, there is still a need to determine whether it is commercializable, and which features of the technology (MVP) are important to customers (the MVP) as benefits: e.g. the customer for a car may really be interested in moving herself from A to B (need) rather than possessing a metal box on wheels (one of many ways to satisfy the need). When a benefit is generalized along these lines, it becomes easier to address (in this case a train, airplane, bicycle, or feet can satisfy this customer need). The major influence of whether or not a technology is commercializable is whether there is a market for it. In order for a market to exist, the technology needs to solve an important enough problem (even a latent need, as is often the case in Technology push).

6.1.3 Interview Potential Customers/Stakeholders

Potential customer feedback and primary research is critical. There are numerous ways to talk to customers to gain valuable feedback. One must also gain an understanding of the buying process – who makes the purchase decision, who is the user, and who has the money and “writes the check” often these are, all different. The “end user” of the technology, which is the one who will either use the technology or benefit the greatest from the technology, should already be generally determined based from prior research. In terms of safety research, one can go to a location where the users would be, such as jobsites, factories, or offices, in order to talk to workers, project managers, etc. Use any resource available to contact customers, including current contacts, social media and blogs, and the phone book. Phone interviews and be a valuable and efficient resource. Trade associations could be an excellent information source for industry contacts or experts.

6.1.4 Ask Questions

Proper discovery of how the world works starts with asking (open-ended) questions. Communicate effectively by taking interests in what problems the customers have to offer. It is important that these meetings are for “discovery” and not “selling” a product.

6.1.5 Evaluate Scientific and Research Hypotheses About the Technology Solution

In many ways, refuting a hypothesis and deciding not to proceed at a step can be very valuable not only to preserve vital resources, but also to inform the market and startup. After gaining feedback, re-evaluate the hypotheses about the problems, customers, markets and any unexpected discoveries. Make shifts in hypothesis to uncover the greater problems. Iterate until the MVP is discovered.

6.2 Step II: Determine the Target Market

Once the MVP is discovered, the next step is to investigate the overall market. It may be better to start small and then expand. To conduct the secondary and primary market research, utilize the same steps and techniques in Step 1, but this time focusing on the market information.

6.2.1 Secondary Market Research

Performing secondary market research will review reports, industry and market analysis/statistics, trade journals, and information available through business databases, including public companies filings or news releases to gather relevant information. Secondary research can help identify major players in industry, industry trends, as well as help estimate the possible market size for proposed product or service. In particular, initiating coverage reports, or S-1 filings of companies required by SEC before IPO include valuable market information in the intended space (the difficult part is finding a good, recently IPO-ed proxy-company).

6.2.2 Primary Market Research

This includes talking to potential customers, experts in the field, competitors, potential partners, other researchers, and overall gathers “primary” source
information. This step overlaps with customer interviews, but here the focus is on understanding the market and space better at a higher level. Although this information is very valuable, it is good to cross-check previous data with interviewees, and make sure they converge or agree. Approximately 8-10 generally converging interviews will provide a good initial assessment. Again, excellent resources for market information are trade associations, as part of their mission is to aggregate and disseminate information about a very specific area.

6.3 Step III: Determine the Channel to Commercialize

It is important to determine what path to use to move the research outside of a university. The channel depends on the type of research, and each has benefits and drawbacks. Unless technology is straightforward fit, large companies purchase more developed third party tech (or smaller companies) that are already de-risked, even though it demands paying a premium – for a defined market, customer need, and a working solutions (and sometimes people talent).

6.3.1 Create Startup

Creating a startup is beneficial since the researches have been the motivation to the commercialization. Researchers take their vision and form a company to solve the problem. This is mostly suited for a tangible technology or invention (such as new safety equipment) but also for early stage inventions that have a limited chance of leaving the university and for success otherwise (not easy to license by a large company at this point, nor amenable to becoming public domain).

6.3.2 License

Licensing research (granting permission at a fee) may be suitable for certain types of research, such as in the field of chemistry (chemical patents) or computer science (software code or programs), where the technology may not be as complex as to require the inventor expertise on the team using it. It will also have to face only minimal implementation challenges into existing processes. An example in construction safety may include a plug-in software program that detects or analyses for safety issues.

6.3.3 Create Open Source (free-share)

Open source is the publishing of the research to grant free access. When there is a limited market to make profit (or there is not much need), or inventors intend to share their invention for public good, or invention can become a de-facto standard, this may be a preferred path. Data, source code, algorithms are popular candidates for open source. This can be a strategic choice when inventors chose not to pursue a technology, yet do not wish anyone else to have an exclusivity on it. Examples include algorithms, data, or source code.

6.4 Step IV: Determine a Business and Revenue Model

This step mostly applies to startups. The key to every successful business is revenue (and therefore profits). There are many potential business models based on any given technology. A startup will need to carefully consider its options and focus on the key expertise areas (or the most-value-added, profitable) functions in the entire product cycle and ecosystem. Other parts of the business (e.g., assembly and manufacturing, packaging, distribution) may best be outsourced. Here, strategic partnerships may play an important role not only in developing the technology but also in terms of channel-expertise. While commercializing a technology, this needs to be taken into account. Of course, all the while startup develops additional prototypes, focuses on benefits and tests or releases a minimally viable product (MVP).

6.5 Step V: Explore Channels to Receive Funding or Support

There are various governmental and private programs that help aid researchers in the commercialization process to bridge the early funding gap between research funding and a saleable product (customer or investor funding). Many universities now have programs for early stage development. Further grants may include various US Departments’ Small Business Innovation Research (SBIR), emerging technology funds (ETFs), and other state, government, and industry consortia incentives. Consider Strategic partnerships to help in development activities, especially between R&D and market, as well as customer channels where funding sources may be limited.

6.6 Step VI: Strategic Choices/Due Diligence with Legal and Financial Aspects

Each type of research, technology, market or customer, will vary. There is no a one approach-fits-all. The following list is outside of the work scope of this paper, but is important to keep in mind when conducting due diligence with legal and financial aspects:

- Intellectual search
- Projected expenses and financials
- Determine scope of IP and IP strategy overall
7 Conclusion

The proposed roadmap was based on an extensive literature review, empirical data, experience, and case studies. The guidelines need to be tested and validated with future case studies. The BIM and safety startup will continue to be examined as it makes progress towards commercialization. The failures will be recorded and changes will be made accordingly as it continues through the commercialization process.

Important research questions to explore throughout this research and address include when one should step to the next stage, when to iterate, and when to abandon a technology for a given market or application. Additional research questions discovered are:

1. What metrics should be defined to determine when to go on to the next step?
2. Is market size/growth sufficient?
3. Is there sufficient market interest from potential customers and at what levels? (i.e. can it be filled with the technology product/service?)
4. How will the product transition from early users (evangelists/lead users) to the early majority?

References


Exploring Local Feature Descriptors for Construction Site Video Stabilization

Jung Yeol Kim\textsuperscript{a} and Carlos H. Caldas\textsuperscript{b}

\textsuperscript{a} PhD Candidate, Dept. of Civil, Architectural and Environmental Engineering, Univ. of Texas at Austin, USA
\textsuperscript{b} Associate Professor, Dept. of Civil, Architectural and Environmental Engineering, Univ. of Texas at Austin, USA
E-mail: jungyeol.kim@utexas.edu, caldas@mail.utexas.edu

Abstract - Recent studies on automated activity analysis have adopted construction videos as an input data source to recognize and categorize construction workers’ actions. To ensure the representativeness of its analysis results, these videos have to be gathered randomly in terms of time and location. In doing so, such videos must be taken with hand-held cameras, a fact that inevitably leads to videos including jittery frames. Such frames can decrease the accuracy of automated activity analysis results. One area of the most recent and effective action recognition methods involves using spatio-temporal action recognition algorithms. The jittery frames, however, are fatal to the recognizing of a human worker’s action using such an algorithm. Jitters can be removed from the videos by using video stabilization technologies. The video stabilization is the pre-processing of action recognition for automated activity analysis. Regarding the video stabilization, local feature descriptor plays a major role in the stabilization process, and the correct selection of proper descriptor is critical. Therefore, the purpose of this study is to identify the best local feature descriptor for the video stabilization. This paper describes detail steps of the stabilization and provides performance analysis of various local feature descriptors in terms of stabilization of videos from construction site.

Keywords - IT applications, video stabilization, video interpretation, activity analysis, work sampling, productivity measurement, action recognition

1 Introduction

Videotaping has long history in gathering on-site data for construction productivity analysis \cite{1, 2, 3, 4, 5}. Recent years, due to the advancement of computer vision technologies, many researchers have studied on automated interpretation methodologies of construction site videos\cite{6, 7, 8, 9, 10, 11, 12, 13, 14}. One of the construction industry’s computer vision application areas is activity analysis. Activity analysis is a continuous measurement and improvement process that helps craft workers increase their time spent on actual construction work. It includes the application of work sampling in its measurement process and requires manual observations of workers \cite{39}. The main focus of the computer vision application in the activity analysis is to substitute, in construction videos, the manual observation of construction workers’ actions with automatic recognition and categorization \cite{9, 12, 18}.

To ensure the representativeness of its analysis results at the site, the construction videos have to be gathered randomly in terms of time and location \cite{39}. The well-planned combination of hand-held and fixed closed-circuit television (CCTV) cameras can be a solution to obtaining those random videos. CCTV camera is a convenient tool to obtain videos at random intervals but has a limitation regarding random locations. It cannot cover all the areas of the construction site. A hand-held camera is useful in gathering those videos at random time intervals and places, but such videos inevitably include jittery frames. Jitters in those videos can decrease the accuracy of automated activity analysis results. One area of the most recent and effective action-recognition methods uses spatio-temporal action-recognition algorithms \cite{22}. The jittery frames, however, are fatal to the recognizing of a human worker’s action when such an algorithm is being used; the jitters in the videos can distort the spatio-temporal volumes, trajectories, or features. Those jitters can be removed from the videos by using video stabilization technologies. The video stabilization is the pre-processing of action recognition for automated activity analysis. Regarding the video stabilization, local feature descriptor is one of the most important elements. Therefore, the purpose of this study is to identify the best local feature descriptor for the video stabilization. This paper describes detail steps of the stabilization and provides performance analysis of various local feature descriptors in terms of stabilization of videos from construction site.
2 Related work

2.1 Automated Action Recognitions for Activity Analysis

In the automation of activity analysis, researchers have studied three types of action recognition technologies: (1) sensor-based [15][16]; (2) 2D image/video-based [6][9][11][12][13]; and 3D vision data- (i.e., depth image, point clouds, and human skeleton) based ([17][19][18]) action recognition. All these approaches have contributed to the automation of activity analysis. However, the 2D and 3D vision-based approaches assume that their inputs are static. Most of the studies use vision data from fixed 2D or 3D imaging sensors. In actual situations, some of the data have to be gathered by hand-held imaging sensors, where jitters are unavoidable. Therefore, it is necessary to adopt the stabilization technologies of the data. This paper focuses on the stabilization of 2D video data.

2.2 Video stabilization

Video stabilization falls into two types: hardware-based during recording and software-based, post-processing digital video stabilization [35][21]. Hardware-based stabilizers consist of complex and expensive sensors and lens systems to reduce the movement of cameras. Cheaper cameras also adopt sensors and firmware to offset camera motions. However, these hardware-based systems fail to provide sufficient stabilization function to compensate for complex camera motions and severe jerking. Therefore, to obtain stable videos, post-processing video stabilization is still required [35]. The Post-processing digital video stabilization is defined as “the process of removing the unwanted motion from input video sequence by appropriately warping the images” [37]. It is not a real-time solution but can be applied to the videos taken by any type of cheap hand-held cameras. This paper focuses on the software-based post-processing digital video stabilization.

Software-based video stabilization (hereafter “video stabilization”) can be divided into two types: (1) 2D and (2) 3D video stabilization [20]. A general 2D video stabilization method is composed of the three steps as shown in Figure 1: 1) motion estimation, 2) motion compensation, and 3) image composition [36][23][35]. Motion estimation means the estimation of motion between two sequential frames (i.e., motion between the previous and current frames). Motion compensation provides the computation of global transformation to stabilize the current frame. Based on the transformation, image composition warps the current image. Recently, more innovative approaches have been introduced such as very stable and anti-distortional.

Figure 1. General video stabilization method [36]

3D video stabilization does not mean the stabilization of 3D vision data. It is for 2D videos, though it uses the estimation of 3D model of input camera motion and scene. It also use image-based rendering techniques to render new frames based on the estimated camera motion path. The new rendered frames are frames of video stabilized [24][25][26]. One interesting study that used this method is a content-preserving warping carried out by Liu et al. [20]. It distinguishes itself from other methods by not having a blank area on the stabilized video.

The 2D video stabilization is limited regarding significant depth variations but is still a simple, robust, and efficient solution [35][20]. The 3D video stabilization could overcome the depth variation problem but is more complex and often depends on unreliable depth estimation [35]. Furthermore, the authors assumed that a cameraman does not walk when taking videos, and those videos consequently have less depth variations. Therefore, this paper focuses on the 2D video stabilization methods instead of 3D based methods.

3 Our Approach

3.1 Overview

Our approach, shown in Figure 2, is a variation of the 2D-based general video stabilization method. It consists of five steps with details given in the following paragraphs.

The first step is to extract local feature descriptors from the first and second frames. Again, these descriptors are one of the most important elements of the video stabilization method. They are used for the estimation of geometrical transform for stabilization. The geometrical transform is estimated by the matched descriptors of sequential frames.

In this paper, the authors selected the following four descriptors: (1) Scale-Invariant Feature Transform (SIFT) [28]; (2) Speeded Up Robust Features (SURF) [29]; (3) Fast Retina Keypoint (FREAK) [30]; and (4) Oriented FAST and Rotated BRIEF (ORB) [31]. The authors selected these features because SIFT is well known for its scale and rotation invariant performance [32]; SURF is inspired by SIFT but is known for its...
higher detection speed and better performance. FREAK is a newer descriptor and shows the faster detection speed and better robustness than SIFT and SURF according to its inventor’s experiments [30]. ORB is a combination of FAST (Features from Accelerated Segment Test) corner detector [42][43] and BRIEF (Binary Robust Independent Elementary Features) descriptors. Rublee et al [31], the inventor of ORB, insisted that ORB outperforms SURF and SIFT. The experiment’s results regarding the stabilization performance by these feature descriptors will be described in the next section of this paper.

Figure 2. Our construction video stabilization approach

The second step is to recognize construction workers in each frame using a human-detection algorithm, histograms of oriented gradients (HOG) [33] and to remove unnecessary local descriptors detected within the workers’ regions in each frame. Figure 3 shows the example descriptors detected in the worker’s regions. Those descriptors can be sources of error during the estimation of geometrical transformation; indeed, the directions of workers’ movements (trajectories) can differ from the camera’s jittering directions. Figure 4 shows the matched local feature descriptors outside of the worker’s regions in the sequential frames. In this case, the estimation of the geometrical transform in the next step will be incorrect [27].

Our approach differs from of Wang and Schmid [27] by eliminating SURF descriptors before identifying matched descriptors. They simply selected SURF descriptor and motion vectors to estimate homography between two consecutive frames and eliminate matched descriptors in the people’s region. Our approach is simpler but effective because there are still sufficient amount of descriptors outside of the worker’s regions that enable the estimating of the geographical transform. Importance to this step is the accurate recognition of human workers.

Figure 3. Local feature descriptors detected in the worker’s regions (SURF Descriptor Used)

Figure 4. Matched Local Feature Descriptors outside of the Worker’s Regions in the Sequential Frames (SURF Descriptor Used, Top 150 Matches Displayed out of around 1,500 Matches)

The authors follow the general process (Figure 1) for the remaining steps [23][34][35]. The third step is to compare the remaining descriptors of the two frames to discover corresponding points. To match the corresponding descriptors, this study used the Nearest Neighbor Ratio for floating point descriptors and Hamming distances for binary descriptors. The fourth step is to estimate the geometrical transform with the corresponding points. Affine transform was used in this study and was compensated. The last step is to warp the second frame with the geometrical transformation and repeat these steps.
4 Experiments and Results

The authors compared the stabilization performance with four feature descriptors. The performance will vary according to the descriptors because they each have a different ability to discover corresponding points with jittery frames. The jittery frames include horizontal movement, vertical movement, rotation, and the combination of all the jitters. The authors used OpenCV, VL-FEAT, C++, and Matlab for the experiments. The videos were gathered from a commercial building, road resurfacing, and building exterior remodelling sites with cheap hand-held camera. The resolution and frame per second (fps) of the video used for this experiment were 573 × 320 and 5 fps. The average computation time per frame is 0.58 seconds. Figure 5 shows the experiment’s results. From the first to the fourth rows correspond to the videos from the commercial building, road resurfacing (the second and the third rows), and building exterior remodelling sites.

Each column of the figure corresponds to the video stabilization results with SIFT, SURF, FREAK, and ORB descriptors. Each image is an overlaid image from the first stabilized frame to the last. Therefore, sharpness can be a proper metric to compare their performances. The sharper the image the better the stabilization performance; it means that objects in each stabilized image are at similar locations. The yellow boxes in each image are the sharpest parts. Figure 5 shows that the overlaid images stabilized with SURF descriptor have the highest sharpness. To the naked eye, however, it is hard to distinguish the relative sharpness of the images. Therefore, the authors measured the sharpness of each overlaid image. There are many methods to estimate the sharpness of images. The authors adopted the Brenner gradient based sharpness measurement method because it is more sensitive to the sharpness changes than other methods [44]. The measurement results are shown in Figure 6. The higher number means a higher sharpness.

Figure 6 shows that SURF always outperforms other descriptors. SIFT follows SURF, and FREAK and ORB demonstrate worse performances. Sometimes, the FREAK descriptor fails to find matched descriptors between two sequent frame images. In terms of computation time, stabilization with FREAK generally showed the shortest computation time, while SIFT showed the longest.

Based on the experiments, it can be concluded that SURF and SIFT are robust to stabilize the jittery videos due to their rotation-invariant characteristics. Therefore, the authors selected SURF for the best descriptor for our construction video stabilization.

![Figure 5. Construction video stabilization results](image_url)
Figure 6. Stabilization performance comparisons by different descriptors.

Figure 7 shows an example of original jittery frames and stabilized frames. Figure 7 (a) and (b) are the 1st and 19th frames of the original video. The long solid line is the basis point of the first image and the dotted line is the vertical difference between the two frames. Otherwise, there is a really small amount of vertical difference in the corresponding stabilized frames (Figure 7 (c) and (d)).

5 Conclusions

The authors have explained a modified video stabilization method for the activity analysis while considering construction workers in videos. The authors also performed an experiment to find out the best descriptor with the video stabilization method. The authors used the sharpness of overlaid stabilized images as a metric to measure stabilization performance. In the experiment, SURF descriptor performed best, followed by SIFT descriptor.

This video stabilization method could pave the way for the use, at a construction jobsite, of cheap hand-held cameras and any other mobile video-recording devices, such as Gopro®, Looxcie, and Google Glass. This would mean that it could become easy, and with less expense, to gather random videos for automated activity analysis.

This study has few limitations. Human detection, a part of the second step to eliminating unnecessary descriptors, needs to be improved in the future. The HOG based human-detection algorithm used in the step has some level of Type-I (False Positive) and Type-II (False Negative) errors. Furthermore, it tends to better detect upright position than other poses. Their effects were not considered because it is not the scope of this study, but the authors believe that using the human detector still could reduce the chance of errors as it stand. Another limitation of this method is that the experiment is performed with only four descriptors and a small number of videos. Experiments involving more descriptors and a greater number of videos need to be performed in the future.

References


A Novel Inference Model for Post-Earthquake Bridge Safety and Failure Probabilities Prediction—a Case Study in Taiwan

Min-Yuan Cheng, Yu-Wei Wu, Yung-Fang Chiu, Yu-Chen Ou and Chien-Kuo Chiu

Abstract - Bridges are a vital and significant component of Taiwan’s transportation infrastructure. Therefore, regular and comprehensive inspections of existing bridges are necessary to prevent damage and traffic disruption and reduce earthquake-related damage and casualties. However, due to the large number of bridges in Taiwan, the time and budget required to perform traditional structural analyses (preliminary assessment, detailed analysis) on every bridge to calculate yield acceleration (Ay) and collapse acceleration (Ac) values make doing so impractical. This paper integrates material degradation, pushover analysis, and artificial intelligence to create a new inference model as an alternative to traditional structural analysis. Historical cases are used to infer Ay and Ac values by mapping relationships between the preliminary assessment factors (input) of historical cases and detailed assessments of Ay and Ac values (output). Using the proposed inference model to predict Ay and Ac values, bridge maintenance planners can quickly and more cost effectively assess bridge earthquake damage probabilities as a guide to identifying priority bridge maintenance projects.

Keywords - Seismic Assessment; Deterioration of Materials; Seismic Capacity; Evolutionary Support Vector Machine Inference Model

1 Introduction

Besides the functionality of crossing rivers and valleys to communicate with the outside world, bridges also play an important role in maintaining the economic arteries. In sparsely populated remote areas, bridges have become an important lifeline of living supplies and product output. Once bridges are damaged by natural disasters, it will generate far-reaching impacts on communication and transportation. It is also imaginable that life and property will be endangered, and social economy will be affected as well. At present, highway bridges in Taiwan are aged from newly built ones to the oldest ones aged more than eighty years. Accordingly, due to different seismic design specifications, as well as deterioration caused by environmental impact, some bridges do not meet the current seismic requirements and may result in damage.

Taiwan reinforced concrete bridges are the most widely used type (95%), and 75% bridge are used more than 20 years [1]. The old bridge will not meet the needs of seismic norms issued by the circumstances which led to destruction because different seismic design used or deterioration phenomenon caused by environmental influences. In other words, the current problems faced by the majority of Taiwan’s bridges are old and deterioration. Taiwan belongs to many earthquakes, once large-scale natural disasters such as earthquakes, high probability of a bridge collapse or breakage. To avoid disaster, causing damage to the bridge leading to traffic disruption, or even residents trapped casualties and other incidents occur, the existing bridge is imperative to conduct a comprehensive inspection.

At present, the traditional structural analysis can be divided into three types, including: 1. brief investigation; 2. preliminary assessment; and 3. detailed assessment. The brief investigation table thus developed is mainly for relevant management personnel to identify buildings with seismic capacity-related problems. Professionals will then perform preliminary assessment of such problematic buildings. Regarding buildings of seismic capacity concerns after the brief investigation, civil engineers are hired to complete the preliminary assessment table to assess the buildings. Such investigation and assessment will result in large sum of precious data relating to the seismic capacity of buildings. Finally, according to the collected bridge data, professional and complex assessment methods, such as the pushover method, are applied to carry out the detailed assessment to obtain the highly accurate yielding acceleration (Ay) and complete damage...
However, there are more than forty thousand bridges in Taiwan. When applying the simple assessment method to carry out visual investigation, the results will be not accurate despite the fast speed of investigation. In the case of applying the detailed assessment, although the results are relatively accurate, it takes much more time and costs. In addition, detailed assessment can only be done by experienced professionals. To conduct detailed structural analysis of each bridge with limited funds and professional labor is impossible. If simple assessment factors and the mapping relation between Ay and Ac for detailed assessment of similar bridges can be identified to infer the Ay and Ac values of other bridges, it will save a great deal of manpower when acquiring Ay and Ac values within the tolerable error range.

However, due to the large number of bridges in Taiwan, the time and budget required to perform traditional structural analyses (preliminary assessment, detailed analysis) on every bridge to calculate yield acceleration (Ay) and collapse acceleration (Ac) values make doing so impractical. This paper integrates material degradation, pushover analysis, and artificial intelligence to create a new inference model as an alternative to traditional structural analysis. Historical cases are used to infer Ay and Ac values by mapping relationships between the preliminary assessment factors (input) of historical cases and detailed assessments of Ay and Ac values (output). Using the proposed inference model to predict Ay and Ac values, bridge maintenance planners can quickly and more cost effectively assess bridge earthquake damage probabilities as a guide to identifying priority bridge maintenance projects.

## 2 Literature review

### 2.1 Seismic assessment of bridge

The static pushover analysis is becoming a popular tool for seismic performance evaluation of existing and new structures. The expectation is that the pushover analysis will provide adequate information on seismic demands imposed by the design ground motion on the structural system and its components [2]. Nonlinear static (pushover) analysis has become a popular tool during the last decade for the seismic assessment of buildings or bridges [3].

The static pushover analysis has no rigorous theoretical foundation. It is based on the assumption that the response of the structure can be related to the response of an equivalent single degree-of-freedom (SDOF) system. This implies that the response is controlled by a single mode, and that the shape of this mode remains constant throughout the time history response. Clearly, both assumptions are incorrect, but pilot studies carried out by several investigators have indicated that these assumptions lead to rather good predictions of the maximum seismic response of multi degree-of-freedom (MDOF) structures, provided their response is dominated by a single mode [2].

The purpose of the pushover analysis is to evaluate the expected performance of a structural system by estimating its strength and deformation demands in design earthquakes by means of a static inelastic analysis, and comparing these demands to available capacities at the performance levels of interest [2]. Large-scale simulations of transportation networks in urban regions have been the focus of numerous projects, such as HAZUS (1999). The goal of such simulations is to provide an economic impact analysis of damage caused by an earthquake event to a transportation network. Direct economic impact can be defined explicitly in terms of monetary losses due to damage to bridges and loss of function. Indirect damage can be defined in terms of damaged and failed links in the transportation network and link costs associated with traffic flow reduction, rerouting, and time delays that cause interruptions in the flow of goods and services.

Damage to bridges in the highway network contributes significantly to both direct and indirect losses. Bridge damage fragility curves, sometimes termed vulnerability curves, describe the conditional probability of exceeding a level of direct or indirect bridge damage, given a level of seismic hazard. The formulation of bridge fragility curves has historically transitioned from empirical to analytical methods [4].

### 2.2 Evolutionary Support Vector Machine Inference Model (ESIM)

Support vector machines and fast messy genetic algorithms represent recently developed AI paradigms. SVM was first suggested by Vapnik [5] and have recently been applied to a range of problems that include pattern recognition, bioinformatics, and text categorization. SVM classifies data with different class labels by determining a set of support vectors that are members of the set of training inputs that outline a hyper plane in a feature space. It provides a generic mechanism that uses a kernel function to fit the hyper plane surface to training data. The user may select the SVM kernel function (e.g. linear, polynomial, or sigmoid) during the training process, which identifies support vectors along the function surface. Using SVM presents users with the problem of how to set optimal kernel parameters. Therefore, SVM parameters must be obtained simultaneously. Proper parameter settings can improve SVM prediction accuracy, with parameters to
be optimized including penalty parameter C and kernel function parameters such as the gamma of the radial basis function (RBF) kernel. In designing a SVM, one must choose a kernel function, set kernel parameters and determine a soft margin constant C (penalty parameter). The Grid algorithm is an alternative to finding the best C and gamma when using the RBF kernel function. However, this method is time consuming and does not perform well [6]. Fast messy genetic algorithms (fmGA) were developed by Goldberg et al [7]. Unlike the well-known simple genetic algorithm (sGA), which uses fixed length strings to represent possible solutions, fmGA applies messy chromosomes to form strings of various lengths. Its ability to identify efficiently optimal solutions for large-scale permutation problems gives fmGA the potential to generate SVM parameters C and gamma simultaneously. Considering the characteristics and merits of each, this paper combines the two to propose the Evolutionary Support Vector Machine Inference Model (ESIM).

The ESIM used here was developed by Cheng and Wu [8]. In the ESIM, the SVM is employed primarily to address learning and curve fitting, while fmGA addresses optimization. This model was developed to achieve the fittest C and gamma parameters with minimal prediction error. The structure of ESIM is shown in Figure 1.

3 Post-Earthquake Bridge Safety Assessment using Failure Probabilities Inference Model

The primary purpose of this study was to develop an earthquake seismic assessment of bridge diagnostic prediction model.

3.1 Historical cases collection

This paper adopted 24 RC bridges in Taiwan (as shown in Table 1).

<table>
<thead>
<tr>
<th>No.</th>
<th>Length (m)</th>
<th>Design year</th>
<th>Design accelerate(g)</th>
<th>Structure type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>1992</td>
<td>0.139</td>
<td>I beam</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>1992</td>
<td>0.139</td>
<td>T beam</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>150</td>
<td>1990</td>
<td>0.187</td>
<td>I beam</td>
</tr>
</tbody>
</table>

3.2 Material deterioration

The seismic performance of bridges can be evaluated using the capacity spectrum method suggested by ATC-40. This paper presents a model of deterioration due to carbonation taking into consideration uncertainty factors to estimate the initiation and the rate of corrosion and to analyze the structural capacity and serviceability of bridge. Then, it goes on to propose a method for evaluating the failure and severe cracking probability during earthquakes and the deterioration risk of members in specified years from construction (Figure 3).
3.3 Pushover analysis

This paper considers material corrosion level (0%, 5%, 10%, 20%, 35%; five degree corrosion) and bridge assessment analysis are applied to carry out the detailed assessment to obtain the yielding acceleration (Ay) and Damage acceleration (Ac). In the figure 4, bridges PGA with 5 corrosion levels are draw.

3.4 Pushover analysis training database

After pushover analysis, 24 bridges detailed assessment are collected. This study analysed 5 degree corrosion levels (0%, 5%, 10%, 20%, 35%) for every bridge. Therefore, 120 training data are established in the Table 2.

3.5 AI seismic assessment prediction model

At present, the traditional structural analysis can be divided into three types, including: 1. brief investigation; 2. preliminary assessment; and 3. detailed assessment. The brief investigation table thus developed is mainly for relevant management personnel to identify buildings with seismic capacity-related problems. Professionals will then perform preliminary assessment of such problematic buildings. Regarding buildings of seismic capacity concerns after the brief investigation, civil engineers are hired to complete the preliminary assessment table to assess the buildings. Such investigation and assessment will result in large sum of precious data relating to the seismic capacity of buildings. Finally, according to the collected bridge data, professional and complex assessment methods, such as the pushover method, are applied to carry out the detailed assessment to obtain the highly accurate yielding acceleration (Ay) and complete damage acceleration (Ac). However, there are more than forty thousand bridges in Taiwan. When applying the simple assessment method to carry out visual investigation, the results will be not accurate despite the fast speed of investigation. In the case of applying the detailed assessment, although the results are relatively accurate, it takes much more time and costs. In addition, detailed assessment can only be done by experienced professionals. To conduct detailed structural analysis of each bridge with limited funds and professional labor is impossible. If simple assessment factors and the mapping relation between Ay and Ac for detailed assessment of similar bridges can be identified to infer the Ay and Ac values of other bridges, it will save a great deal of manpower when acquiring Ay and Ac values within the tolerable error range. Therefore, this study aimed to first use the bridge detailed assessment results of the "highway seismic capacity of bridge assessment and reinforcement project feasibility study" by the Directorate General of Highways, Ministry of Transportation and Communications, Taiwan. The artificial intelligence inference model was applied to
find the mapping relation between inputs (brief investigation seismic capacity influence factor) and outputs (Ay and Ac) via cases (seismic capacity of bridge assessment results) learning. At present, artificial intelligence inference models mainly utilize learning models such as the artificial neural network, and support vector machine (SVM). However, such models have parameter setting and initialization problems. For considerations of the searching speed and inference accuracy, this study developed an “artificial intelligence mechanical learning inference model” by using fast and messy GA (fmGA) integrated with SVM. The model searched for the most appropriate model parameters by fmGA, and applied SVM to find the relationship between inputs (brief investigation seismic capacity influence factors) and outputs (detailed assessment of Ay and Ac), and further developed an optimal inference model. The advantage of such inference model was that it could improve the prediction accuracy by case database updating and increasing number of cases. To understand the accuracy of the model after training, this study employed the root mean square error (RMSE) in calculation equation to measure the model learning accuracy. The 10-folds testing results illustrate that the RMSE was 0.09 and 0.13 (as shown in the Table3).

<table>
<thead>
<tr>
<th>Fold no.</th>
<th>Ay(g) RMSE</th>
<th>Ac(g) RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.085</td>
<td>0.17</td>
</tr>
<tr>
<td>2</td>
<td>0.144</td>
<td>0.194</td>
</tr>
<tr>
<td>3</td>
<td>0.058</td>
<td>0.065</td>
</tr>
<tr>
<td>4</td>
<td>0.066</td>
<td>0.059</td>
</tr>
<tr>
<td>5</td>
<td>0.071</td>
<td>0.071</td>
</tr>
<tr>
<td>6</td>
<td>0.143</td>
<td>0.189</td>
</tr>
<tr>
<td>7</td>
<td>0.09</td>
<td>0.084</td>
</tr>
<tr>
<td>8</td>
<td>0.071</td>
<td>0.11</td>
</tr>
<tr>
<td>9</td>
<td>0.054</td>
<td>0.144</td>
</tr>
<tr>
<td>10</td>
<td>0.117</td>
<td>0.214</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.09</td>
<td>0.13</td>
</tr>
</tbody>
</table>

3.6 Earthquake simulation

User can assign the epicenter, location, and depth of earthquake. For example, 921 Chi-Chi earthquakes can be simulated in the system.

3.7 Earthquake event

According to Central Weather Bureau earthquake alert information collected by agent system, the model can evaluate the PGA of bridge automatically.

3.8 Earthquake PGA analysis

Taiwan is divided into several grids in order to evaluate the PGA of every zone. Distances, site and other parameters are input of seismic attenuation equation. The output is accelerating of every zone.

3.9 Bridge failure probability

Finally, according to bridge’s location, PGA of every bridge can be calculated. If the accelerate is higher than Ay or Ac. The bridge failure probability will be high (as shown in Figure 5). In the Figure 6, 921 Chi-Chi earthquake event is simulated in the system. High failure probability bridges will be listed. Therefore, alert message will send to manager’s phone by SMS (Short message service) automatically.

3.6 Earthquake simulation

User can assign the epicenter, location, and depth of earthquake. For example, 921 Chi-Chi earthquakes can be simulated in the system.

3.7 Earthquake event

According to Central Weather Bureau earthquake alert information collected by agent system, the model can evaluate the PGA of bridge automatically.

3.8 Earthquake PGA analysis

Taiwan is divided into several grids in order to evaluate the PGA of every zone. Distances, site and other parameters are input of seismic attenuation equation. The output is accelerating of every zone.

3.9 Bridge failure probability

Finally, according to bridge’s location, PGA of every bridge can be calculated. If the accelerate is higher than Ay or Ac. The bridge failure probability will be high (as shown in Figure 5). In the Figure 6, 921 Chi-Chi earthquake event is simulated in the system. High failure probability bridges will be listed. Therefore, alert message will send to manager’s phone by SMS (Short message service) automatically.
variable) and the values of \( Ay \) and \( Ac \) (output variable) to build the seismic capacity of bridge diagnostic model.

The highly efficient and accurate bridge seismic capacity prediction model proposed in this study can effectively improve the inability to assess the bridge seismic capacity in real time by traditional calculation approach, and hence considerably reduce time and costs. The prediction results can be a reference for relevant management personnel when doing maintenance.

Acknowledgements

The authors thank the support from Harbor and Marine Technology Center in Taiwan, MOTC.

References

BIM based Schedule Control for Precast Concrete Supply Chain

J. Nissilä\textsuperscript{a}, R. Heikkilä\textsuperscript{b}, I. Romo\textsuperscript{a}, M. Malaska\textsuperscript{b} and T. Aho\textsuperscript{b}

\textsuperscript{a}Skanska Oy, Helsinki, Finland
\textsuperscript{b}Construction Technology Research Center, University of Oulu, Finland

E-mail: jaakko.nissila@skanska.fi, rauno.heikkila@oulu.fi, ilkka.romo@skanska.fi, mikko.malaska@oulu.fi, timo.aho@oulu.fi

Abstract - Development of the schedule control of precast concrete supply chain has been studied. Main idea was to use BIM model created by structural engineer as a user-interface for schedule control, for saving different status information of the real-time schedule situation of the propagation of structural design, element manufacture, delivery and site erection directly to the BIM model by using a cloud-based networked service. Some of the missing software applications were programmed by the software companies participated in the project. Experiments were done in a real construction project in Finland, where the information from design, prefabrication, delivery and erection phases was synchronized between the stakeholders by using the cloud service. The most important observations and results are introduced and analyzed. A future model for intelligent BIM based schedule control concept is concluded.

Keywords - BIM, supply chain, schedule control, precast concrete

1 Introduction

Since the use of building information modelling (BIM) to design tasks has increased a lot in most industrialized countries, more and more attention has been paid to BIM model utilization in production phases, such as prefabrication, transportations and on-site assembly and erection works. Schedule control is one of the key management issues in the control of BIM based production flow. Traditionally schedule planning is made using special schedule software such as Microsoft Project, Primavera’s Project Planner, Vico Control, etc.

The extension of CAD based design to schedule management has been studied and developed as a part of so called 4D CAD research area. The main idea has been in combining 3D CAD models with construction activities to display the progression of construction over time (visual intelligence for construction management). 4D models have used to link components in 3D CAD models with activities from the design, procurement, and construction schedules. The final 4D production model allows stakeholders to view the planned construction on the screen and to review a 3D CAD model for any day, week, or month of the project. Stanford University has also developed a special 4D interface for the interactive use of 4D models for planned schedule examinations. [4] Even though the construction industry has used BIM-based schedule management for year, an automatic information flow management between the project’s stakeholders has not been generalized.

Newt Finnish BIM guidelines (COBIM) require that the critical installation dates for designated structures and systems needs to be saved in the building information model. The BIM-based schedule must also be distributes for other parties’ use in an agreed format. Model views must be shared with other parties without the need for separate BIM-based software. [2]

Precast concrete elements, precast concrete wall panels or other components are usually fabricated in a central plant where industrial production techniques are used, then hauled to the construction site and erected.

The information flow during supply chain of the precast building element is important. Traditionally the propagation of design, prefabrication, delivery and erecting information has been handled by traditional methods, such as e-mail, phone and by fabricators project portal. Generally, traditional forms of information dissemination in construction industry are seen ineffective and time-consuming. [1][3][5][8]

By using 4D CAD / BIM models combined with cloud service, information batches as called as statuses of building element give important
information for project stakeholders of the propagation of the supply chain. Each building element can be followed in the supply chain individually. The status information in the precast building element supply chain is divided usually in four phases: design, fabrication, delivery and erection.

The aim of this research was to study the BIM-based supply chain management of precast concrete elements.

2 Methods

Scheduling has to be integrated as a part of the BIM-based supply chain management. The schedule changes in the supply chain have to be linked to the scheduling program used. All the delays on the critical path should be visualized in the modelling program (integrated modelling and scheduling).

In the experiments, the main idea was to use BIM model created by structural engineer (used Tekla Structures) as a user-interface to the schedule control, save different status information of the real-time schedule situation of the propagation of structural design, element manufacture and site erection directly to the BIM model by using a cloud based networked service. Experiments were done in a real construction project in Finland, where the information from design, prefabrication, delivery and erection phases was synchronized between the stakeholders by using the cloud service.

The basis of the data transfer was an online information management system, Trimble Connected Community (TCC) cloud service, through which Extensible Markup Language (XML) files containing status updates were transferred. To mitigate the risk of losing traceability of the precast elements the precast elements were tracked by Globally Unique Identifiers (GUID) through the supply chain.

An example of the XML language is shown in Figure 1. In the code each attribute has a UDA (User Defined Attribute) name, for example “PLANNED_END_DEL”, type “DATE” and value “2013-10-07T00:00:00”. Value indicates both date and time.

![An example of XML language in Status exchange.](image1)

![User interface of Tekla Structures’ Status Sharing tool.](image2)

Structural designer’s and contractor’s status updates were entered manually into Tekla Structures UDA tab pages. Installed external TS Status Sharing extension operated as collecting and synchronizing tool between TS native model and the cloud service. The Status Sharing tool maps and writes above-mentioned code from all the building elements which have the certain attribute information. Transferred attributes can be settings as shown in Figure 2. Stakeholders can act also only as a recipient part, which enables utilization also for project’s customer.

A link to collect the status information from precast manufacturer’s ERP (Enterprise Resource Planning) was created in order to automatize the exchange of fabrication and delivery information to the cloud service. The link was coded to the database of ERP in PL/SQL. On the basis of defined project, the link was ordered to collect all demanded statuses from the ERP by GUID. Collected data was used to create a XML-file
which complied with TS definition. XML-file was written by using the UTF-8 character set. The fully automatic and daily file transfer to TCC was completed by using FTP-application.

The exchange of status information was tested between the structural engineer, the manufacturer and the general contractor. The method was tested in 2013 in the pilot project (As Oy Kauniasten Kvartetti, Kauniainen, Finland), where the information from the design, fabrication, delivery and erection phases was synchronized between the stakeholders by using the cloud service (Figure 3). [6]

Figure 3. The system used to the status information exchange.

The status information was manually entered into the modelling program in the structural engineering office and in the construction site. An automated link was created between the manufacturer’s ERP and the cloud service in order to automatize the exchange of the fabrication information. All status information was readable and visualized in the participant’s Tekla model using predefined color-coded visualizations (Figure 4).

Figure 4. An illustration about the schedule situation (green – erected, blue – design delayed, light brown – fabrication delayed, pink – delivery delayed, orange – erection delayed), As Oy Kauniasten Kvartetti, Kauniainen.

The most important target in the scheduling process of precast concrete elements (Table 1) was the sending of fabrication drawings to the factory (4-6 weeks before delivery) by the structural engineer and the sending the erection plan to the factory by site personnel (3-4 weeks before delivery). Site personnel ordered the elements on site 1-2 weeks before deliver by email. As the separate scheduling software, Vico Control program was used to make a master schedule of the project. On the basis of Vico’s schedule, more accurate element-specific schedule was done in Tekla Structures (TS).

To ensure a functional supply chain, anticipation needs for the scheduling has to be done carefully. Time value attributes in BIM offer considerable scope into advanced project management. This leads to a chance to calculate predictions of probable duration of the project. Simple assessments of each task durations are adequate in first step. Furthermore, risk analyses could be put into practice using standard methods of probability calculus in some extent. Table 1 presents values for scheduling.

Table 1. Anticipation needs for the scheduling of precast concrete elements.

<table>
<thead>
<tr>
<th>Main phases of supply chain for precast concrete</th>
<th>Weeks before the start of the delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Tender inquiries of the separately detailed concrete elements</td>
<td>13-18</td>
</tr>
<tr>
<td>- Delivery contract</td>
<td>10-15</td>
</tr>
<tr>
<td>- Kick-off meeting for the separately detailed concrete elements.</td>
<td>12-14</td>
</tr>
<tr>
<td>- Source data for the design</td>
<td>9-14</td>
</tr>
<tr>
<td>- Schedule for separately detailed concrete elements</td>
<td>9-14</td>
</tr>
<tr>
<td>- Information about special material and deliveries</td>
<td>8-10</td>
</tr>
<tr>
<td>- Element diagrams</td>
<td>8-9</td>
</tr>
<tr>
<td>- Exploratory elements</td>
<td>6-7</td>
</tr>
<tr>
<td>- Kick-off review for the example of the concrete element</td>
<td>5-6</td>
</tr>
<tr>
<td>- Shop drawings by sections, Starting the fabrication, Rough installation schedule</td>
<td>4-6</td>
</tr>
<tr>
<td>- Erection plan by the sectors and stories</td>
<td>4-6</td>
</tr>
<tr>
<td>- Kick-off meeting for installation</td>
<td>1-2</td>
</tr>
</tbody>
</table>
3 Results

The experiences of the method were very encouraging. The exchange of status information succeeded with stakeholders. The main objectives were achieved – the exchange of status information was easy-to-use, adequate real time and reliable. Reliable, transparent and daily exchange of the status information substantially improved the information flow between the stakeholders. The project stakeholders were able to follow daily design, fabrication, delivery and erection progression of the construction project. Progress information was valuable for all the participants of the project. The structural engineer was able to find out if there was still a possibility to change the plans before the element was fabricated. The precast manufacturer’s main target was to produce reliable data for the use of other stakeholders. Also, the precast manufacturer had an easy method to follow the design phase and also the planned erection dates in order to space out the fabrication. The method to track precast concrete elements onsite has been changed from the use of spreadsheets and color-coded drawings to the real time, transparent method, where site personnel can see at a glance if the manufacturer is producing elements in the correct construction sequence and on schedule.

The next step to improve the scheduling process is to integrate the real-time status information exchange to scheduling program in order to affect the schedule when changes occur. Principle of a future model for intelligent, integrated and social BIM-based design and schedule software is presented in Figure 5. On the basis of status updates, the method would enable the automatic comparison of the changes in the precast supply chain with the project schedule, and the effects of the changes, such as design-delays, could be estimated and visualized. At the moment only the colour visualization of delayed building elements is possible.

Figure 5. An integrated modelling and scheduling.

Also, as the site personnel are obligated to inform the manufacturer about the schedule changes in erection order, BIM-based status information exchange provides a tool to leave more time to make last-minute changes to the fabrication and delivery processes (Figure 6).

Also, the method should enable to verify, which elements has been given a new status. A way to inform the viewer that something has changed is needed. Only visual detection is allowed right now.

4 Conclusion

The experiments using the status information exchange system introduced were successful. The exchange of the status information into the BIM-based supply chain process can be easily implemented. In Finland, the exchange of status information into BIM-based supply chain process has already been successfully implemented. When status information production is properly configured, the cloud based approach enables unlimited numbers of information producers and consumers. Moreover, it does not depend on specific features of any cloud service of ERP system but should be possible to implement with any of them.

In control of the schedule, the saving of different status information into building information model facilitates the tracking of schedule situation in design, fabrication, transportation and site erections. However, the information entering remains manually in the designing and erection phases. In future, in order to ease the information management, the information entering phase should take place as a natural part of normal processes. Huge potential is seen in the integration of BIM-based design software and commercial schedule software is a soluble
problem. Integration and comparison of the changes in the supply chain to the project schedule require more investigation. Especially integrating scheduling software using efficiently BIM information and the method described above is seen to have great potential to improve schedule control in building sector.

References


Information Modelling based Tunnel Design and Construction Process

R. Heikkilä, A. Kaaranka and T. Makkonen

Construction technology Research Center, University of Oulu, Finland

E-mail: rauno.heikkila@oulu.fi, tomi.makkonen@oulu.fi, annemari.kaaranka@oulu.fi

Abstract - The present development situation of 3D and BIM based process model for tunnel design and construction methods has been studied. Tunnel design is generally done using traditional 2D drawing based working methods. As a basement for the research, the new specifications developed in Finland for national BIM requirements for infra sector were used. However, underground tunnels are not yet considered in these specifications. In a research project, a new concept how to use intelligent information modelling in tunnel works have been under development. For comparison, five different tunnel design and construction projects and work areas were studied and evaluated. A technical-economic evaluation for the future model of tunnel design and construction is also presented.

Keywords - 3D, BIM, tunnel design

1 Introduction

In Finland, a special effort has been focused on the development of Building Information Modelling (BIM) based process model and guidelines for infra industry including roads, streets, railways and bridges (www.infrabim.fi) [5]. Due to the resource limitations, underground tunnels in road and railway construction projects or in mining have remained to be out of this work, which was one reason to perform this study. In the world, we have found only one similar activity, which has been done a little bit earlier in Norway with HB138 specifications for infra sector as well [3]. In addition, University of Oulu and VTT Technical Research Centre of Finland have an active project called FlexiMine (Flexible Optimization of Mine Production) that aims to the same area of information modelling and automation based development of current work methods and technologies in tunnelling and underground mining.

The HB183 information modelling specifications of Norway determine also tunnel works including guidelines for documentation, how to model existing situation, models of new structures as well as roles and task for practical works. Tunnels are considered as underground objects, to which a determined level of uncertainty is allowed for 3D coordinates. According to HB 138, the 3D geometry of tunnels has to be modelled in its entirety. There are also accurate specifications set for the information transfer from design software used to the LandXML standard based open information transfer file. [3]

Figure 1. A example for tunnel modelling by Norwegian HB 138 specification [3].

McHattie (2013) examines the intelligent mine concept provided by Bentley. Intelligent mining implies the application of information technology at every phase of the mining value chain, from exploration and geological modelling to equipment, operations and maintenance, and logistics and transportation. According to the paper, the fundamental idea is that the operator of mine has a complete 3D digital representation of the physical world, i.e. the digital asset of the mine. For the infrastructure information exchange, Bentley offers i-model, which gathers all the different information gathered together. [4]

Two major tunnelling projects, the first in London (Crossrail, 21 km tunnels, 37 new stations, value about £15 billion) and the second in Stockholm (Bypass, 18 km tunnels, 3 lanes in each direction, the largest road tunnel project in the world, value about £3 billion) do presently apply and utilize Bentley concept. A BIM Academy has been established for the support Crossrail project needs. In UK, the
Goverment Construction Strategy 2011 announced the intention to require collaborative 3D BIM with all project and asset information, documentation and date being electronic on its projects by 2016.

2 Methods

As a part of our Fleximine research project, a new concept for intelligent information modelling and management for tunnelling and underground mining is to be developed. The principles and different parts and examples of the concept are here introduced. In this future concept, continuous 3D initial data models are measured and automatically modelled including digital terrain model, underground soil model and rock model. The accuracy of measurement is any time evaluated and saved to be transferred with the measured information models. The soil model includes all of the different soil material layers (rock, moraine, sand, silt, clay, etc.), their geomechanical and workability features. All of this information is connected to a combined information model using advanced classification, numbering, nomenclature and information structure. This initial information model is saved using open and intelligent format, which is a wide used international standard.

Tunnel design will be done using fully 3D and information modelling based methods and the measured 3D initial data model as a starting point. For the construction operations, different intelligent 3D machine control models (production models) are created and saved to the database of the construction project. Each of the machine control systems pick up automatically the machine control models from the database. Site managers and foremen use dynamic site control system for the guidance and control tasks of the whole work site. Every moving machine and human are continuously positioned using an accurate underground 3D measurement system. Information transfer is done using wireless network. Automated machine control is based on operator assisted or autonomous and unmanned control principle.

In the experiments, the development and utilization situation in five different tunnel design and construction projects and work areas were studied and evaluated: a tunnel part of West Metro (Länsimetro) project in Helsinki (Pöyry Oy, YIT Construction), Finland, Rantatunneli project in Tampere (A-Insinöörit Oy, Lemminkäinen Infra Oy, Tampere City, Finnish Transport Agency, Saanio & Riekkola Oy [6]), Finland, tunnel works in the Crossrail project in London (UK), tunnel design process in the Stockholm Bypass project in Sweden, and tunnelling works in Outokumpu Kemi Mine in Kemi, Finland.

3 Results

The current design and construction methods in West Metro project of Helsinki and Espoo was studied by interviewing the main design consultant office Pöyry and the tunnel contractor YIT. The design method used was traditional 2D drawing based according to the special tunnel design specification of Helsinki City. Designer has typically created no deliverable 3D design model. For some special examinations, 3D models have, however, been designed, most often to compare the measurement result of 3D laser scanning with the designed tunnel geometry. In these cases, the tunnel parts have had to model after the design work to enable the deviation calculations and tolerance checking. The contractor has used a typical ground based 3D laser scanning system for the measurements. A special illustration method has been developed for the deviation examinations. In the method, the 3D designed tunnel model is opened to 2D level, where colours illustrate the calculated 3D deviations.
Figure 3. A comparison of measured 3D as-built model with designed model. The 3D design model needed to be extruded afterwards during construction work for the calculations.

Figure 4. An example of BIM based coordination in tunnel design (Liverpool Station, Crossrail, London, UK).

In the Crossrail (London, UK) project, BIM is utilized largely to integrate data for design, construction and operation. The project uses a project wide consistent approach to flow and production of information (information management) and 3D models are produced by all team members to common level of detail using common tools (information modelling). Direct benefits of BIM have been reduced waste, improved efficiencies, reduced information loss, improved safety, reduced programme risk, improved performance, collaborative model transfer from designer to contractor, and innovative asset management possibility. According to a Crossrail presentation, in the design of only one station (Farrington Station), a cost of £120k was used to develop the 3D model, but it saved over £8 million from risk contingency due to interfacing complexity. The process model used in the Crossrail project is based on the British Standard BS1192.

Figure 5. An example of BIM in tunnel design (Crossrail, London, UK).

Figure 6. An example of coordination model (Farrington Station, Crossrail, London, UK).

In the Stockholm Bypass project, 3D spatially coordinated and integrated design models are used. This was seen to be necessary for the large project with very demanding 3D tunnel geometry as well as for very international organization for the project activities (over 500 engineers, 19 disciplines, 7 European countries). The main objective of the use of BIM is to create a collaborative project working environment. Also satellite machine guiding systems will be used. Even thought the project locates in Sweden, the process model is based on the same British Standard BS1192.

Figure 7. An example of complicated ramps and interchanges in a tunnel model of Stockholm Bypass (URS, London 2014).

The underground Kemi Mine has been evaluated to be one of the most intelligent mine in the world [4]. Currently this mine includes about 50 km underground tunnels. After mining and rock extraction new tunnel parts have been systematically measured using 3D laser scanning systems. All the measured 3D data has been combined into 3D mine model that forms the basement for the whole mine production control. There are a number of wlan base stations that enable the fluent information transfer in the underground mine part as well. A special
underground production control center uses the 3D mine model as a base user-interface for the control of all of the mining tasks.

Figure 8. 3D Kemi Mine model.

Figure 9. Examination of the measured rock strengths on the level -400 m using the mine model (blue – good areas, red – weak areas).

Rantatunneli project in Tampere, Finland, is executed as an alliance project. The alliance project is planned and carried out together by all parties from the development phase to the implementation phase. The alliance enables better co-operation, information sharing and innovation between the alliance parties. [6]. Roads are designed using 3D modelling and the Infra BIM guidelines. Also bridges and challenging interfaces of different structures (road junctions and tunnel endings) are 3D modeled carefully, but tunnels not. That was not considered to be necessary or cost-effective. According to the contractor, there have been planned some flexible tolerances for tunnel structures that BIM modeling was not needed.

4 Conclusions

The Finnish Infra BIM guidelines do not yet determine the process model and detailed specifications for tunnel design and construction work phases. Instead, the corresponding Norwegian HB138 does determine some basics for the BIM process of tunnels. In principle, the specifications are similar as the guidelines for road structures.

The findings of the considered tunnel work in Europe shows diversity in design and construction methods: typically in small projects the design method is still traditional 2D drawing based like in West Metro and Rantatunneli projects in Finland. Currently, contractors do not see clear and direct benefits in the utilization of BIM methods. Instead, in more large Crossrail and Bypass projects everything information management has been clearly based on the BIM based method and solutions. Significant savings and benefits have already been reported and shown in the Crossrail project. The British Standard BS1192 seems to be important for implementing the functional information modelling based process into practice. The information mobility concept provided by Bentley offers many new possibilities into the development of information modelling, management and utilization of tunnel works. In the well-established Kemi mine underground environment, the advanced production control uses the accurate 3D mine model as the main user-interface to the information management. There are different challenges to develop wireless information transfer as well as 3D positioning.

References

Abstract -
A three-dimensional (3D) laser scanner is one of the most well known devices when it comes to creating a 3D model of existing bridges and buildings. Most 3D scanners can pick up point clouds very accurately, which can be used to create an accurate 3D model of existing objects. One may speculate that a 3D laser scanner can also be used to pick up the status of a construction project on the job site and transform them into a 3D computer model. However, most 3D laser scanners are still expensive and they are not easy use yet especially on a congested construction site. In addition, it takes a significant amount of time to create a 3D computer model using point clouds picked up by the laser scanner. As emerging photogrammetry techniques demonstrated the use of photos instead for rapid 3D modelling, one may be wondering 1) if this technique can be used to create a 3D model of a construction site quickly, and 2) if this model is accurate enough to help project managers make some decisions. This paper presents our test demonstrating the process of creating a 3D model of an existing building using photos. It presents some challenges we faced when taking photos and creating a 3D model. This paper also presents the method we came up with to create a 3D model of the entire building without using any control points.

Keywords -
3D CAD Model; Photogrammetry

1 Photogrammetry

Photogrammetry is a technique of creating a 3D model of an object from one or more photographs of that object. Software tools introduced recently facilitate object reconstruction and creation of a 3D computer model from digital images without requiring the domain knowledge of photogrammetry.

Photogrammetry has been used in many areas including architecture, heritage preservation, engineering, forensics and accident reconstruction, and medical applications. Almagro et al. [1] produced a model of the Otto Wagner Pavilion in Vienna using photos, which is used by CIPA (the International Council on Monuments and Sites) as a reference building for testing modern methods of measurement and processing in architectural photogrammetry. Arias et al. [2] combined the graphic and metric documentation on the traditional agro-industrial buildings, which are an important part of the heritage of Galica (Northwest of Spain) using close-range photogrammetry techniques.

Mills et al. [3] measured deformation of a pavement within the Newcastle University Rolling Load Facility. Precise 3D measurements of the pavement have been produced from stereo-imagery taken with different cameras, using both analytical and digital photogrammetric instrumentation.

Przybilla et al. [4] presented all the stages involved in the procedure for the determination of the shape of a fuel assembly, which is an essential part of a nuclear power plant and can only be handled underwater. Photo-triangulation was used to obtain orientation elements, taking into account light refracting surfaces. Stereo-models were then set up for analytical restitution and the shape parameters for the object can be obtained.

Fraser and Riedel [5] monitored the deformation a series of super-hot steel beams using digital close-range photogrammetry. An on-line configuration of three CCD cameras was established to measure both stable reference points and targets subject to positional displacement.

Fenton and Ziernicki [6] presented a method of determining a vehicle crush and equivalent barrier speed using digital photogrammetry. Close-range photogrammetry allows engineers and accident reconstructionists to create 3D computer models of damaged vehicles utilizing photographs. Utilizing photogrammetric software PhotoModeler, engineers can digitize accident scene photographs and create accurate 3D computer models of the vehicles, which then can be used to quantify structural damage sustained by the vehicles. Knott Laboratory utilized these techniques on a case of Princess Diana accident in France.

Lynnerup and Vedel [7] analyzed surveillance images from a bank robbery, and the images were compared with images of a suspect.

Burke and Beard [8] monitored facial shape as it
changed over an extended period of time through growth. Walton [9] involved photogrammetry in the therapy of various gait problems arising primarily from deformities or injuries.

2 Photogrammetry in Construction

Abeid et al. [10] integrated the site construction progress bar chart in MS Project with a database of digital site pictures showing the building process and building elements at particular points of time. The digital pictures taken from up to four cameras are placed on a website, where a remote computer can capture and store the pictures in the database. The system enables management staff of contractors and owners to follow developments at the construction site in real time. Additionally, time-lapse films of activities at the construction site taken by multiple cameras can be played back in synchrony with dynamic graphs showing planned versus actual schedules. A new concept in time-lapse photography has been introduced. It enabled a reasonable playback time as well as the implementation of the technology for long-term construction projects using standard personal computers.

Kim and Kano [11] suggested a method for determination of the 3D viewpoint and the direction vector of a construction photograph to perform comparison of the construction photograph and the corresponding virtual reality (VR) image. They developed photo images in 3D computer graphics showing the as-built site situation. These photos were compared against the corresponding as-planned CAD images. Application cases in foundation excavation, refill, scaffolding and steel erection proved their methodologies to be convenient and effective in checking actual site progress against as-designed or as-planned models.

Memom et al. [12] identified the techniques, which were used in the construction industry for monitoring and evaluating the actual physical progress, and discuss the Digitalizing Construction Monitoring (DCM) model. The DCM model is an interactive system integrating 3D CAD drawings and digital images. The authors made a practical attempt to automate the process of producing as-built construction schedule by applying modern photogrammetry techniques to photographs and integrating with CAD drawings. The application of a DCM model in monitoring the progress would enable project management team to better track and control the productivity and quality of construction projects.

Quinones-Rozo et al. [13] explored the use of two image-based techniques to perform semi-automated tracking of excavation activities. An Enhanced Pattern Detection and Comparison (EPDC) technique was introduced to quickly identify changes in poor contrast excavation surfaces.

Luhmann and Tecklenburg [14] used site photos to measure 3D geometries of buildings adjacent to a construction site in order to preserve forensic evidence against potential construction-caused damage claims.

Kamat and El-Tawil [15] discussed the feasibility of using augmented reality (AR) to evaluate earthquake-induced building damage. In the proposed approach, previously stored building information is superimposed onto a real structure in AR. Structural damage can then be quantified by measuring and interpreting key differences between the real and augmented views of the facility. Proof-of-concept tests were performed in conjunction with large-scale cyclic shear wall. They measured and interpreted the drifts between the original walls in 3D CAD images and the actual wall specimens for post assessing any earthquake-induced building damages.

3 Issues on Accuracy

Photogrammetry has been seldom used actively to achieve a 3D model in the construction industry although it has been regarded as the most cost-effective, flexible, and portable approach in terms of getting a 3D model. Among many reasons keeping industry practitioners from actively using photogrammetry for 3D modelling is a doubt on the accuracy of the 3D model created from photos.

Burt [16] investigated factors affecting the accuracy of digital photogrammetry. From his study on historic adobe wall ruins located at Fort Davis in Texas, he demonstrated that photogrammetry is a suitable method for obtaining measurements of adobe erosion.

Randles et al. [17] compared traditional technique and photogrammetry for measurement of targeted damaged vehicle. The points on each vehicle were measured using both techniques, and compared. After all the calculations and comparisons, they concluded that both methods effectively measured the vehicle points, with a mean difference between the baseline and hands-on measurements of 0.6±1.4cm, and a mean difference between the baseline and photogrammetry measurements of 0.1±1.0cm.

However, a test conducted by Bhatla et al. [18] reported that a 3D model of a bridge created from photos was not accurate enough for construction professionals to use for decision-making. They created a 3D model of a 2,000ft bridge in southern United States, which was under construction, using 351 photos taken on the job site. They compared it against a 3D model developed manually from the 2D drawings. The average height of exterior girders, average distances between the holes for electrical fixtures, and the average length and width of the floor beams were compared between two
models. They then found about 2 to 5% of differences between two models on the length of beams, height of exterior box girder, and distance between the holes for electrical fixtures. They concluded that photogrammetry was not suitable yet for modelling infrastructure projects.

4 Empirical Test

Our research team was wondering if current photogrammetric computer applications would enable us to create a decent 3D model that can be used by construction professionals for their decision-making. Seeking the answer for this question, our team decided to create a 3D model of an existing campus building using photos. The following figure shows the plan of the campus building we used for our test.

Figure 1 Floor plan of a building used for the empirical test

A digital single-lens reflex camera with 18-55mm lens attached was used for taking photos. The camera was placed on a tripod at multiple locations in the building, and photos were taken while the camera was getting rotated horizontally by 45 degrees and tilted vertically by 30 degrees in order to pickup all objects around the camera. As many as 1329 photos were taken for the test, and we spent a total of 278 minutes for taking those photos.

We then used Autodesk Stitcher to generate a panoramic image of the objects around the camera. Autodesk Stitcher is designed to generate a panoramic image automatically. However, some photos were not stitched automatically, to some extent because of lack of texture and contrast of some objects. A total of 597 minutes were consumed for image processing.

For 3D modelling, we used Autodesk ImageModeler 2009. This application enabled us to pickup the boundary lines of building components including columns and walls. We drew lines on the edges of these objects and use them to create a 3D model of building components picked up by the camera at one location. We then map the photos on the surface of the 3D model. This process took a total of 87 minutes.

5 Accuracy of an Empirical Model

A total of 38 dimensions were extracted from our empirical model, and then compared with dimensions extracted from the 3D model created using CAD drawings. Differences between each dimension are defined:

- Difference = Dimension from CAD model – Dimension from photo model
- Difference (%) = Difference / Dimension from CAD model x 100

The following table presents some of differences we figured out.

Table 1 Differences in dimension between photo model and CAD model

<table>
<thead>
<tr>
<th>Items</th>
<th>Photo model (m)</th>
<th>CAD model (m)</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.39</td>
<td>2.39</td>
<td>0.11</td>
</tr>
<tr>
<td>2</td>
<td>5.05</td>
<td>5.02</td>
<td>-0.67</td>
</tr>
<tr>
<td>3</td>
<td>2.58</td>
<td>2.44</td>
<td>5.81</td>
</tr>
<tr>
<td>4</td>
<td>11.16</td>
<td>11.17</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>2.45</td>
<td>3.06</td>
<td>19.94</td>
</tr>
<tr>
<td>6</td>
<td>2.67</td>
<td>2.60</td>
<td>2.55</td>
</tr>
</tbody>
</table>

The mean and the standard deviation of all differences are 0.87% and 8.14% respectively. After eliminating the outliers with 95% confidence, the mean and standard deviation are -0.43% and 4.33% respectively. Then, the dimensions were categorized into four groups, and the mean and standard deviation for each group were calculated again to compare with each other. The result shows that when a length of a section is less than 3.81m (150 inches) or more than 10.16m (400 inches) the accuracy level of the model drops. The reason for this inaccuracy may be the result of distortion of panoramas, or the operator’s mistakes when creating the model using panoramic photos.

6 Conclusions

In order to see if a 3D model created from photos is...
accurate enough for construction professionals to make decisions, we empirically created a 3D model of an existing campus building using 1,329 photos. It took 16 hours to produce a 3D model using these photos. A total of 278 minutes were consumed for photo taking, 597 minutes for stitching photos together, and 87 minutes for 3D modelling.

Dimensions were extracted from 38 difference locations in the photo model and compared with the dimensions extracted from the CAD model. It turned out that the differences between these dimensions were 0.87% in average. However, the accuracy level dropped when the length of the object was either shorter than 3.81 m or longer than 10.16 m.

Our test shows that a 3D model created from photos are fairly accurate for construction managers to figure out the status of job site. When the photo model and CAD model were both presented to industry professionals in the BIM CAVE (Computer Aided Virtual Environment for BIM), they indicated that they gained more sense of presence in the 3D model while walking though the photo model.

References

Using Benders Decomposition for Solving Ready Mixed Concrete Dispatching Problems

Mojtaba Maghrebi, Vivek Periaraj, S. Travis Waller, Claude Sammut

School of Civil and Environmental Engineering, The University of New South Wales (UNSW), Sydney, Australia
Department of Systems and Industrial Engineering, The University of Arizona, Tucson, AZ, USA
National Information and Communications Technology Australia (NICTA), Sydney, Australia
School of Computer Science Engineering, The University of New South Wales (UNSW), Sydney, Australia
E-mail: {maghrebi, s.waller, c.sammut}@unsw.edu.au, vivek.periaraj@gmail.com

Abstract

Large scale dispatching problems are technically characterized as classical NP-hard problems which means that they cannot be solved optimally with existing methods in a polynomial time. Benders decomposition is recommended for solving large scale Mixed Integer Programming (MIP). In this paper we use the Bender Decomposition technique for reformulating the Ready Mixed Concrete Dispatching Problem (RMCDP). Benders decomposition involves separating the original RMCDP formulation into the master (lower bound) and sub-problems (upper bound). The master problem only deals with integer variables and the sub problem is usually a linear programming problem. Benders optimally cuts and Benders feasibility cuts are added to the master problem upon solving the sub-problem at each iteration. The proposed method is tested on a single real instance and results are reported.

Keywords: Benders Decomposition; Ready Mixed Concrete (RMC), Dispatching

1. Introduction

During the past 10 years a growing body of literature has been devoted to Ready Mixed Concrete Dispatching Problems (RMCDP); however, this area still suffers from a lack of practical solutions [1-6]. In RMCDP it is desirable to find the best truck and depot allocation for each delivery. A few attempts have been made to acquire the exact solution of RMCDP; nevertheless, as a result of increasing the size of the problem the complexity is increased exponentially [4] and cannot be solved in a polynomial time. Heuristic solutions have been implemented widely in the literature to alleviate this problem. Among the introduced methods, Genetic Algorithms is the most promising heuristic solution in the RMCDP literature [1, 3, 5, 7-10]. Other heuristic methods also have been tested in this context, such as Ant Colony [11], Particle Swarm Optimization (PSO) [12, 13], Bee Colony Optimization (BCO) [14] and Tabu Search (TS) [14]. Despite developments in implementing heuristic methods in RMCDP, the solution structure of most of the mentioned techniques is pretty much same. Moreover, the main drawback for these techniques is that there are a number of infeasible allocations in the outcomes of these techniques. Thus, via supplementary algorithms, obtaining a viable solution has been attempted. To overcome this issue, [3] presented an evolutionary based method which can solve the RMCDP without the need for any supplementary algorithm.

Rather than simply looking at heuristic methods some other numerical approaches have been studied. Yan, Lai [15] introduced a numerical method for solving the RMC optimization problem. They proposed a method that works by cutting the solution space iteratively and as well is integrated with branch-and-bound. Lin, Wang [16] introduced a new RMCDP formulation inspired by the job shop problem. Yan, Lin [17] used decomposition and relaxation techniques coupled with a mathematical solver to solve the problem. Variable Neighbourhood Search (VNS) was tested by Payr and Schmid [18] to deal with RMCDP. One of the robust RMCDP formulations was proposed by Asbach, Dorndorf [19]. In this method, depots and customers are divided into sub-depots and sub-customers. More recently, Maghrebi, Periaraj [20] implemented a Column Generation (CG) method which is amenable to the Dantzig-Wolfe reformulation for solving large scale models which with available computing facilities cannot optimally solve in polynomial time. However, the Benders decomposition [21] has not been used in RMCDP, which is the main contribution of this paper.
2. RMC Benders Decomposition

In 1962, Benders introduced a decomposition method for solving MIP which later was generalized by Geoffrion. Benders’ methodology involves decomposing the compact formulation into master (lower bound) and sub-problems (upper bound). The master problem is usually an integer programming problem and the sub-problem is usually a linear programming problem. In each iteration, the sub-problem is solved for a given solution of the master problem. If the sub-problem is optimal, then an extreme dual solution is used to form what is called a Benders optimality cut and added to the master problem. If the sub-problem is primal infeasible (or dual unbounded), then an unbounded extreme dual ray is used to form what is called a Benders feasibility cut and added to the master problem. The master and sub-problems are solved iteratively in this way until the bounds are strengthened and the algorithm converges.

A few RMCDP formulations have been introduced, such as [5, 15-17, 19, 23-25]. To simplify the formulation in some methods [15, 17, 19, 26] the depots and customers are divided into a set of sub-depots and sub-customers, respectively based on the number of loads at depots and the number of required deliveries. The compact formulation of RMCDP can be stated as follows [4, 19] if we assume RMCDP as a graph \( G = (V, E) \) in which \( V \) is the set of vertices belonging to start points, customers, depots and end points \( V = \{ u_s \cup C \cup D \cup v_f \} \). Additionally, \( E \) is the set of edges belonging to the distance between vertices.

\[
\text{Minimize } \sum_{u} \sum_{v} \sum_{k} z_{uvk} x_{uvk} - \sum_{c} \beta_{c} y_{c} \tag{1}
\]

Subject to:

\[
\sum_{u \in u_s} \sum_{v} \sum_{k} x_{uvk} = 1 \forall k \in K \tag{2}
\]

\[
\sum_{u \in v_f} \sum_{v} \sum_{k} x_{uvk} = 1 \forall k \in K \tag{3}
\]

\[
\sum_{u} \sum_{v} x_{uvk} - \sum_{v} \sum_{j} x_{ujk} = 0 \tag{4}
\]

\[
\forall k \in K, v \in C \cup D \tag{5}
\]

\[
\sum_{u \in u_s} \sum_{k} x_{uvk} \leq 1 \forall v \in C \tag{6}
\]

\[
\sum_{u \in D} \sum_{k} x_{uvk} \leq 1 \forall u \in D \tag{7}
\]

\[
-q_k x_{uvk} \geq q_c y_c \forall c, v \in C \tag{8}
\]

In the RMCDP, the master problem consists of customer only constraints, depot only constraints, demand constraints, delivery constraints and perishability constraints involving only depot to customer arcs. The sub-problem consists of truck start constraints, truck finish constraints, customer flow constraints, depot flow constraints, time truck start constraints, time return constraints and time truck finish constraints.

2.1. Benders Master Problem:

The master problem in RMCDP is a mixed integer programming model and involves assignment of depots and customers subject to the time requirements. The Benders master problem can be formally stated as follows:

\[
\text{Minimize } \sum_{u} \sum_{v} \sum_{k} z_{uvk} x_{uvk} - \sum_{c} \beta_{c} y_{c} + Z \tag{10}
\]

Subject to:

\[
\sum_{u \in u_s} \sum_{k} x_{uvk} \leq 1 \forall v \in C \tag{11}
\]

\[
\sum_{u \in D} \sum_{k} x_{uvk} \leq 1 \forall u \in D \tag{12}
\]

\[
-q_k x_{uvk} \geq q_c y_c \forall c, v \in C \tag{13}
\]
3. Solution Approach

The master problem is a mixed integer problem and can be solved using branch-and-cut [27]. The master problem can be considered as an assignment problem of depot to customer arcs, subject to demand constraints and time restrictions. The sub-problem is solved at the truck level for a given set of optimal assignments associated with a given truck from the master problem. The optimal depots to customer arcs from the master problem determine the bounds of the customer service times at given customer locations. The delivery constraints affect the lower bound, while the perishability constraints affect the upper bound of the customer time. The start to depot arcs, customer to depot arcs and customer to finish arcs in the sub-problem are adjusted for their bounds based on the new bounds of the customer time obtained from the optimal solution of the master problem. The prerequisite for the sub-problem in the Benders solution framework is that it needs to be a linear programming model, such that its weak duality property can be used to derive the Benders optimality and feasibility cuts. The sub-problem can be solved either by the network simplex method [28] or by using the LP optimizer (primal or dual) to obtain the extreme dual solution or extreme unbounded dual ray solution.

The following two integrality properties motivate solving the sub-problem using the network simplex optimizer or the LP optimizer (primal or dual). The first type of integrality property can be formally stated as follows:

\[
\text{Minimize } \sum_{u} \sum_{v} \sum_{k} z_{uvk} x_{uvk}
\]

Subject to:

\[
\sum_{u \in u_s} \sum_{v} \sum_{k} x_{uvk} = 1 \forall k \in K
\]

\[
\sum_{u \in u_t} \sum_{v} \sum_{k} x_{uvk} = 1 \forall k \in K
\]

\[
\sum_{u} \sum_{v} x_{uvk} - \sum_{v} \sum_{j} x_{ujk} = 0
\]

\[
\forall k \in K, v \in C \cup D
\]

0 \leq x_{uvk} \leq 1 if x_{uvk} is feasible

0 \leq x_{uvk} \leq 0 if x_{uvk} is not feasible

u \in C, v \in D, k \in K

u \in u_t, v \in D, k \in K

\[
\forall u \in D, v \in C, k \in K
\]

3.1. Optimality Cuts

From an extreme dual solution of the sub-problem, the following Benders optimality cut is added to the master problem. Let, \( \Pi_k \) be the extreme dual associated with the truck start constraint (2) for the truck \( k \). \( \lambda_k \) be the extreme dual associated with the truck finish constraint (3) for the truck \( k \). \( \theta_{dk} \) be the extreme dual associated with the depot flow constraint (4) for the truck \( k \). \( \theta_{ck} \) be the extreme dual associated with the customer flow constraint (4) for the truck \( k \).

\[
\sigma_k = \text{the sum of extreme duals associated with bound constraints of the arcs in the sub-problem for the truck} k.
\]

\[
\theta_{uk} x_{uvk} - \theta_{vk} x_{uvk} + Z_k \geq (\pi_k - \lambda_k) + \sigma_k
\]

Each optimality cut for a given truck \( k \) is added to the set \( \Pi_k \). The convergence of the algorithm is defined by
the value of the variable $Z$. This variable is sometimes called the approximation variable or the recourse function in the Benders decomposition context and is a measure of the dual objective of the sub-problem. As each optimality cut added to the master problem attempts to improve the value of this bound and thus convergence of algorithm can be stated when the value of $Z$ is equal to or within a specified tolerance of the objective value of the sub-problem.

### 3.2. Feasibility Cuts

From an unbounded dual ray solution of the sub-problem, the following Benders feasibility cut is added to the master problem. Let, 

- $\pi_k$ be the extreme ray associated with the truck start constraint (2) for the truck $k$.
- $\lambda_k$ be the extreme ray associated with the truck finish constraint (3) for the truck $k$.
- $\theta_{u_k}$ be the extreme ray associated with the depot flow constraint (4) for the truck $k$.
- $\theta_{v_k}$ be the extreme ray associated with the customer flow constraint (34) for the truck $k$.
- $\sigma_k$ be the sum of extreme rays associated with the bound constraints of the arcs in the sub-problem for the truck $k$.

Each feasibility cut for a given truck $k$ is added to the set $\Phi_k$.

This process is terminated when the model converges.

### 4. Case Study

The proposed Benders decomposition is tested by field data which belong to an active RMC with three active depots and around 50 trucks. From the available dataset, the data from a day on which 22 customers were to be supplied was selected for further studies. Among the customers, 2 needed 3 deliveries and 4 needed 2 deliveries, while the remainder only needed 1 delivery. The authors believe that this instance is not a very complex RMCDP problem; however, the main goal in selecting a small instance is to provide an opportunity to investigate all aspects of RMC resource allocations in detail.

The algorithm was developed in C++ and tested on a RedHat(R) CentOS(R)5.9 Linux server with 8 3.60GHz Intel(R) Xeon(R) CPUs and a 188 GB physical memory. The IBM CPLEX ™ version 12.5.0.0 with parallel optimizers using up to 8 threads was used in the study.

The most important criteria in optimization is the value of the objective function. In Figure 1 the trend of the objective function inclusive of $Z$ is shown and in Figure 2 the best solution as well as elapsed time over iterations were illustrated. This paper aimed to show how Benders decomposition can be implemented in RMCDP to obtain an optimum solution or near optimum in a practical time. The future research could involve improving the convergence rate for larger problems by devising hybrid methods to minimize the effect of combinatorial aspect of the RMCDP.

### 5. Conclusion

The application of Benders’ decomposition to the Ready Mixed Concrete Dispatching Problem (RMCDP) has been studied in this paper. Optimally solving larger scales of RMCDP was the main motivation for this approach. Benders decomposed the original RMCDP formulation into the master (lower bound) and sub-problems (upper bound). In the master problem only discrete variables are dealt with, and the sub-problem is usually a linear programming problem. A Benders optimal cut is added to the model in each iteration, if the sub-problem obtained a feasible solution. Also, a Benders feasibility cut is added if the sub-problem is unbounded or infeasible. This process is iteratively continued until the problem converges. The Benders formulation of RMCDP was presented in this paper and tested by a real instance. Moreover, the trends of objective function, best solution as well as elapsed time over iterations were illustrated. This paper aimed to show how Benders decomposition can be implemented in RMCDP to obtain an optimum solution or near optimum in a practical time. The future research could involve improving the convergence rate for larger problems by devising hybrid methods to minimize the effect of combinatorial aspect of the RMCDP.

### Acknowledgement

NICTA is funded by the Australian Department of Communications and the Australian Research Council through the ICT Centre of Excellence program.
Figure 1. Objective of master problem

Figure 2. Best solution obtained over iterations
Figure 3. Elapsed time (cumulative) over iterations

Figure 4. Travel between depots (blue dots) and customers (red dots)
Figure 5. Travel between customers (red dots) and depots (blue dots)

Figure 6. Schedule of Trucks
Notations

C Set of customers
C_k Set of customers visited by a truck k
D Set of depots
D_k Set of depots visited by truck k
K Set of vehicles
U_u Set of starting points
V_f Set of ending points
S_u Service time at the depot u
t_vk Travel time between u and v with vehicle k
q_k Maximum capacity of vehicle k
q_c Demand of customer c
w_u Time at node u
\( \beta_c \) Penalty for not satisfying the customer c
M A large constant
Y Maximum time to haul the concrete
x_{uvk} 1 if route between u and v with vehicle k is selected, 0 otherwise
y_c 1 if total demand of customer c is supplied, 0 otherwise
z_{uvk} Cost of travel between u and v with vehicle k

References


Mining Rules for Satellite Imagery Using Evolutionary Classification Tree

L.C. Lien\textsuperscript{a}, Y.N. Liu\textsuperscript{b}, M.Y. Cheng\textsuperscript{c} and I-C. Yeh\textsuperscript{d}

\textsuperscript{a, b} College of Civil Engineering, Fujian University of Technology, China  
\textsuperscript{c} Department of Construction Engineering, National Taiwan University of Science and Technology, Taiwan  
\textsuperscript{d} Department of Civil Engineering, Tamkang University, Taiwan  
E-mail: lclien@gmail.com, ynliu@fjut.edu.cn, myc@mail.ntust.edu.tw, icyeh@chu.edu.tw

Abstract -
Classification Tree (CT) can establish explicit classification rules of Satellite Imagery (SI). However, the accuracy of explicit classification rules are poor. Back-Propagation Networks (BPN) and Support Vector Machine (SVM) both can establish a highly accurate model to predict the classification of SI but cannot generate the explicit rules. This study proposes a novel mining rule method named Evolutionary Classification Tree (ECT) which is composed of Particle Bee Algorithm (PBA) and Classification Tree (CT) that automatically produce self-organized rules to predict the classification of SI. In ECT, CT plays the architecture to represent explicit rules and PBA plays the optimization mechanism to optimize CT to fit the experimental data. 600 experimental data sets were used to compare accuracy and complexity of four model building techniques, CT, BPN, SVM and ECT. The results showed that ECT can produce rules which are more accurate than CT and SVM but less accurate than BPN models. However, BPN is black box models while ECT can produce explicit rules which is an important advantage to mining the explicit rules and knowledge in practical applications.

Keywords -  
Satellite Imagery (SI), Back-Propagation Networks (BPN), Support Vector Machine (SVM), Evolutionary Classification Tree (ECT), Particle Bee Algorithm (PBA).

1 Introduction

Due to vigorous economic development, the change of land usage severely causes the destruction of natural environment and land resources. Thus, how to effectively manage land resources to achieve the purpose of sustainable usage is an important topic. Satellite Imagery (SI) was a record and testing information technology which explores through a sensor to indirect survey with objects [1-3]. Since 1972, the United States launched land satellite to reflect object by the sensor receiving surface of solar electromagnetic radiation. All the raw data was sent back to the earth in the form of numeric data and provide detection of environmental resources information for researcher. SI mining has the characteristic for real time survey to covered extensive area. It has become an effective survey tools to build environmental resource database. The mining steps of SI are: (1) the satellite scans surface spectral reflectance intensity from the sensor's spectrum to obtain the image data. (2) the staffs investigate on site to obtain the surface classifies information. (3) establish the relationship between the surface spectral reflectance intensity data and the surface classifies information with appropriate statistical methods. (4) the established relationship can be directly applying on other surface only based on its surface spectral reflectance intensity data to determine its surface classification. Thus, the staffs do not need to do investigation on site and can save considerable manpower and funding. In other hand, with a quick grasp of the ability of a region-wide data, it can be applied to land use, agriculture and forestry planning, environmental monitoring, disaster assessment, scientific research and other purposes. However, different surface classifications of spectral reactions on SI mining are extremely similar, so that to distinguish surface classification will be confusing. Therefore, how to solve SI classification problems through artificial intelligence (AI) mining technique is the purpose of this study.

In the past few years, (1) artificial neural networks (ANN) have been done a lot in science field. There were also much literature [4-7] proposed complex nonlinear models with highly accuracy for predicting material behavior. But these “black box” models are unable to generate explicit formulas or rules which can explain the essence of the models. Besides, there are a lot of research have been used in SI classification area such as (2) the nearest neighbor classifier (NNC) [8] and (3)
Inductive decision tree (IDT) [9]. However, those methods mostly focus on accuracy (accurately predict the performance of classification model) but ignore the understandability of classification model.

In recent years, some researchers have employed genetic operation tree (GOT) that comprise genetic algorithms (GA) and operation tree (OT) in order to build material model that can accurately predict material behaviors and explain the substance of material models [10-12]. Operation tree is a tree structure that expresses a mathematical formula. Optimizing the operation tree can produce a self-organized regression formula. In general, the accuracy of GOT generated model are lower than those produced by neural networks, but more accurate than those produced by RA [10-12].

The strength of GA lies in its ability to locate the global optimum using random yet directed searching operators. Therefore, the GA is less likely to restrict the search to a local search [13]. Thus, GA was risk finding a suboptimal solution. Another main disadvantage of GA is the excessively long run-time that is needed to deliver satisfactory results for large instances of complex design problems.

A hybrid swarm algorithm, the particle bee algorithm (PBA) was proposed to instead GA that imitates a particular intelligent behavior of bird and honey bee swarms and integrates their advantages [14, 15]. PBA improves BA neighborhood search using PSO search [14, 15] and can solve discrete optimization problem, representing one paradigm of evolution computation. It is based on natural evolution and derived from the ideas of the survival of the fittest and successful applied to many case studies [14, 15]. PBA has some advantages, such as global optimization, local optimization, exploration process, exploitation process, flexibility, and parallelism [14-15].

Besides, due to previously studies [10-12], the researchers applied GOT only on producing self-organized regression formula. There still have classification and clustering problems have to mining. Thus, this study focus on propose a novel self-organized classification tree idea namely evolutionally classification tree (ECT) that optimize the tree rules structure by PBA.

A large number of 600 experimental [16] datasets were used to compare accuracy and complexity of the five model building techniques (CT, BPN, SVM and ECT) and evaluate whether ECT can produce simpler and understandability but accurate classification trees to mining satellite imagery rules.

2 Particle Bee Algorithm (PBA)

Particle bee algorithm (PBA) was proposed by Cheng and Lien [14, 15]. It has been successful applied to many case studies [14, 15]. In PBA, the particle bee colony contains four groups, namely (1) number of scout bees (n), (2) number of elite sites selected out of n visited sites (e), (3) number of best sites out of n visited sites (b), and (4) number of bees recruited for the other visited sites (r). The first half of the bee colony consists of elite bees, and the second half includes the best and random bees. The particle bee colony contains two parameters, i.e., number of iteration for elite bees by PSO (Peitr) and number of iteration for best bees by PSO (Pbitr). PBA flowchart is shown in Figure 1.

---

**Figure 1. Particle bee algorithm flowchart**

Step (1) Initialize scout bees

PBA starts with n scout bees being randomly placed with respective positions and velocities in the search space.

Step (2) Evaluate fitness

Start the loop and evaluate scout bee fitness.

Step (3) Select elite sites (e) from scout bees

Elite sites are selected for each elite bee, whose total number is equal to half the number of scout bees.

Step (4) Elite bees initiate the PSO procedure by Peitr iteration for neighborhood-windows (NW)

In this step, new particle bees from elite and best bees are produced using Eq. (1). Elite and best bee velocity update are performed as indicated in Eq. (2). This study further proposes a neighborhood-windows (NW) technique to improve PSO searching efficiency as show in Eq. (3). Thus, after $x_{i\omega(t+1)}$ is substituted into Eq. (1) and Eq. (2), the NW ensures PSO searching...
within the designated $x_{id\min}$ and $x_{id\max}$. In other word, if the sum of $x_{id}(t+1)$ exceeds $x_{id\min}$ or $x_{id\max}$, then $x_{id}(t+1)$ is limited to $x_{id\min}$ or $x_{id\max}$.

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1)$$

(1)

where $x_i$ is $i$th $x$ and $d = 1$ to $n$; $v_i$ is $i$th $v$; $d$ is dimension in $x_i$ or $v_i$ and $d = 1$ to $D$; $t$ is iteration; $x_{id}(t)$ is $d$th dimension in $i$th $x$ and in $t$ iteration; $v_{id}(t+1)$ is $d$th dimension in $i$th $v$ and in $t+1$ iteration; $x_{id}(t+1)$ is $d$th dimension in $i$th $x$ and in $t+1$ iteration; $n$ is number of particles.

$$v_{id}(t+1) = w \times v_{id}(t) + c_1 \times \text{Rand} \times [P_{id}(t) - x_{id}(t)] + c_2 \times \text{Rand} \times [G_{id}(t) - x_{id}(t)]$$

(2)

where $v_{id}(t)$ is $d$th dimension in $i$th $v$ and in $t$ iteration; $w$ is inertia weight and controls the magnitude of the old velocity $v_{id}(t)$ in the calculation of the new velocity; $P_{id}(t)$ (this $d$th dimension in $i$th, local best particle and in $t$ iteration; $G_{id}(t)$ is $d$th dimension global best particle in $t$ iteration; $c_1$ and $c_2$ determine the significance of $P_{id}(t)$ and $G_{id}(t)$; Rand is a uniformly distributed real random number within the range 0 to 1.

$$x_{id\min} \leq x_{id}(t+1) \leq x_{id\max}$$

(3)

where $x_i$ is $i$th $x$ and $d = 1$ to $n$; $d$ is dimension in $x_i$ and $d = 1$ to $D$; $t$ is iteration; $x_{id}(t+1)$ is $d$th dimension in $i$th $x$ and in $t+1$ iteration; $n$ is number of particles.

Step (5) Select best sites ($b$) from scout bees

Best sites are selected for each best bee, the total number of which equals one-quarter of the number of scout bees.

Step (6) Best bees start the PSO procedure using the NW Phitr iteration

In this step, new particle bees from elite and best bees are produced using Eq. (1). Elite and best bee velocity updates are acquired using Eq. (2). The NW technique improves PSO search efficiency, as shown in Eq. (3).

Step (7) Recruit random bees ($r$) for other visited sites

The random bees in the population are assigned randomly around the search space scouting for new potential solutions. The total number of random bees is one-quarter of the number of scout bees.

Step (8) Self-parameter-updating (SPU) for elite, best and random bees

Furthermore, in order to prevent being trapped into a local optimum in high dimensional problems, this study proposes a solution, i.e., the self-parameter-updating (SPU) technique, the idea for which came from Karaboga [17]. Eq. (4) shows the SPU equation.

$$x_{id(new)} = x_{id(cur)} + 2 \times (\text{Rand} \times 0.5) \times (x_{id(cur)} - x_{id(cur)})$$

(4)

where $x_i$ is $i$th $x$ and $i = 1$ to $n$; $d$ is dimension in $x_i$ and $d = 1$ to $D$; $x_{id}(cur)$ is $d$th dimension in $i$th $x$ and in current solution; $x_{id}(new)$ is $d$th dimension in $i$th $x$ and in new solution; Rand is a uniformly distributed real random number within the range 0 to 1; $j$ is the index of the solution chosen randomly from the colony as shows in Eq. (5), $k$ is the index of the dimension chosen randomly from the dimension as shows in Eq. (6); $n$ is number of scout bees.

In step (8), after elite, best and random bees have been distributed based on finesse, finesse are checked to determine whether they are to be abandoned or memorized using Eq. (4). Therefore, if finesse of elite, best or random bees are both improved using Eq. (4) and improved over previous finesse, the new finesse are memorized. In step (3) through step (8), this differential recruitment is a key operation of the PBA.

Step (9) Convergence?

In this step, only the bee with the highest fitness will be selected to form the next bee population. These steps are repeated until the stop criterion is met and bees are selected to be abandoned or memorized.

3 Mining Rules of Satellite Imagery

3.1 Experimental data

There are three features in the satellite imagery dataset include angular second moment (ASM), contrast (CON) and entropy (ENT). The three features both survey by four sources (twelve variables) include raw light, green light, infrared and red light. Thus, there are totally twelve input variables in the satellite imagery dataset. The outputs of the dataset are six different types of images include water, betel palm, building, cloud, orchard and wood. This study collected 600 experimental satellite imagery data, 200 data were randomly selected as the training set, and the remaining 400 data as the testing set [16]. All the variables were normalized into 0 to 1 by Eq. (7). The training set was employed to build the classification rules model and the testing set was employed to evaluate model generalizations. Table 1 and Figure 2 present some descriptive statistics of the satellite imagery dataset.
Table 1 Variables of satellite imagery

<table>
<thead>
<tr>
<th>Variables</th>
<th>Range</th>
<th>Unit</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green light raw survey source (G_SOURCE)</td>
<td>35~159</td>
<td>Pixel</td>
<td></td>
</tr>
<tr>
<td>Green light on second-order differential angular survey momentum (G_ASM)</td>
<td>21~100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green light on contrast survey source (G_CON)</td>
<td>0~352</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green light on entropy survey source (G_ENT)</td>
<td>0~69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrared raw survey source (I_SOURCE)</td>
<td>15~135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrared on second-order differential angular survey momentum (I_ASM)</td>
<td>21~100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrared on contrast survey source (I_CON)</td>
<td>0~469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrared on entropy survey source (I_ENT)</td>
<td>0~69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red light raw survey source (R_SOURCE)</td>
<td>25~171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red light on second-order differential angular survey momentum (R_ASM)</td>
<td>21~100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red light on contrast survey source (R_CON)</td>
<td>0~465</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red light on entropy survey source (R_ENT)</td>
<td>0~69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>0, 1</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Betel palm</td>
<td>0, 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building</td>
<td>0, 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cloud</td>
<td>0, 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orchard</td>
<td>0, 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>0, 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
X_{\text{new}} = \frac{X_{\text{old}} - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} (D_{\text{max}} - D_{\text{min}}) + D_{\text{min}} \quad (7)
\]

Where \(X_{\text{old}}\) is the \(X\) value before normalization; \(X_{\text{new}}\) is the maximum value of \(X\) value before normalization; \(X_{\text{min}}\) is the minimum value of \(X\) value before normalization; \(D_{\text{max}}\) is the maximum \(X\) value after normalization. This study setting \(D_{\text{max}}\) is 1; \(D_{\text{min}}\) is the minimum \(X\) value after normalization. This study setting \(D_{\text{min}}\) is 0; \(X_{\text{new}}\) is the \(X\) value after normalization.

3.2 Rules and encoding of operation tree

This study adopted classification tree to express classification rules and employed particle bee algorithm (PBA) to optimize the tree to produce self-organized rules. In this study, a five-layered classification tree was adopted, as shown in Figure 3. In Figure 3, variables \(X_1\) to \(X_{31}\) were external tree branch and variables \(K_1\) to \(K_{15}\) were internal tree branch. The external and internal tree branch encoding variables and constants are listed in Table 2 and Table 3, respectively. The encoding rule was designed to adhere to the following rules:

- The first to fourth layers of external tree branch (\(K_1\) to \(K_{15}\)) must be variables or constants. The encode must be between integer 1 to 12 (see Table 3).
- The first to fourth layers of internal tree branch (\(K_1\) to \(K_{15}\)) must be variables or constants. The encode must be between integer 13 to 25 (see Table 3). When the gene encoding is 25, it represents a constant, and a constant between 0 to 1.
- Between each layer, on the left of internal and external tree branch is smaller mathematical operator, on the right of internal and external tree branch is bigger or equal mathematical operators.

The fifth layer of internal tree branch (\(Y_1\) to \(Y_{15}\)) is determined by the classification result. For an example, if \(Y_1\) classification includes 20 water datasets, 1 wood dataset and 2 cloud datasets. The \(Y_1\) classification will be assigned for water classification.

Table 2 Encode of internal tree branch

<table>
<thead>
<tr>
<th>Encode</th>
<th>X_1~X_{15}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G_SOURCE</td>
</tr>
<tr>
<td>2</td>
<td>G_ASM</td>
</tr>
<tr>
<td>3</td>
<td>G_CON</td>
</tr>
<tr>
<td>4</td>
<td>G_ENT</td>
</tr>
<tr>
<td>5</td>
<td>I_SOURCE</td>
</tr>
<tr>
<td>6</td>
<td>I_ASM</td>
</tr>
<tr>
<td>7</td>
<td>I_CON</td>
</tr>
<tr>
<td>8</td>
<td>I_ENT</td>
</tr>
<tr>
<td>9</td>
<td>R_SOURCE</td>
</tr>
<tr>
<td>10</td>
<td>R_ASM</td>
</tr>
<tr>
<td>11</td>
<td>R_CON</td>
</tr>
<tr>
<td>12</td>
<td>R_ENT</td>
</tr>
</tbody>
</table>

Table 3 Encode of external tree branch

<table>
<thead>
<tr>
<th>Encode</th>
<th>K_1~K_{15}</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>G_SOURCE</td>
</tr>
<tr>
<td>14</td>
<td>G_ASM</td>
</tr>
<tr>
<td>15</td>
<td>G_CON</td>
</tr>
<tr>
<td>16</td>
<td>G_ENT</td>
</tr>
<tr>
<td>17</td>
<td>I_SOURCE</td>
</tr>
<tr>
<td>18</td>
<td>I_ASM</td>
</tr>
<tr>
<td>19</td>
<td>I_CON</td>
</tr>
<tr>
<td>20</td>
<td>I_ENT</td>
</tr>
<tr>
<td>21</td>
<td>R_SOURCE</td>
</tr>
<tr>
<td>22</td>
<td>R_ASM</td>
</tr>
<tr>
<td>23</td>
<td>R_CON</td>
</tr>
<tr>
<td>24</td>
<td>R_ENT</td>
</tr>
<tr>
<td>25</td>
<td>I</td>
</tr>
</tbody>
</table>

3.3 Fitness function and PBA parameters

The correct rate (CR) was used to evaluate the models accuracy. Therefore, this study focus on producing an accurate model to predict satellite imagery classification, the CR was adopted as the evaluation function (fitness function) of solutions. This study adopted PBA to optimize the classification tree to fit the data set to produce the self-organized classification rules. There are some parameters may affect the performance of PBA. Reference [14, 15] suggested the parameters following as Table 4. In this study, these parameters were determined according to maximizing the CR on the training set.
First layer

Second layer

Third layer

Forth layer

Fifth layer

Figure 3. Five layers of evolutionary classification tree

Table 4 Parameter values used in the experiments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n)</td>
<td>50</td>
</tr>
<tr>
<td>(e)</td>
<td>(n/2)</td>
</tr>
<tr>
<td>(b)</td>
<td>(n/4)</td>
</tr>
<tr>
<td>(r)</td>
<td>(n/4)</td>
</tr>
<tr>
<td>(w)</td>
<td>0.9–0.7</td>
</tr>
<tr>
<td>(v)</td>
<td>(X_{\text{min}}/10–X_{\text{max}}/10)</td>
</tr>
<tr>
<td>(P_{\text{eitr}})</td>
<td>15</td>
</tr>
<tr>
<td>(P_{\text{bitr}})</td>
<td>9</td>
</tr>
</tbody>
</table>

where \(n\) is population size (colony size); \(w\) is inertia weight; \(v\) is limit of velocity; \(e\) is elite bee number; \(b\) is best bee number; \(r\) is random bee number; \(P_{\text{eitr}}\) is PSO iteration of elite bees; \(P_{\text{bitr}}\) is PSO iteration of best bees.

4 Results

4.1 Evolutionary Classification Tree (ECT)

This study adopted particle bee algorithm (PBA) to optimize the operation evolutionary classification tree (ECT) to produce the self-organized classification rules. Figure 4 is the result of satellite imagery classification by ECT. The mining classification rules are as follow. The correct rate of training set and testing set as shown in Table 5 and Table 6. Thus, ECT not only can produce the satellite imagery classification but also can self-organized the classification rules.

Figure 4. Five layers of satellite imagery evolutionary classification tree

ECT mined 11 classifies rules as the followed:

RULE (1): IF \(G_{\text{SOU}} \geq 46.87\) AND \(I_{\text{SOU}} \geq 100.58\) THEN Cloud

RULE (2): IF \(G_{\text{SOU}} \geq 46.87\) AND \(I_{\text{SOU}} \geq 77.51\) AND \(I_{\text{SOU}} < 100.58\) THEN Building

RULE (3): IF \(G_{\text{SOU}} \geq 46.87\) AND \(I_{\text{SOU}} < 77.51\) AND \(I_{\text{SOU}} \geq 49.91\) THEN Orchard

RULE (4): IF \(G_{\text{SOU}} \geq 46.87\) AND \(I_{\text{SOU}} < 77.51\) AND \(I_{\text{SOU}} < 49.91\) THEN Water

RULE (5): IF \(G_{\text{SOU}} < 46.87\) AND \(R_{\text{SOU}} \geq 32.50\) AND \(I_{\text{ENT}} \geq G_{\text{ENT}}\) AND \(R_{\text{CON}} \geq 1.62\) THEN Building

RULE (6): IF \(G_{\text{SOU}} < 46.87\) AND \(R_{\text{SOU}} \geq 32.50\) AND \(I_{\text{ENT}} \geq G_{\text{ENT}}\) AND \(R_{\text{CON}} < 1.62\) THEN Betel palm
RULE (7): IF G_SOU < 46.87 AND R_SOU >= 32.50 AND I_ENT < G_ENT AND I_ENT >= 40.00 THEN Betel palm
RULE (8): IF G_SOU < 46.87 AND R_SOU >= 32.50 AND I_ENT < G_ENT AND I_ENT < 40.00 THEN Orchard
RULE (9): IF G_SOU < 46.87 AND R_SOU < 32.50 AND R_SOU >= 29.34 AND I_SOU >= 79.50 THEN Wood
RULE (10): IF G_SOU < 46.87 AND R_SOU < 32.50 AND R_SOU >= 29.34 AND I_SOU < 79.50 THEN Orchard
RULE (11): IF G_SOU < 46.87 AND R_SOU < 29.34 THEN Wood

Table 5 Training set of CR for ECT

<table>
<thead>
<tr>
<th>Training set</th>
<th>Actual classes</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td></td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td></td>
<td>0</td>
<td>17</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Building</td>
<td></td>
<td>0</td>
<td>3</td>
<td>32</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Cloud</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Orchard</td>
<td></td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>3</td>
</tr>
<tr>
<td>Wood</td>
<td></td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Correct rate</td>
<td></td>
<td>100.00%</td>
<td>50.00%</td>
<td>91.43%</td>
<td>96.97%</td>
<td>76.47%</td>
<td>62.50%</td>
</tr>
</tbody>
</table>

Table 6 Testing set of CR for ECT

<table>
<thead>
<tr>
<th>Testing set</th>
<th>Actual classes</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td></td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td></td>
<td>0</td>
<td>26</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>Building</td>
<td></td>
<td>0</td>
<td>15</td>
<td>51</td>
<td>5</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Cloud</td>
<td></td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>62</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Orchard</td>
<td></td>
<td>0</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>Wood</td>
<td></td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td>Correct rate</td>
<td></td>
<td>100.00%</td>
<td>39.39%</td>
<td>78.46%</td>
<td>92.54%</td>
<td>60.61%</td>
<td>54.41%</td>
</tr>
</tbody>
</table>

4.2 Classification Tree (CT)

This study adopted classification tree to self-organize the classification rules and to fit the data set of satellite imagery. The CR of training data and testing data of these models are presented in Table 7 and Table 8.

4.3 Back-propagation Networks (BPN)

This study adopted back-propagation neural networks (BPN) [16] to fit the data set of satellite imagery. In this study, network parameters such as number of hidden neurons, learning rate, momentum factor, and number of learning cycles were determined according to maximizing the CR on the testing dataset. The best network structure is the network with one hidden layer containing six hidden units; and the optimum learning parameters are 1.0 for learning parameters and 0.5 for momentum factor. The CR of training data and testing data of these models are presented in Table 9 and Table 10.

4.4 Support Vector Machine (SVM)

This study adopted nu-SVM to fit the data set of satellite imagery. In this study, Leave-One-Out (LOO) is the searching method for searching best value of nu and gamma according to maximizing the CR on the testing dataset. The best nu-SVM structure is the structure with nu=0.48 and gamma=0.0625. The CR of training data and testing data of these models are presented in Table 11 and Table 12.
Table 9 Training set of CR for BPN

<table>
<thead>
<tr>
<th>Training set</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict classifies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td>23</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Building</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Cloud</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>Orchard</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>25</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Correct rate</td>
<td>100.00%</td>
<td>67.65%</td>
<td>82.86%</td>
<td>90.91%</td>
<td>73.53%</td>
<td>81.25%</td>
</tr>
</tbody>
</table>

Table 10 Testing set of CR for BPN

<table>
<thead>
<tr>
<th>Testing set</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict classifies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td>29</td>
<td>10</td>
<td>3</td>
<td>62</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Building</td>
<td>10</td>
<td>19</td>
<td>10</td>
<td>46</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Cloud</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>13</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Orchard</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Correct rate</td>
<td>100.00%</td>
<td>43.94%</td>
<td>56.92%</td>
<td>92.54%</td>
<td>69.70%</td>
<td>73.53%</td>
</tr>
</tbody>
</table>

Table 11 Training set of CR for SVM

<table>
<thead>
<tr>
<th>Training set</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict classifies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td>29</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Building</td>
<td>2</td>
<td>32</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cloud</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Orchard</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td>Correct rate</td>
<td>100.00%</td>
<td>85.29%</td>
<td>91.43%</td>
<td>84.85%</td>
<td>82.35%</td>
<td>81.25%</td>
</tr>
</tbody>
</table>

Table 12 Testing set of CR for SVM

<table>
<thead>
<tr>
<th>Testing set</th>
<th>Water</th>
<th>Betel palm</th>
<th>Building</th>
<th>Cloud</th>
<th>Orchard</th>
<th>Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict classifies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Betel palm</td>
<td>28</td>
<td>7</td>
<td>41</td>
<td>7</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Building</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Cloud</td>
<td>0</td>
<td>12</td>
<td>10</td>
<td>0</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td>Orchard</td>
<td>19</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Correct rate</td>
<td>100.00%</td>
<td>42.42%</td>
<td>63.08%</td>
<td>89.55%</td>
<td>60.61%</td>
<td>63.24%</td>
</tr>
</tbody>
</table>

The model accuracy and understand ability comparison between those four methods are shown in Table 13 to Table 15 and Figure 5. Highest model accuracy (correct rate) means model has the better forecasting ability. Model understand ability means model can produce the rules for understanding and explaining by user. The result of Table 13 to Table 15 and Figure 5 as following:

- **Model accuracy**: ECT produces the satellite imagery classification model which model accuracy only lower than BPN but better than CT and SVM.

- **Model understand ability**: ECT and CT can produce the self-organized explicit classification rules but BPN and SVM cannot.

Thus, ECT not only can produce accurate satellite imagery classification model but also can self-organized the classification rules.
The results showed that ECT can produce explicit rules which are more accurate than CT and SVM but less accurate than BPN model. However, BPN is black box models, while ECT can produce explicit rules which are an important advantage to mining the explicit rules and knowledge in practical applications.

If the user requirement is understandable satellite image classification models and rules, then ECT is an method which can produce accurate self-organized classification models and rules; If the user understand ability for satellite image classification model is not very important, then BPN is a more accurate and rapid method which can establish satellite image classification model.

Acknowledgements

This study was partially supported by the National Natural Science Foundation of China, under the Grant No. 51308120.

References

Roadwork Site 3D Virtual Visualization Using Open Source Game Engine and Open Information Transfer

T. Makkonen, R. Heikkilä, A. Kaaranka, K. Nevala
Construction Technology Research Center, University of Oulu, Finland
E-mail: tomi.makkonen@oulu.fi, rauno.heikkila@oulu.fi, annemari.kaaranka@oulu.fi, kalervo.nevala@oulu.fi

Abstract
Machine control systems generate huge amounts of data to be used to build smarter. This information is still not widely used as there are no general methods to access this massive spread of data as these are strictely closed systems. Main tools in our study were an open source 3D game and simulation engine for visual 3D presentation (Panda3D), a machine control system on an excavator which was opened by manufacturer for data transfer (Novatron) and a communication link between these main components Extensible Messaging and Presence Protocol (XMPP). Visualization was performed in 3D presenting model of the road layers and the excavator working on the site using GPS positioning transferred in XML through XMPP. The result shows how to use open source tools to successfully manage data transfer and data presentation using easily accessible open source software and also show potential what these tools could provide for construction software development. For future we hope that every machine control system used will send the joint data through internet and work file is sent automatically to the machines.

There is a catch however. These systems are usually closed in nature, for example Topcon’s systems are planned to work only with their products. For this approach to work every machine on the site should use only one brand of machine control system, including subcontractors. Infrakit [3] demonstrates how small company can rapidly bring new innovations to practical use on construction site when communications, mainly machine control systems, are opened. Another good example is SVS innovations’ Field3D [4] which rabidly brought visualization of building construction project to tablets, a task only possible because of open IFC standard [5].

The work here studies how hard and what is needed to set up communication channel through internet and link between machine control system and open source 3D game engine, to enable real-time visualization of an excavator working on the site. Real-time 3D construction machine visualization is widely studied [6],[7], [8].

For added depth, we study the possibilities of building new kind of tools easily for construction industry, a rude 4D visualization of the work schedule of a road site pre construction is performed. For other 4D related construction visualizations following resources are presented [9], [10], [11], [12].

1 Introduction

Machine control systems for construction machines, like those used in excavators, can be seen as a passive information presenters for a driver, but fast progress on implementing new tools are in a away. Topcon has a very advanced product sitelink3D [1] which, for example, uses machine control system data to show graphically the difference between planned and processed surface. Novatron’s system [2] can send the joint data through internet and work file is sent automatically to the machines.

Keywords
open data transfer; 3D visualizations; 4D visualization; machine control system

2 Roadwork visualization

2.1 Experiment setup
Main components used in this research:

1. Extensible Messaging and Presence Protocol (XMPP), originally called Jabber, for information exchange, open standard
2. Machine control system Novatron Vision3D for an excavator, real time communications opened using XMPP.
3. Panda3D game engine, programmed with python language, open source
4. SLeekXMPP a python XMPP library, MIT license
5. Eclipse python programming environment, open source
6. Blender 3D computer graphics software, open source
7. Matlab, proprietary software
8. Tekla Civil, proprietary software

2.1.1 Building communications

The machine control system was reprogrammed by the company Novatron to send data using Extensible Messaging and Presence Protocol (XMPP) [13]. Protocol is build using XML and the form is decentralized client – server architecture. For the work here the most important transferred information is GPS position data, but also data about the construction project like vehicles on the project and organization are included.

On Eclipse IDE SDK 3.7 with python, communication software was programmed using SleekXMPP library [14]. The main components of the program are:
1. Handshaking
2. Registering to a project with a password
3. Reading GPS position data from the wanted construction machine
4. Relaying the GPS position data for virtualization

2.1.2 Building visualization

A 3D-road with four layers and a second one with one layer was constructed using Tekla Civil. The lengths of the road were 50 m and 30 m and the width about 5 m. For visualization the layers were divided to shorter pieces as it would give more possibilities to present visual information. For example area under work can be highlighted in more detailed manner than just the whole layer.

The dividing the road layers to shorter pieces is done semi automatically. From Tekla Civil DXF-file was exported to AutoCAD where road central line was divided to short intervals. The more logical choice of the file format open standard Inframodel [15] was omitted because in theory all the tools to build models directly with DXF were in use without extra programming, however complications arise and open standard should have been used. DXF-file with layers and the central line points where then imported to Matlab where algorithm was created to divide layers to triangle groups. In Figure 1 we can see a road layer regrouping: Red rhombus marks a centre points of triangles forming a layer, green lines perpendicular to centreline ending to red x marks the selection box, centreline is formed using points acquired from AutoCAD and are used to determine the length of the layer triangle selection box.

To present road layers in game engine Panda3D [16] accurately build-in data format GeomVertexData was used. Using the import tools through Blender [17] to Panda3D resulted not eye candy enough results, as seen in Figure 2, where top we see a tool chain import to Panda3D and bottom use of build-in data format GeomVertexData.

At the time of work 3D objects were taken from the Google SketchUp’s 3D Warehouse which is today called Triple SketchUp 3D Warehouse [18]. Using COLLADA [19] file format these models were then
imported to Blender for animation and for export to Panda3D in egg format [20]. For a ground model an aerial flat picture was used.

2.2 Implemented visualizations

2.2.1 Excavator position at the site

Excavator working at the site is visualized in 3D. The machine position is moving according the GPS position data received from the machine control system. The test was performed at the Oulu University Construction Automation Test Site. Data was updated every 30 seconds. Excavator was animated to do constant digging movements to please the eye. The program includes an option to change colour to match the current state of the machine, for example red = broken, green = working and orange = on a pause and also colouring the road. At the actual test this was not done however, as the machine control system, at the time of the test, did not support sending status data.

2.2.2 3D objects on the site

There are objects on the road site whose location is important to know when flow of the work is monitored. For visualization freight container as storage for equipment and pile of gravel as buffer storage were included, in Figure 4 blue boxes and brown pile of rocks. These objects can be anything what is decided to be important as adding them is relatively easy task.

2.2.3 4D Timeline visualization – in design phase

This is tool to go through work schedule visually before actual work is done. The idea is to use construction time schedule and visualize in 3D what is planned for the site. This should help the person who makes the schedule to find possible bottlenecks and errors in timetable.

From construction timeline a schedule matrix was created and then used to run visualization. Table shows the components of the matrix. There is one schedule matrix for every individual machine (Machine) which includes start and stop times for planned operation (Task) in selected road layer (Layer) in selected work area (Span).

Table 1. Schedule Matrix – simplified example

<table>
<thead>
<tr>
<th>Machine 1</th>
<th>Task 1</th>
<th>Task 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>Stop</td>
</tr>
<tr>
<td></td>
<td>Span</td>
<td>Layer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Day 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Schedule matrix generation was not automatized, but doing so from project management software was studied and found possible, but was seen unnecessary to this demonstration.

Figure 4 shows a grader driving according to schedule matrix. The green areas of the road layer are ones where the work has been already done and red area is the one where the current work is performed, and white marks layers and areas where work is not done yet. Visualizations were performed to also to excavators (Figure 5) or multiple different machines at the same time as there is no limit how many or what machines are working on the site.
3 Results

Extensible Messaging and Presence Protocol was successfully used to transfer position data through internet from the machine control system. This information was then merged to the 3D visualization using open source game engine Panda3D, so that anyone with internet access and the made software, regardless the distance to the work site, can track the position of the excavator.

Because of the work of the manufacturer, Novatron, this task was relatively easy to implement. The machine control system will send the information to the cloud and our task was implement procedures to read the XMMP message using readily available tools, like SleekXMPP library. All main components were using python language making the work even easier.

The negative side of the small open source software was the neglected documentation and possibility to get support right then when needed. On the positive side, support by the users was always there, if sometimes a bit late and the technical skill level of the support was in high level. Main stream open source software like Eclipse and Blender were easy to get support.

Timetable visualization to help to find errors and bottlenecks from the project plan was performed in 3D from the planned work schedule. Graders and excavators were moving on the road work site simulation according to the schedule. Road layers were changing color according the current work phase.

Our work showed that rude 4D simulation using game engine is straight forward engineering work. The object oriented programming made scalability for different and also for multiple construction machines a somewhat easy task.

Finally a group of 3D objects were added to point the important locations on the site. Freight container demonstrates equipment storage and pile of gravel works as a buffer storage. Adding 3D objects is a relatively easy task as long as one has a 3D model ready for it.

4 Discussion

By opening machine control system data transfer to us to use, in fast phase, we were able to present the position of the excavator real time in 3D visualization running on the open source game engine Panda3D [21]. The data was transferred using Extensible Messaging and Presence Protocol [13] using SleekXMPP [14] library.

Big machine control system manufacturers are keeping the data for themselves, building a business model over the idea that every machine on the site is equipped with their product. This is valid point, but misses the fact that even if the main contractor, in some cases, is fully equipped with only one brand tools the subcontractors are usually not.

If the data is opened for small companies there is a chance that technology will develop really fast. Not only the large manufacturer’s software for managing whole constructions digitally would be possible, but there would also be room for smaller companies to fill the caps. At the moment product like Infrakit [3] is working in integration with some machine control systems (Novatron [2], DigPilot [22], Prolec [23]), but is relying on file transfers with other manufacturers. Manual transfers are not real-time, laborious and also leaves room for errors. Gained automation level rise is also not that high when compared to integrated systems.

A rude 4D schedule simulation using open source game engine was performed. From a work schedule a schedule matrix (Table 1) was created, which was then used to drive construction equipment in virtual 3D work site. The virtual site was constructed using layered road model, 3D objects for interesting subjects we wanted to know location for and a flat aerial photo.

The used open source programs were adequate to the job in hand. The grief we had with them, was the not so perfect documentation, which is not so uncommon complain with small software [24].

We suggest that real-time machine control system data should be integrated openly to the building information management process. How exactly one does that is a topic for further study.
References


The Implementation of BIM
In a Large European Construction Company

Ger Maas

Abstract -

Background: Information Technology (IT) has reached a level of maturity, where IT can really meet the complexity of the construction industry. IT has reached the ability to support construction not just any more in the facilitating processes as financial services and human resources management (HR), but also to play a key role in the primary process due to the development of the building information modeling (BIM). However, implementation of BIM in a company is still a challenging exercise. Purpose of this paper: The development of an implementation strategy based on the needs of project support. Method: Desk research gives insights in the roles and phases in a project, levels of hierarchy in a company and two case studies in two large European construction companies deliver insight in the actual situation. Results & Discussion: The complexity of BIM implementation will be presented in a clear analysis of BIM for the different roles, activities, phases of an building project. Moreover a BIM implementation approach will be presented useful for small and large construction companies.

Keywords -
Building and Architecture, BIM, Construction

1 Introduction

Latest IT such as BIM is rapidly spreading to architecture, civil engineering, engineering and maintenance with the aim of simplifying complex system in buildings and facilities, minimizing operation cost, providing high value services and advancing Facility Management. In particular, BIM-based management framework design to assist the development of sensor information operation system, which adds convenience to facility maintenance and management, has emerged as a key issue.

In addition, a technology for real-time connectivity of geographic information systems (GIS) with building information modeling (BIM) linking technology, sensor information and three-dimensional spatial information in smart construction operation and management is necessary.
Figure 1 shows the potential role that a BIM tool can play in the life cycle of a building more than just replacing the traditional 2D drawings. It supports information management and decision making during the whole life of the building.

The survey of Lui and Issa [3] shows results, that indicate that the industry practitioners believed that maintainability issues should be considered in the design and construction phases.

BIM has changed the way the Architecture, Engineering and Construction industry (AEC) industry communicates and cooperates. Knowledge sharing between the facility management and design professionals has become possible with BIM. BIM technology has been used effectively in the design and construction phases. There is a need to expand BIM beyond the design and construction phases and to consider using BIM for facility management such as in maintenance activities. However, the research on BIM use for Facility Management is lagging behind the study of BIM in design and construction phases.

Maintenance costs, although the largest cost over a building’s life cycle, are currently rarely considered in the early design phase. Some design errors that make maintenance activities impossible to perform are always hard to visualize in the design phase. As the next advancement for Facility Management (FM), design for maintenance (D4M) should be considered in the early design phase.

The authors in [4] presented the use of BIM in design optimization and design-scenario development and the development of an automated Decision Support System (DSS) for optimizing the selection of the best design according the Leadership in Energy and Environmental Design index (LEED). This LEED index has an upgrade for decisions in Existing Buildings (LEED-EB). The DSS provides decision makers with the flexibility to minimize the required total upgrade costs to achieve a specified LEED-EB certification level such as gold or silver; or maximize the number of LEED-EB points that can be achieved within a specified limited budget. The developed DSS utilized linear programming to perform the optimization computations because of its guarantee to generate a global optimal solution and its reasonable computational time and effort compared to other optimization techniques. An application example was analyzed to illustrate the use of the developed DSS and evaluate its performance. The developed DSS was able to identify the optimal upgrade decisions for minimizing total upgrade costs for achieving Certified and Silver LEED-EB levels. Furthermore, the DSS was able to identify the optimal upgrade decisions for maximizing the number of LEED-EB points within a range of specified upgrade budgets. The DSS offers unique and important capabilities to aid decision makers in achieving the highest benefits for upgrading their buildings within the specified budgets. It provides a practical tool to evaluate and optimize various green upgrade options effectively and efficiently.[4].

2 Case studies

2.1 NCC Sweden

Patrick Lindvall, NCC Construction Group Sweden, reported the strategies adopted by NCC and introduces how they have integrated corporate and ‘legacy data’ into the BIM for use during the Build Phase.

NCC is today implementing the internal services and data layers required to reach maturity Level 3 BIM, ensuring they have the infrastructure to support the organizations use of Virtual Design and Construction not only today but tomorrow as well.

2.2 BAM, The Netherlands

BAM’s vision on BIM is that with BIM the information of a building project will be secured in one or more 3D models and databases. By relating the 3D models with the dimensions, time, budget requirements, maintenance data, etc. the information of various expertises will be stored unambiguously BIM created.
BIM creates opportunities for collaboration during the whole life cycle of the project.

Figure 3: BIM in the life cycle

Figure 2 shows the life cycle of a building from program to design, construction, demolition or renovation and re-programming again. The circle isn’t closed yet, but efforts are made to improve the use of BIM and to increase the value during the whole life cycle. At the moment most of the influence of BIM has been limited to design and execution.

BIM supports on the projects:
- To develop opportunities and to limit risks;
- To create insights in complex structures and bottle necks;
- To promote integral and multidisciplinary working;
- To collaborate simultaneously;
- To reduce failure costs;
- To shorten the project time of the building process;
- To enlarge the accuracy;
- To improve interface management;
- To save resources (paper).

Some projects examples are:

At Leeds Arena (UK) some improvements were achieved:
- Reduction of failure costs by interface management;
- Saving of manhours by more efficient working;
- Limiting the number of 2D drawings by using 3D models;

The use of BIM in project management and execution will be extended, when a contractor has a long time commitment to a project or if the owner likes to use design and construction data during the life span of the use of a building. The next use of BIM is implemented now:
- Simulation of the use phase by comparing energy costs and maintenance efforts as a consequence of different design solutions.
- Simulation of different design and construction solutions and their consequences of aspects as logistics and cash flow.
- 100% validation and verification at the moment of handover, transparent for the owner. It avoids hidden failures, that have to be solved during the operation phase.
- By linking maintenance management systems is real time reporting on key performance indicators (KPI) possible.
- As build data with output specifications and technical specifications are conditionally for a quick response in case of a failure during the use and operation.

Generally the following BIM functions are usefully used in Royal BAM Group companies: visualization, engineering, 3D reinforcement, 4D simulations, clash detection, measurements, global positioning systems (GPS), steering of equipment, quantity calculations, CAD-CAM connection.

The use of BIM needs initiatives at 4 levels in the company:
1. The employee
2. The departments
3. The group companies and
4. The company as a whole.

At the level of individual employees and at the level of departments applications for the following activities have to be developed: planning, design, scheduling, cost estimation, work preparation, execution, operation and maintenance.

Facilitating effort has to be spent at the level of group companies: development of BIM techniques, knowledge and experiences, implementation in projects, influence at HR and investment in ICT.

At group level: overarching facilitation is necessary as: development of BIM techniques, sharing knowledge and experiences and ICT investments. Every hierarchical level has its own distinguished responsibilities. The initiatives at the different levels have to be related very carefully.

3 Proposal for BIM implementation program

The NCC BIM development program as reported [5] presents 3 steps.

In case of BAM a few mismatches with the 3 steps development program can be reported from this BAM case study.

Firstly in a large company the development speed is not equal in all countries where the company has activities. Market circumstances, the involvement of other parties in the supply chain and local regulations influence the implementation speed of BIM as well as the possibilities to use BIM on site.

Secondly: the more a company has been involved in the life cycle of a project, the more useful it is to invest in BIM applications. The traditional use of BIM for design support is usually the start, but more and more you can...
see the link of the need of Facility Management data to the design phase. The gathered data from Facility Management can be useful as design input as well. It creates the option to make scenario analyses in the early stage of a project. This needs to be added to the NCC development steps.

In case of BAM the following phases of the implementation of BIM can be discovered.

The first step is always the use of 3D models for visualization, coordination and engineering.

Secondly the use of an information model for quantities, cost estimation, save specifications and verification can be introduced.

Thirdly documents with its build information ready for maintenance support can be created.

Fourthly model based scheduling, planning and validation of requirements is reachable.

The fifth phase is to use full size BIM models to create simulations for different aims such as optimizing cash flow, performance requirements and logistics on site.

The last phase is to have fully integrated BIM support in your project. All information that people need in the life cycle can be extracted from the BIM model. To steer the implementation of all BIM activities the following process is used in the BAM case.

<table>
<thead>
<tr>
<th>Year</th>
<th>3D Visualization</th>
<th>Coordination</th>
<th>4D Estimation</th>
<th>Specifications</th>
<th>Handover</th>
<th>BIM</th>
<th>As-Build</th>
<th>Model Based 60 Planning</th>
<th>Progress Monitoring</th>
<th>Validation</th>
<th>Simulation to optimize cash flow</th>
<th>Performance Support</th>
<th>Full Integrated BIM Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2019</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>2018</td>
<td>100</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>2017</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2016</td>
<td>100</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>2015</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2014</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
</tbody>
</table>

Figure 4: BIM implementation process of BAM (2013)

Facilitation of the implementation is necessary at group company and at group level to share knowledge and experiences and to decide on ICT investments.

The differences in complexity of successive BIM modules delivers a guidance in the order of the implementation process. The more mature, employees and organizations are in the use of BIM, the further they can go in the implementation and use of BIM tools from 3D visualization and design through cost and quantity estimation and as build information to model based planning, design simulation and full integrated BIM support.

References


4 Lessons

It creates transparency in a company to distinct activities in BIM development process for the different hierarchical levels in a company. The employee and his/her department have to do a lot of development work for all modules of BIM to integrate these modules in other company systems.
Visualization of As-built Progress Data using Construction Site Photographs: Two Case Studies

H.Jadidi and, M.Ravanshadnia, M.H.Alipour

Department of Construction Engineering and Management, Science and Research Branch, Islamic Azad University, Tehran, Iran

Department of Construction Management, Shahid Beheshti University, Tehran, Iran

E-mail: hossein.jadidi@gmail.com, ravanshadnia@gmail.com, m-hosseinalipour@sbu.ac.ir

Abstract

The collection of as-built data for construction progress monitoring remains challenging. This paper develops two case studies on image-based modeling, in which point cloud models are created based on the photo collection of the construction site. The first case considers 399 unordered construction images previously taken for purposes other than progress monitoring, whereas the second case considers 118 photos that have been taken based on the results of the first case study. The results of the first case study are employed to improve the quality of the point cloud model in the second case, using the site photo collection captured by the first author for the purpose of establishing an enhanced point cloud model. The results of the two case studies are compared. Furthermore, the results are compared with those of other researchers and found that they are in a good agreement with other reported results. Finally, some suggestions are proposed to improve the image-based model for construction progress monitoring, particularly for industrial projects that involve a large construction site and various work packages.

Keyword

Construction site photographs, Image-based modelling, Visualization, Progress monitoring, Decision making, Information technology; Huge construction sites

1 Introduction

Decision making during the construction phase significantly depends on accessible as-built and as-planned information. Daily construction site photographs are robust sources of as-built data that can be easily captured by either the construction manager or any site staff member. A 3D image-based model can also be automatically created daily using computer vision algorithm and image-processing techniques. Such models can aid in the visualization of discrepancies between as-built and as-planned data in an augmented reality environment, which facilitates progress monitoring. A project manager’s effective decision making in selecting corrective actions during the construction phase significantly depends on the immediate detection of schedule delay, and such corrective actions can prevent delays and budget deficiencies [1]. Figure 1 shows that the project control process mainly occurs during the construction phase. This process comprises three steps, namely, monitoring, comparing, and corrective action selection. In the first and second steps, members of the project management team prepare information products that visualize important data to facilitate decision making by the project manager. Figure 1 also shows various graphs (e.g., gaunt charts or s-curves) and images that can be used to present the status of the construction process. From it, we can thus determine whether such material can help the project manager to immediately identify discrepancies between actual and as-planned performance.

![Figure 1. Using information products for decision making](image-url)
Several researchers highlighted the limitations of current manual data collection approaches in terms of speed and accuracy. According to Akinci et al. (2006), field staffs in construction sites spend 30% to 50% of their time recording and analysing field data [2]. Moreover, data transfer from a site to a field office requires additional time, because most data items are not captured digitally [2]. Daily construction site photographs, which are robust sources of as-built data for construction progress monitoring, comprise a usable and easily accessible source of information for as-built progress data [1]. Digital images can be easily captured without additional cost for construction projects [3]. According to Section 4.2.1-b of the FIDIC series book (Red Book)—Conditions of Contract for Construction, photographs are among the progress reporting requirements that a contractor should regularly (e.g., monthly) send to the owner [4]. Aside from progress documents, photographs have more applications, and can be easily captured using a handheld digital camera by a contractor staff, construction manager, superintendent, the owner’s representative, the subcontractor, or other project team members. In this approach, collections of photographs are used to reconstruct the 3D as-built scenes using computer vision algorithm and image processing techniques [1], [3].

This paper focuses on the creation of 3D as-built point cloud models using construction site photographs. By registering these 3D models on a 4D as-planned model, the progress of construction projects can be visualized in an image.

2 Background

In 2006, augmented reality was proposed as a technique for the visualization of construction progress monitoring [5]. The use of augmented reality enables the visualization of construction progress in an image by superimposing a 3D model on the actual construction scene, and then highlighting discrepancies from the schedule by color coding any part of a structure. Golparvar-Fard et al. (2009) then proposed a visualization system called 4D Augmented Reality (DAR) for the automatic visualization of construction progress monitoring [1]. In this system, daily site photographs are captured using a digital handheld camera by anyone involved in a construction project [1]. In DAR, the location of a photographer and the orientation of each camera are computed based on the images using computer vision algorithm and image processing techniques. The generated 3D image-based model is then used as an overlay on the 4D as-planned model for the visualization of progress monitoring; here, color coding makes it easier for the user to understand what the model represented [1]. Finally, Golparvar-Fard et al. (2011) reported that the identification, processing, and communication of progress discrepancies are enhanced by the integration of the visualization of as-built and as-planned performance and can thus serve as a powerful remote project management tool for remote decision making in the A/E/C and FM industries [3].

3 Computer vision techniques for image-based modeling (IBM)

Several computer vision techniques for IBM have recently been used to create 3D models from a collection of input images (i.e., unordered daily construction site photographs in our study). In this approach, the locations of the photographer who captures the images are unidentified, and images are captured under various illumination, resolution, zoom, and quality conditions [6], [7], [8]. Then, correspondences between images should be estimated for scene reconstruction from an image collection. The goal of correspondence estimation relative to construction progress monitoring or to any use of the image collection is to take a raw set of images and then identify sets of matching 2D pixels across all the images [6]. Each set of matching pixels represents a single point in 3D [7]. For correspondence estimation, the distinctive local features of each image are initially identified, after which similar-looking features in different images are determined [6], [7]. Once the correspondence problem is solved, the structure from motion (SFM) procedure is used to estimate the location of the camera and 3D points [1], [6], [7], [8]. The SFM procedure studies both structure (i.e., 3D view of the construction site) and motion (i.e., motion of the camera within the construction scene) [6]. SFM estimates the extrinsic and intrinsic parameters of a single image pair [8]. Thus, the process must start with an ideal initial image pair with good estimates of camera parameters for the chosen pair [1], [8]. In the current paper, the initial image pair is selected manually. In our case study, the sparse model becomes vague and difficult to understand when the inappropriate image pair is selected. To estimate the intrinsic parameters of a camera, the focal length must be extracted from the exchangeable image file format (EXIF) tags of JPEG images to initialize the focal length of the new camera [6], [8]. In this paper, the sizes of original images are changed using Xnview [9] software to maintain the EXIF tags. All computer vision techniques are applied using VisualSIM software [10], [11]. We use the SIFT of Lowe [12] for feature detection as well as the functions available in
VisualSfM [10],[11] for matching and SfM. The process of capturing images and as-built point cloud model are shown schematically in Figure 2.

Determining correspondence matches between each image pair
Feature detection of each photo
Running the SfM procedure

Photography by construction field staff
Daily photo collection

Figure 2. Process of collecting as-built data resulting in the reconstructed 3D as-built point cloud model

4 Case study

Two case studies are implemented. The first uses an unordered photo collection previously captured for purposes other than progress monitoring and image-based modeling.

The photographs in this collection are captured randomly from any part of the construction site for documentation and for presentation in weekly and monthly progress reports. Some of these photographs are unrelated to project management tasks or are captured under poor conditions. For instance, some are taken from afar, or with the glare of the sun. More importantly, some are captured without any overlay, which is needed for good image-based modeling.

The second study uses the results of the first study and the images captured by the first author for the purpose of progress monitoring. This case study is conducted to investigate the quality of the point cloud model, with the aim of using the results for enhanced image capture and, consequently, improved IBM. In this section, a collection of images for a recent project is used to establish an as-built point cloud model of the construction scene. In this case, the photographs captured the scenes related to project management tasks only. Moreover, the photographer tried to capture photographs with a good overlay to achieve proper image-based modeling. Figure 3 shows the development of the models for these studies.

Figure 3. Development of the models for the two case studies

4.1 Case study 1

The project is the construction of a gas compressor station. The construction site covers approximately 20 ha, and the image collection includes approximately 3000 photos. A cutoff date is chosen, and images captured after this date are eliminated, resulting in a final number of 399 remaining images. These images were captured with different resolutions and under different illumination conditions. Furthermore, different staff members captured these images using different digital cameras.

Many of the images in the collection are not relevant to project management tasks. Moreover, the photographer was not able to choose the appropriate frame to capture the photographs of the scenes related to project management tasks. Figures 4 and 5 show examples of faulty photography. Figure 4 shows an image with the minimum number of SIFT features in the collection, with the location visualized by red points. This photograph is captured far from the structure and under the glare of the sun, which caused a significant decrease in the number of SIFT features. Figure 5 shows an image sample with an inappropriate frame of
photography. This image has the maximum number of SIFT features in the collection, but many of these features (which can be seen on the ground) are not used for image-based modeling for the purpose of project progress monitoring.

This paper uses VisualSfM software [10], [11] to solve the correspondence problem and bundle adjustment for the 3D sparse modeling of a construction scene. The first step is feature detection, which is implemented using the SIFT algorithm [12]. Figure 6 shows the number of SIFT features determined by the SIFT detector [12] for each image. As can be seen, the tolerance between the maximum and minimum numbers of features is 15783. The average number of detected SIFT features is 5904. Table 1 shows the maximum and minimum values of SIFT features. These values and the domain between them show that the scenes chosen by different photographers are vastly different, disproportional, and lack rational dependency.

<table>
<thead>
<tr>
<th>Image no.</th>
<th>No. of SIFT features</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>462 (min)</td>
</tr>
<tr>
<td>264</td>
<td>16245 (max)</td>
</tr>
</tbody>
</table>

The images and their feature matrix are then imported to VisualSfM [10],[11], which computes for the missing matches. Figure 7 shows one image pair from the collection with the correspondence matches which are shown by solid lines.

Finally, the appropriate initial image pair is manually selected, after which the 3D reconstruction of the construction scene is established by running the SfM procedure. The reconstructed sparse scene with the estimated camera location is shown in Figure 8.
The reconstructed scene, shown in the Figure 9, has relative coordination, such that the absolute coordinates of these points can be measured using SfM Georef v.2.3 [13] by defining coordination points (X,Y,Z) for at least three points. This transformed model is then overlaid on the station plan shown in Figure 10. According to Figure 10, the reconstructed scene covers only a small area of the site, although the photo collection includes a scene representing the entire construction site. The recall characteristic (portion of the number of images used to number of all images) for this photo collection is 0.25, a value that is lower than that obtained by previous studies conducted by Golparvar-Fard et al. (2011) [3].

The low quality of this 3D sparse model can be attributed to the small number of images relative to the large area of the construction site (20 hectares) for the gas compressor station. In addition, failure to overlay the images properly hinders the establishment of a good image-based model. Figure 4 and Figure 5 show that some images have been captured in poor conditions (e.g., with the glare of the sun) or from a far distance or with a poor selection of frame, resulting in a small number of SIFT features or a huge amount that is not relevant to the project management scope. As a result, few correspondence matches are obtained, thus leading to a low-quality 3D sparse model.

4.2 Case study 2

After performing Case 1 and confirming the low quality of the 3D point cloud model (recall=0.25), the construction site used for other project is the same as that used in Case 1 is divided into a large number of small areas. For example, industrial projects, such as gas compressor stations, involve various building and work packages (e.g., buildings, piping area and air coolers). In Case 2, we focus on one of the buildings (control building) with the size of 17 m x 40. A total of 118 photographs are captured in 6 min, and the conditions under which these images have been captured are shown in Table 2. These images have been captured for IBM with the aim of a proper overlay between images, which is in contrast to the randomly captured Case 1 images. According to McCoy et al. (2012), the images were captured approximately 9 m from the building to obtain good results for IBM [14].
Figure 11 shows the number of SIFT features for all 118 photos. According to Figure 11, the tolerance between the maximum and minimum numbers of features is 10752, which is less than the tolerance between the maximum and minimum numbers of SIFT features in the Case 1. This condition indicates that the method used to capture the scenes is the key factor for creating an enhanced point cloud model. If we use a larger number of photographs in Case 1, an improved model can be obtained, and recall can be increased. However, the process time required to identify features and solve the correspondence problem is increased. The average number of detected SIFT features is 12457, considerably higher than the average number in the Case 1. This finding indicates that in Case 2, we have more features on average, such that we expect to have increased correspondence among all photos in the collection and, ultimately, an improved point cloud model. The values listed in Table 3 show the maximum and minimum numbers of SIFT features. Figure 12, 13 show these images with the visualized SIFT feature location.

Table 2. Characteristics of the captured images

<table>
<thead>
<tr>
<th>Photography Characteristic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera model</td>
<td>Nikon Coolpix P510</td>
</tr>
<tr>
<td>Number of captured photos</td>
<td>118</td>
</tr>
<tr>
<td>Image resolution(pixel)</td>
<td>4608 x 3456</td>
</tr>
<tr>
<td>Photography duration</td>
<td>6 minutes</td>
</tr>
<tr>
<td>Lighting condition</td>
<td>Sunny</td>
</tr>
</tbody>
</table>

Table 3. Maximum and minimum numbers of SIFT features in the image collection (Case study 2)

<table>
<thead>
<tr>
<th>Image no.</th>
<th>No. of SIFT features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6645 (min)</td>
</tr>
<tr>
<td>264</td>
<td>17397 (max)</td>
</tr>
</tbody>
</table>

Figure 12. The image in the collection with minimum number of SIFT features location visualized by red points

Figure 13. The image in the collection with maximum number of SIFT features location visualized by red points

Figure 14 shows the reconstructed point cloud model with a reconstructed camera denoted by 100 frusta. Figure 15 shows the 3D sparse model in dense form, as determined by the CMVS/PMVS module in VisualSfM software [10],[11].
5 Conclusion

As-built progress data collection is among the most challenging tasks in progress monitoring. However, daily construction site photographs can be used as a robust source of as-built progress data. This paper develops two case studies, and an as-built point cloud model is created for both cases. The results of the first study are used to improve the second case. The results of the 3D point cloud model in both cases are then compared, from which we have drawn several conclusions.

First, a group of images in good condition and with proper overlaying taken from only one view of the site or a group of images sporadically taken from various parts of the construction sites without proper overlaying or with faults during capturing. Such faults include the glare of the sun or capturing a scene that is unrelated to the project management task can be used to reconstruct only a small area of a large construction site. Therefore, the sparse reconstructed model which is resulted from Case study 1 uses only 25% of the images (Recall=0.25). The model can show only a particular view despite the existence of images taken from another view. Therefore, Case 1 (Recall=0.25) exhibited lower recall than Case 2 (Recall=0.85).

Second, in such a large industrial project (site area is approximately 20 ha), some parts of the site in which photographs are captured should be divided into smaller parts so that images can be captured more accurately. For example, every building or piping area must be considered separately. Therefore, the site is divided into small areas in Case 2. In these areas, photographs are taken from one of the buildings included in the project considered in Case 1. The recall characteristic in the Case 2 is 0.85, a value that is in a good agreement with other reported results obtained by previous studies conducted by Golparvar-Fard et al. (2011) [3].

Table 4 summarizes the results. The results are benchmarked on a computer with 2.4 GHz Intel® Core i5 CPU, 2.00 GB of RAM, and a Windows 64-bit platform. According to the results shown in Table 4, changing the strategy of capturing construction site photographs in case 2 increases the recall considerably. The average number of detected features in case 2 also increases considerably. Although the collection of images in both cases are not similar and are from two different projects, the higher number of features means that the proper image based model we expect that the results show this expectation.

<table>
<thead>
<tr>
<th>Case study 1</th>
<th>Case study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total # of images</td>
<td>399</td>
</tr>
<tr>
<td># of used</td>
<td>99</td>
</tr>
<tr>
<td># of points recovered</td>
<td>375,273</td>
</tr>
<tr>
<td>Recall(#of used/Total)</td>
<td>0.25</td>
</tr>
<tr>
<td>Average number of detected features</td>
<td>5904</td>
</tr>
<tr>
<td>Computation time</td>
<td>375 min</td>
</tr>
</tbody>
</table>

Finally, the authors do not refute the capability of the computer vision algorithm to establish 3D point cloud models for visualizing as-built progress data from unordered construction site photographs. However, we need to capture numerous photographs, all with good overlaying, when using this method. This condition is impossible for huge construction sites. Even if such goal is possible, computation time will be extended by up to more than a day, in some cases. Therefore, the authors strongly suggest the division of huge construction sites to smaller parts. Then, 3D point cloud models can be
constructed and visualized separately for every working package.

6 Acknowledgment

The writers would like to thank Oil Turbo Compressor Construction Company (OTCC) and department of construction engineering and management in Science and Research Branch of Islamic Azad University for their contributions to this research project. A special thanks is extended to Dr. Hosseini, OTCC’s managing director, and Mr. Azimi, OTCC’s project manager, for providing the first author with the opportunity of capturing images from the construction site. Any findings and conclusions expressed in this paper are those of the authors and do not reflect the views of the companies and individuals mentioned. Moreover, we would like to thank the reviewers who suggested some points to improve the quality of this paper.

References


A Computational Framework for Estimating the Carbon Footprint of Construction

Z.S. Moussavi Nadoushani* and A. Akbarnezhad*

*School of Civil and Environmental Engineering, University of New South Wales, Australia

E-mail: z.moussavinadoushani@student.unsw.edu.au, a.akbarnezhad@unsw.edu.au

Abstract - Estimation of the carbon emissions in the construction phase of the building life cycle requires acquiring and analysing a great deal of project specific data including the amount of the works performed by different equipment and the unit emissions of various machinery at different work conditions. Such information are very project specific and vary from one stage of the construction phase to another. Building Information Models (BIM) have been proposed as efficient sources of information which can be utilized to computerize the construction carbon estimation. However, an easy-to-use methodology and framework to perform such analysis is still lacking. In this paper, a BIM-based framework is proposed to estimate the construction carbon emissions using the quantity take-offs from BIM and the information obtained from WBS and carbon inventories. The proposed method also includes a predefined simulated construction process for various types of buildings which can be used to estimate the carbon emissions when construction plans are not available. Therefore, the method proposed in this paper can be used to compare carbon emissions of the various potential designs at early stages of the building design. A case study is presented to illustrate the advantages of the proposed framework.

Keywords - CO2 emission, Construction stage, Automation, BIM, Quantity takeoffs

1 Introduction

According to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), the building and construction sector is one of the seven dominant sectors that greatly contribute toward global greenhouse gas (GHG) emissions [1]. The building and construction industry consumes up to 40% of total worldwide used energy, and contributes up to 30% to the total GHG emissions annually. The United Nations Environment Programme (UNEP) estimates show that as a result of the sharp surge in urbanization and the inefficiency of existing building stock, GHG emissions will more than double in the next 20 years unless preventive measures are taken [2].

The total energy and GHG emissions of buildings constitute all those incurred in various phases of the building life cycle including the construction, use and end of life phases. In general, the life cycle emissions of buildings may be divided into two groups: embodied and operational. Embodied carbon (EC) refers to all primary carbon emissions associated with the energy used, directly or indirectly, throughout the life-cycle of a building component; and operating carbon (OC) which is emitted in maintaining the inside environment through processes such as heating and cooling, lighting and operating appliances. With this definition, the embodied carbon may include the carbon emitted by production and transportation of building components and by construction operations. The carbon emitted during production and transportation of building components is referred to as cradle to site embodied carbon. The available literature on reducing the embodied carbon of buildings is mainly focused on reducing the cradle to site embodied carbon of buildings through selection and use of low carbon materials. Significantly less has been done to investigate various methods to reduce the carbon emissions incurred in the construction and end of life phases of building life cycle.

With achievements in reducing the operational energy of buildings and growing hopes for making net zero energy buildings dream a reality, the relative share of the embodied energy and its associated GHG emissions in the life cycle carbon of buildings tend to increase. Therefore, estimating the total embodied carbon of buildings including the carbon emissions incurred in the construction phase has attracted a great deal of attention and is the topic of worldwide research.

However, estimation of the carbon emissions in the construction phase of the building life cycle requires acquiring and analysing a great deal of project specific data including the amount of the works performed by different equipment and the unit emissions of various
machinery at different work conditions. Such information are very project specific and even vary from one stage of the construction phase to another, rendering the analysis difficult and time consuming using traditional estimation techniques. The lack of a computational tool and the complexity of the construction stage highlight the need for an automation framework to enable the early estimation of the GHG emissions of this stage. A number of computing methods based on building information modelling and 4D modelling of buildings have been proposed to improve the estimation precision and reduce the computational time of the construction carbon estimations. However, an easy-to-use methodology and framework that considers all influencing parameters including the specific construction requirements of different construction components, emissions of individual construction operations, distance between the construction site and different material/component processing/fabrication sites and, the availability of equipment is lacking. In addition, all previously proposed computerized estimation methods require a great deal of details related to the availability of resources and construction plans and thus are of little value during the design stage of buildings when such details may not be yet available. In this paper, a BIM-based framework is proposed to estimate the construction carbon emissions using the quantity take-offs from BIM and the information obtained from WBS and carbon inventories. The proposed method uses the simulated construction process for various types of buildings to estimate the carbon emissions when construction plans are not available. Therefore, the method proposed in this paper can be used to compare the carbon emissions of the various design and construction method alternatives at early stages of planning. A hypothetical example is presented to show the applicability of the proposed framework.

2 Available Methods for Estimating the Carbon Footprint of Construction

Estimation of carbon emission of construction has attracted a great deal of attention in the past two decades. The study conducted by Cole (1998) [3] was one of the first attempts to estimate the energy and greenhouse gas emissions associated with the construction of alternative structural systems. In this study, a detailed examination of the energy and greenhouse gas emissions associated with the on-site construction of a selection of alternative wood, steel and concrete structural building assemblies was performed to investigate the share of the construction process in the total initial embodied carbon and the effects of different structural alternatives on the latter. Results of this study showed considerable differences between the GHG emissions associated with construction of alternative wood, steel and concrete structural assemblies and highlighted the construction of the concrete assemblies as the highest GHG emitter. The largest proportion of construction energy use for most structural assemblies was reported to be associated with the transportation of workers to and from the construction site [3].

Guggemos and Horvath [4, 5] developed a Construction Environmental Decision-Support Tool (CEDST) to evaluate the environmental effects due to the construction of commercial buildings. CEDST follows a predefined detailed process diagrams [5] to quantify the energy use and carbon emissions of the construction stage based on the designer’s and builder’s choices of structural materials, temporary materials and employed equipment. Through a case study, they showed that a single decision such as using a concrete mixer truck with a 335 hp engine rather than a 565 hp engine (with the same capacity) can reduce the total construction energy demand by as much as 12%. This highlights the importance of availability of a computerized carbon estimation tool to evaluate the effect of various decisions on the carbon emissions in construction and thus embodied carbon of the building.

The increasing interest on evaluating and comparing the environmental impacts of different construction scenarios led to several studies aimed at developing methods for quantitative estimation of the energy usage and GHG emissions in the construction phase. Yan et al. [6] and Mao et al. [7] evaluated the GHG emissions and energy consumption of the construction phase by considering different sources of emission (different boundaries). Unlike the previous studies [3, 4, 5], in these studies the embodied energy of the permanent building component’s materials was considered as a part of the construction embodied energy. However, one of the drawbacks of such methods is the considerable amount of information required to perform the analysis. This information includes the amount of the fuel, electricity, water and various materials used by different contractors. Therefore, such methods are not applicable to predict the carbon emissions before the completion of an actual construction activity. The need for the availability of actual site data is also a limiting factor in a number of other proposed frameworks including that proposed by Wong et al. [9]. The framework proposed by Wong et al. [9] uses a wide range of data collected on the needed equipment and their respective fuel consumptions as well as planned activity durations and start and finish times to estimate the construction related emissions incurred in a specific period of time. The latter is then combined by a virtual prototyping tool to allow the project teams to visualise the predicted carbon
emissions at different times in the construction process. To eliminate the dependence on fuel consumption data, Hong et al. [8] proposed a model to assess the energy consumption and GHG emissions in the construction phase by reflecting the type and energy efficiency of the equipment needed, based on the amount of the materials used and characteristics of the building project and construction site.

While the models developed in the previous studies reviewed above can be applied to achieve relatively satisfactory estimates of carbon emissions during construction, such methods have a number of disadvantages. First, the data entry requirements for the user are quite extensive, rendering the developed methods unsuitable for application in the early design stages. Second, quantity take-off and analysis of the collected information using traditional estimating methods are time consuming. The investors are inherently reluctant to spend money and time to perform carbon estimation of various construction alternatives, especially when reduction in carbon footprint is not an organizational or project objective. Third, while the carbon inventories for materials and activities are becoming more accessible, there is a lack of interoperability between design and analysis software and datasets that enable full life cycle analysis (LCA) of the building [10]. A number of computational based methods have recently been proposed to overcome or alleviate these problems. Russell-Smith and Lepech [10] proposed a framework to integrate BIM and LCA. The proposed framework uses the quantity take-offs from building information models and the unit carbon emissions of various materials and activities obtained from various data inventories to estimate the carbon emissions associated with the construction. However, feasibility and accuracy of the proposed method has not been yet investigated. Moreover, a major drawback of the method proposed by Russell-Smith and Lepech [10] is its inability to capture the effects of off-site and in-site transportation. The transportation of equipment or material from the processing/fabrication sites to the construction site is responsible for a great deal of carbon emissions during the construction process. Transportation has been reported as the second dominant parameter in the embodied energy/carbon of buildings after materials manufacturing [6, 7, 8]. The LCI databases mainly report on the cradle to gate embodied carbon of materials which should be modified to include the transportation impact to obtain cradle to site embodied carbon. To achieve a more accurate estimate of carbon emissions during construction, there is a need to identify the transportation requirements of a particular project and the associated carbon emissions based on the mode of transportation and quantity of the materials to be transported both in-site and off-site.

In this paper, a comprehensive computational framework is proposed to eliminate some of the major drawbacks associated with the existing methods of construction carbon estimation. The proposed computational framework uses the quantity take-offs from BIM to estimate the amount of the work performed by each construction equipment and transportation vehicle by taking into account the project specific information available including the availability of equipment. The estimated work hours and user input information about likely locations of various material/component processing and prefabrication facilities are then used by an analyzer unit to estimate the construction carbon emission. The effects of both on-site and off-site transportation are considered by estimating the transportation requirements. The construction simulator module of the proposed framework and tool enables designers to obtain an estimate of the carbon emission of various construction and design alternatives in early design stages. The latter can be highly beneficial in decision making about selection of best project alternatives. In addition, when coupled with optimization tools, the proposed methodology can be used to optimize the construction process in order to minimize the carbon emissions of the project.

3 Proposed Framework

By considering the objectives outlined earlier, the main challenge in this study was to develop methods for estimating the quantity of the works performed and materials used using the limited information available at the early stages of design phase. We assume that the only resource available is a rough 3-D representation of building modelled and analyzed using a state-of-the-art building information modelling software. We also assume that the information on the construction methodology and sequence of operations to perform a specific job is not available at the time of the analysis. As shown in Figure 1, the proposed framework consists of four main modules: (1) data collection unit; (2) project database; (3) carbon database; (4) construction process simulator and carbon emission analyser. The following sections provide a brief explanation of each module. For simplicity, the framework and case study presented in this paper only consider the construction of the structural elements (foundation, columns, beams, slabs and walls). However, the concepts developed in this study are general and can be easily extended to include all construction activities.
3.1 Data Collection Unit

As mentioned earlier, this study assumes that the analysis is performed in the early stages of design when the actual construction data are not available at the time of the analysis. By taking this into consideration, the proposed framework uses the quantity take-offs from a customized Building Information Model (BIM) developed in any state-of-the-art software including Autodesk Revit or Tekla Structures. For instance, Autodesk Revit has a powerful schedule and quantity take-offs tool which can be used to collect geometric data and properties of the building elements including the material, element type, cast unit, etc. Revit also has built-in equations to calculate parameters such as volumes (volume = length × width × height) of various materials used. Another feature of most state-of-the-art BIM software is the possibility to add new attributes to the building elements. For instance, the “shared parameter” function in Revit can be used to add a specific property, which is not already available in the default property listing of the software, to the building elements. The framework proposed here makes use of this feature to perform basic calculations within the BIM environment rather than in an external processor. For instance, a built-in equation is defined for concrete structural elements to determine their perimeters which are then used to calculate the area of the formworks required. Moreover, a shared parameter named “PWP coding” is assigned to all structural elements to refer to their Predefined Work Packages. With this, any element family in Revit has a PWP code which determines the specific series of the activities required to construct the element. The PWP code is assigned to element families through predetermined formulas which use basic properties of the element to determine the construction method needed. The use of PWP as a built-in attribute can considerably reduce the computational time by eliminating the process which would have been otherwise performed by the external processor to identify the activities required to construct the element. Further details about the Predefined Work Packages are presented in the following sections.

The quantities of materials (including reinforcing steel, and concrete) required for erecting columns, beams, walls, and slabs are extracted and exported as separate .txt files which can then be imported by the data processor to create the database required for the analysis.

3.2 Project Database

Project database comprises information imported from BIM as explained in the previous section as well as any project specific data available. Although the framework presented in this study requires minimal project specific information, the project database was programmed to import any available actual project data including the equipment list, schedule or work breakdown structure (WBS) and modified construction procedures for particular element families (PWP codes). If available, this information will then replace the results of the simulation and the default values. To be imported to the project database, such information should be saved in the prescribed formats and stored in the prescribed file directory. The analyzer developed in the present study then searches for the specific file formats saved in the predefined directory, imports the required data from the files available and sorts them out into the data structure which could be later used by the processor to perform various calculations.

A user interface was developed to collect the user preferences where various alternatives to conduct a particular work are available. Many construction activities can be conducted using different combinations of equipment or methods. In such cases, a set of alternative options are generated based on the resource availability indicated in the project database and the user is asked to select the preferred options based on his/her engineering judgement. The options provided in our developed tool are related to the vehicle type, alternative suppliers of a material (or inputting the estimated travelling distance for critical materials) which can be selected from the embedded dropdown menus generated using the results of the construction process simulation.

3.3 Carbon Database

The carbon database consists of three sub-databases: (1) the “material” sub-database comprising embodied carbon factors for temporary materials used in the construction process; (2) the “transportation” sub-database comprising carbon emission factors for
different vehicles commonly used to transport materials, building components and equipment to and from construction site; (3) the “equipment” sub-database comprising carbon emission factors for different equipment used commonly in construction. The carbon database was developed using various international carbon inventories available [12, 13].

3.4 Construction process simulator and carbon emission analyser

The construction process simulator is a computer program developed in MATLAB which identifies the construction process (series of construction operations) required to construct a particular element in the building. The program first identifies the general WBS for the construction of the particular building type by analysing the type of the individual elements in the building. For instance, the predefined WBS for the construction of a cast-in-place concrete building comprises the following tasks: hoisting and setting of column steel, setting of column forms, pouring the columns concrete, hoisting and setting of wall steel, setting of wall forms, pouring the walls concrete, hoisting and setting of beam steel, setting of beam forms, setting of slab forms, hoisting and setting of slab steel, pouring the beams and slabs concrete. If available at the time of the analysis, the actual WBS and/or schedules overwrite the predefined WBS. In the next step, the program uses the PWP attribute imported from BIM to identify the type and number of the construction operations required to complete one unit of a particular job (e.g. construction of a cast-in-situ concrete column). The PWPs used in the present study were developed based on RSMeans [11]. Table 1 illustrates an example of a PWP for construction of cast-in-place concrete elements. It is worth mentioning that at this stage of the work only the tasks involving the use of special fuel consuming equipment or vehicle are considered. As a part of the future work, the labour requirements will be added to the PWPs to estimate emissions due to the transportation of labours to and from the construction site.

By considering the capacities and efficiencies of the required equipment available, the analyser uses the quantity take-offs from BIM to estimate the amount of the work performed by each construction equipment and transportation vehicle. The latter is then used to calculate the emitted carbon of the individual construction operations using the equations deduced from reference [8]. The emissions associated with individual operations are then summed up to calculate the total emissions associated with the predefined/given WBS.

4 Case Study

The proposed framework was adopted to predict the CO₂ emissions due to the construction of a three story hypothetical building with a total area of 960 m². The 3D view of the case structure comprising various structural elements (foundations, columns, walls, beams and slabs) is shown in Figure 2. The model was developed in Autodesk Revit. The travelling distance between the material processing sites and construction site was assumed to be 30 km for all the materials. With this simplifying assumption, the transportation emissions are estimated for round trips comprising of 30 km travel of loaded vehicles to the construction site and 30 km return travel of the unloaded vehicles to the processing plant or supplier’s facility. These assumptions can be modified by the user through inputting travelling distance estimates manually or selecting the preferred suppliers from the list of local suppliers available in the project database. It was assumed that plywood formworks can be reused for a maximum of 3 cycles of casting. The overall work breakdown selected by the analyzer includes the use of crane and bucket for pouring concrete, two vibrators for vibrating the poured concrete and a truck crane to displace and hoist the forms and reinforcements.

<table>
<thead>
<tr>
<th>Tasks included in Predefined Work Package</th>
<th>Considered vehicles and equipment</th>
</tr>
</thead>
</table>
| 1. hoisting and setting of steel        | 1.1. Transportation of steel bars to the construction site (truck…)  
|                                         | 1.2. Hoisting and setting the bars for each level (crane…) |
|                                         | 2.1. the amount of used form |
| 2. setting of forms                    | 2.2. Transportation of form to the construction site (truck…)  
|                                         | 2.2. Setting the form for each level (crane…) |
|                                         | 3.1. Transportation of concrete to the construction site (concrete mixer truck) |
| 3. pouring concrete                    | 3.2. Pouring the concrete for each level (crane…) |
| 4. vibration of concrete               | 4.1. vibration of the poured concrete (vibrator) |
Various stages of the analysis performed are shown in Figure 3. Figure 3a and 3b show the project database units developed in MS Excel. The quantity take-offs from the building information model including the volumes of concrete and reinforcements as well as the geometrical properties of different elements were imported from the text files generated by Revit. As can be seen in Figure 3b, each structural element has a PWP attribute which refers to its element family and thus determines the chain of the operations required for its construction. The quantity of total reinforcements, concrete and formworks required as estimated by Revit is shown in Table 2.

After importing the quantities and the PWP codes to the project database, the analyzer program developed in MATLAB was run to simulate the construction operations and calculate the associated carbon emissions using the methodology presented in this paper.

<table>
<thead>
<tr>
<th>Type of element</th>
<th>Foundation</th>
<th>Column</th>
<th>Beam</th>
<th>Wall</th>
<th>Slab</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcement (ton)</td>
<td>1.29</td>
<td>4.80</td>
<td>8.61</td>
<td>7.82</td>
<td>9.89</td>
<td>32.41</td>
</tr>
<tr>
<td>CO₂ emissions from reinforcement transportation (kg)</td>
<td>4.14</td>
<td>15.42</td>
<td>27.61</td>
<td>25.08</td>
<td>31.73</td>
<td>103.98</td>
</tr>
<tr>
<td>CO₂ emissions from reinforcement hoisting (kg)</td>
<td>1.12</td>
<td>4.18</td>
<td>7.48</td>
<td>6.80</td>
<td>8.60</td>
<td>28.18</td>
</tr>
<tr>
<td>Formwork (m²)</td>
<td>80.88</td>
<td>431.52</td>
<td>614.20</td>
<td>610.48</td>
<td>830.00</td>
<td>2567.08</td>
</tr>
<tr>
<td>CO₂ emission from formwork transportation (kg)</td>
<td>2.38</td>
<td>12.71</td>
<td>18.09</td>
<td>17.98</td>
<td>24.45</td>
<td>75.61</td>
</tr>
<tr>
<td>CO₂ emission from formwork hoisting (kg)</td>
<td>0.65</td>
<td>3.44</td>
<td>4.90</td>
<td>4.87</td>
<td>6.63</td>
<td>20.49</td>
</tr>
<tr>
<td>Concrete (m³)</td>
<td>22.20</td>
<td>16.74</td>
<td>51.60</td>
<td>91.57</td>
<td>99.47</td>
<td>281.58</td>
</tr>
<tr>
<td>CO₂ emission from concrete transportation (kg)</td>
<td>652.15</td>
<td>491.75</td>
<td>1515.80</td>
<td>2689.96</td>
<td>2922.03</td>
<td>8271.69</td>
</tr>
<tr>
<td>CO₂ emission from concrete pouring (kg)</td>
<td>60.76</td>
<td>83.30</td>
<td>313.84</td>
<td>278.48</td>
<td>286.58</td>
<td>1022.97</td>
</tr>
<tr>
<td>CO₂ emission from concrete vibration (kg)</td>
<td>12.03</td>
<td>16.50</td>
<td>62.15</td>
<td>55.15</td>
<td>56.75</td>
<td>202.58</td>
</tr>
</tbody>
</table>

Figure 2. 3D view of the structural elements and their embedded reinforcement

Figure 3. a, b) Project Database c) Carbon Database d) Construction process simulator
It should be noted that in this study, only the emissions associated directly with construction of concrete structural elements were estimated. The unit emissions of the required equipment identified and selected are summarized in Table 3. The estimated CO₂ emissions for construction of individual elements present in the building information model are summarized in Table 2.

Table 3. CO₂ emission factor for different Vehicle/Equipment used in construction of the case structure [8]

<table>
<thead>
<tr>
<th>Usage</th>
<th>Type of Vehicle / Equipment</th>
<th>CO₂ emission factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle for reinforcement</td>
<td>Trailer 20</td>
<td>0.0713 (kg/ton.km)</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle for formwork transportation</td>
<td>Trailer 20</td>
<td>0.0713 (kg/ton.km)</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle for concrete transportation</td>
<td>Concrete mixer truck 6</td>
<td>0.153 (kg/ton.km)</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment used for hoisting</td>
<td>Truck crane 25</td>
<td>15.956 (kg/hr)</td>
</tr>
<tr>
<td>reinforcement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment used for hoisting</td>
<td>Truck crane 25</td>
<td>15.956 (kg/hr)</td>
</tr>
<tr>
<td>formwork</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment used for pouring</td>
<td>Truck crane 50</td>
<td>26.157 (kg/hr)</td>
</tr>
<tr>
<td>concrete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment used for vibrating</td>
<td>Concrete vibrator 2.5</td>
<td>2.59 (kg/hr)</td>
</tr>
<tr>
<td>concrete</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4a shows the breakdown of the relative proportions of CO₂ emissions corresponding to the transportation of materials, operation of on-site equipment and the embodied carbon of the temporary materials used. As shown, transportation was identified as the largest contributor to the carbon emissions incurred in the construction phase, accounting for 67.7% of the total construction emissions estimated. This is while the assumed average distance of 30 km between the material processing/production plants and construction site is an optimistic assumption and may be exceeded in many actual construction projects.

The objective of the case study presented here is solely to illustrate the advantages of the method illustrated in this paper. In addition, due to the sensitivity of emission estimates to the project specific parameters, the results presented here cannot be used as indicators of average contribution of various parameters to the overall construction emissions. However, by taking into account the reasonable or even optimistically short (in some cases) travelling distances considered for materials transportation, the relatively high transportation emissions observed for the case project highlights the importance of considering the contribution of transportation in estimating the carbon emissions of construction. It should be noted that the material embodied carbon values reported by the majority of inventories available are cradle to gate values and therefore do not account for transportation emissions.

The results presented in Figure 4 show that temporary materials used for formworks are the second contributor to the construction carbon emissions, accounting for about 22.1% of the total construction emissions. The contribution of temporary materials to construction emissions may vary significantly depending on the type of the formwork material used and number of times the formworks are reused. Results showed that the contribution of construction operations to construction carbon emissions is smaller than that of transportation and temporary materials. As shown, construction operations were estimated to account for about only 10.2% of the construction emissions incurred. Figure 4b illustrates the contribution of different elements to the carbon emissions of the construction stage. As shown, for the case project presented, slabs seem to be responsible for a higher amount of CO₂ emission in the construction phase compared to the other considered.

The relatively high contributions of transportation and temporary materials to the construction emissions suggest that the local availability of materials and embodied carbon of temporary materials should be considered as two important parameters during the design and planning stage to reduce the construction emissions. As shown, the availability of a construction simulator and carbon estimation tool can be highly beneficial in early stages of building design by providing designers and construction managers with a tool to compare various design and construction plan alternatives.

Figure 4. The contribution of different; a) parameters b) elements; in the carbon emission of the construction stage
5 Conclusion

In this study, a computational framework was proposed to estimate the carbon emissions incurred in the construction phase at early stages of design and project planning when detailed construction plans and schedules are not available. The results of a case study were presented to illustrate the advantages of the method proposed. The proposed framework may be used to estimate the construction carbon emissions before the start of the actual construction process and therefore can be utilized by designers and construction managers to compare various design and construction plan alternatives.

References


A Mixed-integer Nonlinear Programming Model for Minimising Construction Site Noise Levels through Site Layout Optimisation

A.W.A. Hammad\(^a\), D. Rey\(^a\) and A. Akbarnezhad\(^a\)

\(^a\)School of Civil and Environmental Engineering, University of New South Wales, Australia
E-mail: a.hammad@unsw.edu.au, d.rey@unsw.edu.au, a.akbarnezhad@unsw.edu.au

Abstract – Activities undertaken on a construction site are often accompanied with high levels of noise. Addressing the issue of noise pollution in construction is gaining significance with the growing awareness about the social and environmental components of sustainable construction and the increasing numbers of projects being undertaken in congested urban areas. The documented methods for reducing noise pollution in construction include controlling (1) the noise produced at the source; (2) noise levels reaching a receptor; (3) noise propagated along the transmission path. Methods addressing the latter points use the fact that attenuation of noise increases as the transmission path gets longer. Thus the efficiency of such methods can be improved considerably through optimising the arrangement of temporary facilities on construction sites, with respect to a receptor, making use of noise attenuation due to distancing noisy facilities away from noise-sensitive receivers. The building under construction can also be used as a barrier to the noise transmission path, where obstruction of particular facilities from a given receiver can help in producing lower levels of sound as measured at the receptor. The available literature on site layout planning is extensive but limited to only achieving traditional construction project objectives (travel and material handling cost, safety, etc.). This paper presents a mixed integer non-linear programming (MINLP) model that optimises the location of temporary facilities on site in order to minimise the sound levels measured at a pre-defined receptor. The present model is expressed in three stages: (1) defining the noise objective function; (2) implementing model constraints; and (3) application of COUENNE to solve the MINLP for a case study.

Keywords – construction site layout; mixed integer nonlinear programming; MINLP; optimisation; construction noise estimation; sustainability

1 Introduction and Literature review

Minimising the pollution caused by construction activities, in all its categories, is an important aspect of sustainable construction. Pollution on a construction site occurs when contamination or harm at any level is caused to the surrounding communities [1]. In particular, construction activities are renowned for the high levels of noise that they produce. Hence, the reduction of sound produced on a construction site happens to be an important objective in sustainable construction.

Construction noise will most often cause a nuisance to the residents surrounding the construction site. The response of a human to a particular sound depends on how the ear perceives that sound, since the frequency response of the ear varies with sound intensity. At low intensities the ear’s response to a change in frequency of the sound is more varied compared to the almost flat response made for high intensity sounds. To take into account this variation in the reaction of the ear, the A-weighted frequency response curve is adopted, where sound level measurements are quoted in dB (A) (A-weighted decibels). The A-weighted frequency response curve ensures that sound measurements are made with one frequency response that reacts well to sound at any frequency or intensity [2].

The impacts arising from noise pollution can be categorised into two main streams, namely occupational, which is concerned with the health hazards faced by on-site workers, and environmental, where the people and wildlife neighbouring the noise source are targeted [2]. Thus, noise, propagated from a construction site surrounding different communities, may lead to a disturbance in several ways. For example, manufacturing industries relying on precise measuring equipment could potentially be disrupted when exposed to high levels of sound [3]. Productivity of employees in buildings adjacent to noisy highways has been reported to be adversely affected [4]. Belojević et al. [5] conducted a study to determine the impacts noise levels had on the cognitive functioning of 45 subjects. Results indicated that subjects exposed to high sound levels had
their mental capabilities lowered. A number of studies have reported findings on disruptions to both patients and healthcare workers caused by high sound levels initiated within a hospital. High levels of noise around hospitals, where quiet operation theatres and accurate medical equipment are vital, may have a detrimental impact on the health of hospital administered patients [6]. Morrison et al. [7] reported in their study that data analysed statistically indicated a high correlation between above-average noise levels within a hospital setting and an increase in the level of stress experienced by nurses. The authors concluded that this could lead to the possibility of medical errors being committed by healthcare workers.

Though the regulatory authorities in most developed countries have set out criteria to control the levels of noise produced by construction activities, the laws and regulations yet remain vague when addressing the issue [8]. In Australia the rules set out are usually controlled by the relevant councils within which the construction work is being undertaken. The main focus of the regulations set out by the councils is on 1) restricting working hours of machinery on construction sites between a specified time frame; and 2) setting out criteria that noise levels not exceed background levels by 5 dB(A) and 10dB(A), during the first working hour and during other working hours, respectively [9], [10]. Standards aimed at addressing noise level reduction and measurements primarily deal with the effect noise has on workers and machine operators on site [11], [12]. Nevertheless these standards can be adopted to address noise pollution subjected at the environment surrounding the construction site.

Several studies in the literature have investigated the methods applicable for controlling sound levels produced by construction equipment. Suter [13] presented a number of techniques aimed at attenuating sound levels affecting the on-site workers. The author discussed the use of electronic ear muffs that adjust noise levels across a given frequency range. The author also discoursed that a cheap and conventional way of attaining noise reductions for construction equipment was through proper maintenance of the equipment. Schneider et al. [14] included in their study the application of retrofits in construction equipment as a procedure for direct noise level reductions at the noise source. Schneider et al. [15] also advocated the use of retrofits on heavy equipment for achieving substantial noise level reductions. Along with the methods presented above, Standards Australia [11] details the use of barriers on construction sites to control noise levels. Attenuation is achieved either through the use of acoustic screens layered with noise absorbent materials, such as chipboard or compressed straw, which enclose the machinery under use, or through creating physical barriers, by piling up construction materials available throughout the different construction phases, to obscure a particular receiving point from the emitting noise source.

However, while considerable literature exists on estimating construction sound levels, reducing noise levels of equipment and identifying the health hazards caused by construction noise pollution, little effort has been made to investigate how construction planning affects noise pollution. An important factor affecting sound levels on a construction site is the adopted layout of the site. The level of noise experienced by a receiver located within a construction site or on the outskirts of the site will vary depending on the noise control mechanism executed on site. As discussed above, these mechanisms can be applied at the noise source, between the noise source and the receiver or at the receiving point itself.

The available literature on site layout focus mainly on optimising the locations of temporary facilities through addressing objectives such as travel cost, material handling cost, safety breaches and wildlife preservation [16]–[21]. Literature addressing noise levels on construction sites mostly emphasise the quantification of the sound levels produced. Gannoruwa and Ruwanpura [22] used simulation and a stochastic model to predict noise levels at receivers, subject to the construction stage specified by the user. Their work focussed on altering noise levels along the transmission paths by optimising sizes of barriers to be placed in obstruction to the noise path. Li [23] presented a computer system capable of calculating noise levels produced by different categories of construction equipment, with a strong emphasis on taking into account attenuation due to weather and ground conditions of the construction site. These two factors were argued to cause disturbance to the accuracy of noise levels predicted at the receiver whenever the distance of the source from the receiver exceeded 100 m. Zaiton Haron and Khairulzan Yahya [24] used Monte Carlo methods in order to predict the noise levels from equipment on a construction site. Simulation was run to collect data and the probability density functions (PDF) and cumulative distribution functions (CDF) of noise levels based on the samples obtained were plotted. The overall mean of noise levels was acquired by combining the PDFs from the multiple noise sources. The results were compared to those derived using the methodology outlined in BS 5228 [12]. Gilchrist et al. [3] implemented Monte Carlo simulation to predict noise mitigation in urban environments. Their method relied on specifying a value to represent the probability of each operating state of the equipment in use (i.e. Idle, working, and not operating). Maximum noise levels, produced by their deterministic model, were obtained
from the statistical data of the simulation run. The results were validated against on-site measurements taken around a hospital surrounding the construction site.

A diverse range of equipment, including heavy equipment such as pile drivers and excavators, as well as the smaller and more frequently used equipment like saw chains, hammers, electric drills, concrete mixers and vibrators and steel benders and cutters, is employed on construction sites, each with a different size, different operating durations and different levels of generated sound. The latter group of equipment are usually stationed in temporary facilities such as steel bending yards, formwork assembly yards and concrete batching areas. These facilities can be positioned on site in a certain way as to reduce the overall noise levels reaching low-tolerance limit receivers. This is especially important if the construction site happens to be in the vicinity of noise-sensitive receptors such as hospitals, schools, etc. To the best of our knowledge, there has been no effort on optimising the construction site layout to minimise noise pollution caused by construction activities.

This paper presents a novel mathematical optimisation model for site layout planning, where the objective function is minimised to reduce noise levels on construction sites. The proposed optimisation model will be applied to and solved for an illustrative case project. The problem is formulated as a mixed integer nonlinear programme (MINLP) on General Algebraic Modelling System (GAMS) and will be solved using COUENNE, a solver originally developed at Carnegie Mellon University and IBM Research. While this paper’s main focus is on applying the proposed optimisation model to construction site layout planning, the model can be extended to areas such as industrial and urban planning, where sound level reduction is critical. To achieve a more realistic optimisation model, the noise objective function should also be considered along with a cost objective function, rendering the problem a multi-objective one. The latter is currently the subject of future study by the authors.

2 Construction site layout model

BS 5228 and AS 2643 present a method for estimating noise levels produced on site [11], [12]. The general format of the objective function that will be used in this section has been obtained from these standards; nevertheless some modifications were made in order to reflect the changes that happen over time during construction projects.

2.1 Notation

This section lists the index notations, scalars, parameters and variables defined in the model.

2.1.1 Sets and indices

\( F \) : set of facilities to be located, indexed by \( f \) and \( j \).

\( L \) : set of predetermined locations on the construction site, indexed by \( l \).

\( ns \) : Noise source within a temporary facility on construction site

\( b \) : Insulation material type

\( e = (ns, f) \) : Tuple to map each noise source, \( ns \), to a facility \( f \)

\( s \) : Slab on grade (SOG) (includes ground preparation and excavation stages)

\( c \) : 1st floor columns

\( completion \) : Project completion

2.1.2 Scalars

\( W \) : Width of construction site in x-direction.

\( B \) : Length of construction site in y-direction.

\( R refl \) : Reflection effect, on noise levels measured at the receiver, due to buildings surrounding the receiving point, equal to 3 dB (A).

\( RLX \) : X-coordinate of receiver point, located around construction site.

\( RLY \) : Y-coordinate of receiver point, located around construction site.

\( H \) : Very small positive value, equal to 5e-5.

\( AM \) : Maximum attenuation due to barrier effect, equal to 10 dB (A), as specified in BS 5228.

\( T \) : Noise Assessment period, measured as working hours on construction site per day.

\( TPD \) : Total project duration, in months.

2.1.3 Parameters

\( LAeq \) : Continuous equivalent sound pressure level, measured at 10 m from the source, obtained from BS 5228.

\( t_e \) : Time duration, over assessment period \( T \), where noise source \( ns \), defined over tuple \( e \), is on.

\( \gamma' \) : X-coordinate of border of building under construction, closest to receiver point.

\( \gamma'' \) : X-coordinate of border of building under construction, furthest away from receiver point.
2.1.4 Continuous and discrete Variables

\( c_{x_f} \): X-coordinate of centroid of facility \( f \).

\( c_{y_f} \): Y-coordinate of centroid of facility \( f \).

\( R_f \): Euclidean distance between facility \( f \) and receiver point.

\( K_h \): Distance adjustment factor for each noise source assumed to be located at the centroid of facility \( f \).

\( Att_{sc_f} \): Average attenuation, as measured over the whole duration of the project, due to screening effect of building under construction. This variable can take on one of three values, depending on the projected construction stage and on the location of the facility from which noise is emitted.

\( L_{eq} \): Equivalent continuous sound pressure level over tuple \( e=(ns,f) \), measured in dB (A).

\( DBCSLX_{f,j} \): The absolute value of the distance between centroids of facilities \( f \) and \( j \) in the x-direction.

\( DBCSLY_{f,j} \): The absolute value of the distance between centroids of facilities \( f \) and \( j \) in the y-direction.

2.1.5 Binary Variables

\( z_{f,\ell} \): Equals one if facility \( f \) is located at location \( \ell \).

\( \mu_{x_{f,j}} \): Equals one if facilities \( f \) and \( j \) do not overlap in the x-direction.

\( \delta_{y_{f,j}} \): Equals one if facility \( f \) is located at a coordinate greater than \( y' \) in the x-direction, and zero otherwise.

\( \phi_{x_{f,j}} \): Equals one if attenuation \( Att_{sc_f} \) over construction stage \( n \) for facility \( f \) applies, and zero otherwise.

2.2 Noise objective function formulation and constraints

2.2.1 Objective function

\[
\text{Minimise } 10 \log_{10} \left( \frac{1}{T} \sum_{e} t_e 10^{\frac{L_e}{10}} \right)
\]

Where,

\( e = (ns,f) \)  \hspace{1cm} (2)

\[
L_e = L_{eq} - K_h + Refl - \sum_{n} \rho_{n,f} Att_{sc_f} - \sum_{b} v_b \theta_b
\]

\( K_h = 20 \log_{10} (R_f) - 8 \)  \hspace{1cm} (4)

\[
R_f = \sqrt{(RLX - c_{x_f})^2 + (RLY - c_{y_f})^2}
\]

Objective function (1) is the noise level equation for various activities taking place within each of the temporary facilities during the assessment period \( T \), and it incorporates equations (3), (4) and (5). It gives the combined equivalent A-weighted sound pressure level for all noise sources included in the summation operator. The summation operator functions over the tuple described by equation (2), as this maps each given noise source to a particular temporary facility. Equations (3) calculate the individual continuous A-weighted sound pressure for each noise source and associated facility in tuple \( e \), taking into account all factors and attenuations altering the sound levels. Equations (4) calculate the distance allowance factors for each facility, assuming the noise sources at a particular facility are located at the centroid of that facility. Equations (5) calculate the Euclidean distance between the receiver point and the
centroids of the temporary facilities.

### 2.2.2 Facility-Location Constraints

\[ \sum_{f} z_{f,j} = 1 \quad \forall f = 1, \ldots, F \]  
(6)

\[ \sum_{f} z_{f,t} \geq 1 \quad \forall t = 1, \ldots, L \]  
(7)

\[ z_{f,t} = 0 \quad \text{for} \ t = P \]  
(8)

Where \( P \) represents the location of the permanent building under construction.

Constraints (6) ensure that one location is assigned to each facility. Constraints (7) ensure that each predefined location contains at least one facility. This is to try to spread out the allocation of facilities within the predefined locations, hence making use of all the available locations. Constraint (8) is for excluding the building construction area from being used for temporary facility assignments.

### 2.2.3 Facility non-overlap constraints

\[ DBCSLY_{f,t} = |x_{f} - cx| \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(9)

\[ DBCSLY_{f,t} = |y_{f} - cy| \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(10)

\[ DBCSLY_{f,t} \geq 0.5WF + WF/\mu x_{f,t} \lambda \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(11)

\[ DBCSLY_{f,t} \geq 0.5(1-F + LF/\mu y_{f,t} \lambda) \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(12)

\[ (0-z_{f,j}) + (0-z_{f,j}) + \mu x_{f,t} + \mu y_{f,t} \leq 1 \quad \forall t = 1, \ldots, L \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(13)

\[ \mu x_{f,t} + \mu y_{f,t} \leq 1, \quad \forall f = 1, \ldots, F - 1 \quad \forall j = 1, \ldots, F \quad f \neq j \]  
(14)

Constraints (9)-(14) ensure that two facilities located at the same location do not overlap. In particular, constraints (9) and (10) are to prevent negative distances in the x and y directions, respectively. Constraints (11), (12), (13) and (14) prevent the facilities from overlapping in the x-direction and y-direction at the same time.

### 2.2.4 Location boundary constraints

\[ cx_{f} + (0.5WF) \leq (CW_{f} + (0.5WF_{t})) z_{f,j} \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, L \]  
(15)

\[ cy_{f} + (0.5WF) \leq (CW_{f} + (0.5WF_{t})) z_{f,j} + (1 - z_{f,j}) \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, L \]  
(16)

\[ cy_{f} + (0.5WF) \leq (CW_{f} + (0.5WF_{t})) z_{f,j} + (1 - z_{f,j}) \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, L \]  
(17)

\[ cy_{f} - (0.5WF) \leq (CW_{f} + (0.5WF_{t})) z_{f,j} \quad \forall f = 1, \ldots, F \quad \forall j = 1, \ldots, L \]  
(18)

Constraints (15), (16), (17) and (18) ensure that the facilities are located within the boundaries of the locations to which they have been assigned, as determined by the binary variable \( z_{f,j} \).

### 2.2.5 Constraints for defining \( Attsc_{f} \)

\[ \delta_{f} + \delta_{c} \geq 1 \quad \forall f = 1, \ldots, F \]  
(19)

\[ Attsc_{f} \geq 0 \quad \forall f = 1, \ldots, F \]  
(20)

\[ Attsc_{f} \leq M^{4}(1 - \phi_{a,c}) \quad \forall f = 1, \ldots, F \]  
(21)

\[ cx_{f} \geq (a + \varepsilon)(1 - \delta_{f}) \quad \forall f = 1, \ldots, F \]  
(22)

\[ \delta_{f} + \phi_{a,c} + (1 - \delta_{f}) \geq 1 \quad \forall f = 1, \ldots, F \]  
(23)

Constraints (19), (20), (21) and (22), are to ensure that if the distance between the receiver and \( cx_{f} \) is less than the distance between the receiver and \( \gamma \) then \( Attsc_{f} \) will equal zero. Facilities positioned at such locations are close to the receiver point, and so no attenuation of noise due to the building under construction occurs.
Constraints (23), (24), (25), (26) and (27) are to ensure that if \( cx \leq \gamma^e \) and \( cx > \gamma^d \), then the overall \( \text{Attsc}_f \) will equal \( \sum_{i=1}^{\text{construction}} \frac{p_i}{TPD} \), broken down as follows: 1) zero for the stages in construction preceding the casting of the SOG; and 2) \( \frac{p_{\text{cons}} + \ldots + p_{\text{complete}}}{TPD} \), after the casting of the SOG occurs, until the completion point of the project. Facilities located within the aforementioned boundaries tend not to be fully obscured by the building. Hence, the 5 in the latter equation represents the value of partial blockage of noise from a given source, as noted in BS 5228 [12].

Constraints (28), (29), (30) and (31) ensure that for facilities where the distance between the receiver and \( \gamma^d \) is less than the distance between the receiver and \( cx \), then their corresponding \( \text{Attsc}_f \) will equal:

\[
5 \left( \sum_{i=1}^{\text{construction}} \frac{p_i}{TPD} \right) + 10 \left( \sum_{i=1}^{\text{construction}} \frac{p_\text{cons} + \ldots + p_{\text{complete}}}{TPD} \right),
\]

broken down as follows: 1) zero for the period up until the SOG is casted; 2) \( \frac{p_{\text{cons}} + \ldots + p_{\text{complete}}}{TPD} \) after the casting of the SOG occurs and up until the first floor columns are casted; 3) \( \frac{p_{\text{cons}} + \ldots + p_{\text{complete}}}{TPD} \) after the first floor columns are casted up until completion of the building. In such case partial blockage of noise due to the building under construction happens only in the stages preceding the casting of the first floor columns. Once the first floor columns are casted, full blockage is assumed; hence the value of 10 is used for the later stages in the project. Facilities governed by constraints (28), (29), (30) and (31) are the ones located furthest apart from the receiver point.

It is assumed in this model that \( \text{Attsc}_f \) takes on one of three values. Therefore constraints (32) ensure that \( \text{Attsc}_f \) is equal to a single value for each facility \( f \).

3 Application and numerical results

The model developed in this study was tested using a case project. The case study is a hypothetical project involving the construction of a multi-storey building. Values for the \( L_{Aeq} \) sound levels, measured in dB(A) at 10 m from the source, obtained from BS 5228 and AS 2436, for different construction equipment, is shown in table 1. It should be noted that these figures are used as indicative values only since the actual level of sound generated may vary depending on a number of different equipment-related factors such as the manufacturer, model, age, condition of equipment, and the way the equipment is being used [2].

Table 2 shows the x and y coordinates of the centroids of the facilities, along with the noise sources at each of the given facilities. As shown in table 3, ten facilities were considered in the case study, with three being rectangles and the rest squares. A mixture of squares and rectangles is used to showcase the capability of the model in handling various dimensions of common facility shapes. Facilities considered include a site office, concrete batch plant, false-work yard, formwork assembly yard, steel welding and cutting yards, toilets, labour residence and a material warehouse. Table 4 presents the dimensions of the predefined locations on the construction site. Table 5 shows in which facilities acoustic screens are applied as a noise reduction measure.

Figure 1 presents the optimal layout of temporary facilities on site. Noise levels are measured at the receiver, assumed to be a hospital, located on the outskirts of the building at coordinates (36, 32). It is apparent from this figure that facilities are placed as far away from the receiver point as permitted by the boundaries of each of the predefined locations. The noisiest facilities are located at the opposite end of the site away from the receiver.

The model was executed on GAMS and solved using COUENNE. An overall optimal noise level of 75.9481 dB (A) was obtained from the solver. Computations were performed on a desktop computer running on Microsoft Windows 7 operating system, with Intel core i7 processor at 3.4 GHz and 16 GB of RAM. The model took 923 seconds for it to be solved to optimality, which is reflected in the fact that the
objective function is non-convex, hence requiring multiple branching.

Table 1. Noise sources and their levels

<table>
<thead>
<tr>
<th>Noise source symbol</th>
<th>Noise source</th>
<th>L_{Aeq} @ 10 m (BS 5228)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Club Hammer</td>
<td>79</td>
</tr>
<tr>
<td>S2</td>
<td>Hand held electric drill</td>
<td>81</td>
</tr>
<tr>
<td>S3</td>
<td>Concrete mixer</td>
<td>75</td>
</tr>
<tr>
<td>S4</td>
<td>Welder</td>
<td>73</td>
</tr>
<tr>
<td>S5</td>
<td>Generator</td>
<td>73</td>
</tr>
<tr>
<td>S6</td>
<td>Hand grinder</td>
<td>150</td>
</tr>
<tr>
<td>S7</td>
<td>Gas cutter</td>
<td>89</td>
</tr>
<tr>
<td>S8</td>
<td>Concrete Vibrator</td>
<td>78</td>
</tr>
<tr>
<td>S9</td>
<td>Drill</td>
<td>85</td>
</tr>
<tr>
<td>S10</td>
<td>Nett gen</td>
<td>73</td>
</tr>
<tr>
<td>S11</td>
<td>Angle grinder</td>
<td>80</td>
</tr>
<tr>
<td>S12</td>
<td>Toilet</td>
<td>75</td>
</tr>
<tr>
<td>S13</td>
<td>Air-con</td>
<td>50</td>
</tr>
<tr>
<td>S14</td>
<td>Normal Conversation</td>
<td>60</td>
</tr>
<tr>
<td>S15</td>
<td>Printer</td>
<td>80</td>
</tr>
<tr>
<td>S16</td>
<td>Quiet room</td>
<td>80</td>
</tr>
<tr>
<td>S17</td>
<td>Material Hoist</td>
<td>68</td>
</tr>
</tbody>
</table>

Table 2. Noise sources and centroids of facilities

<table>
<thead>
<tr>
<th>Noise source at each facility</th>
<th>Facility symbol</th>
<th>Location</th>
<th>x-centroid coordinate (m)</th>
<th>y-centroid coordinate (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1, S2</td>
<td>F1, F2</td>
<td>L6</td>
<td>3.5000</td>
<td>20.5000</td>
</tr>
<tr>
<td>S3, S4</td>
<td>F3, F4</td>
<td>L3,L2</td>
<td>29.5000</td>
<td>47.5000</td>
</tr>
<tr>
<td>S5, S6, S7</td>
<td>F5, F6, F8</td>
<td>L5, L6</td>
<td>2.5000</td>
<td>12.5000</td>
</tr>
<tr>
<td>S10, S11</td>
<td>F6, F7</td>
<td>L1, L4</td>
<td>10.5000</td>
<td>2.5000</td>
</tr>
<tr>
<td>S12</td>
<td>F7</td>
<td>L4</td>
<td>10.5000</td>
<td>2.5000</td>
</tr>
<tr>
<td>S13</td>
<td>F8</td>
<td>L1</td>
<td>2.5000</td>
<td>47.5000</td>
</tr>
<tr>
<td>S14, S15, S17</td>
<td>F8, F9, F10</td>
<td>L5</td>
<td>20.5000</td>
<td>2.5000</td>
</tr>
</tbody>
</table>

Table 3. Dimensions of facilities

<table>
<thead>
<tr>
<th>Facility</th>
<th>Facility Name</th>
<th>Width of facility in m (W)</th>
<th>Length of facility in m (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>Falsework</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>F2</td>
<td>Forklift</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>F3</td>
<td>Concrete Batch Plant</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>Steel working yard</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>F5</td>
<td>Steel cutting yard</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>F6</td>
<td>Forklift</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>F7</td>
<td>Toilets</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>F8</td>
<td>Labour residence</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>F9</td>
<td>Offices</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>F10</td>
<td>Warehouse</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4. Dimensions of predefined locations

<table>
<thead>
<tr>
<th>Location</th>
<th>Width of location in m (W)</th>
<th>Length of location in m (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>L2</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>L3</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>L4</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>L5</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>L6</td>
<td>8</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 5. Application of acoustic screens

<table>
<thead>
<tr>
<th>Acoustic Screen</th>
<th>Noise reduction achieved (AS 2436)</th>
<th>Facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compressed Straw</td>
<td>28</td>
<td>F2, F6</td>
</tr>
<tr>
<td>Plywood</td>
<td>16</td>
<td>F1</td>
</tr>
<tr>
<td>Chipboard</td>
<td>26</td>
<td>F3</td>
</tr>
</tbody>
</table>

4 Conclusion

In this study a mixed integer nonlinear mathematical optimisation model was presented, which is aimed at minimising noise levels measured at a specified receiver by optimising the layout of temporary facilities on a construction site. The model was tested out on a case study, where it was solved to optimality. Future work will involve improving the computational efficiency of the model, considering multiple receivers with different noise-sensitivity thresholds around the site, upgrading the model to incorporate more noise reduction mechanisms as decision variables and the inclusion of a cost minimisation objective to establish a multi-objective optimisation construction site layout problem.

5 References


MINING, BUILT INFRASTRUCTURE AND HUMAN FACTORS
Abstract -

A new LMS algorithm with a variable step-size and its application to improve the quality of broadcast-telephone system use in underground coal mines is presented in this paper. The step-size of the proposed algorithm is proportional to not only the norm of the smoothed gradient vector but also the smoothed mean-square error. Simulation results are presented to support the analysis and to compare the performance of the new LMS algorithm with the ordinary LMS, NLMS algorithms. They show that the performance of the new algorithm in stationary as well as in non-stationary environment is superior to ordinary LMS, NLMS algorithms.

Keywords -
Adaptive noise cancellation; Broadcast-telephone system; LMS algorithm; NLMS algorithm

1 Introduction

The broadcast-telephone system is used for managing, dispatching and ensuring safety in underground coal mines. Mine workers can communicate by two-way audio intercom with their manager through this system. The system is based on Ethernet platform in coal mine and UDP protocol, and converts voice signal to standard IP packet to transmit in Ethernet by using technologies of computer, network, VOIP and embedded system.

However, during the dialogue, the noise generated by the devices in underground coal mine will affect quality of human voice. According to the spectral analysis, the noise includes harmonic components and its spectrum overlaps a spectrum of the human voice and the band of the noise varies with time, so traditional filters as FIR or IIR are not effective. The results of our experiments show that, we can reduce this noise by using an adaptive filter.

One of the most popular algorithms in adaptive signal processing is the least mean square (LMS) algorithm with a constant step-size of Widrow and Hoff [3]. In practical application, a key parameter of LMS algorithm is the step-size. It is well known that, if the step-size is large, the convergence rate of the LMS algorithm will be fast, but the steady-state mean square error (MSE) will be large. On the other hand, if the step-size is small, the steady-state MSE will be small, but the convergence rate will be slow. Thus, the step-size provides a trade-off between the convergence rate and the steady-state MSE of the LMS algorithm. In recent years, a several variable step-size algorithms based on ordinary LMS algorithm have been proposed [1], [4], [5], [6]. Costa et al. in [1] proposed a noise resilient variable step-size LMS (NRVSS). The NRVSS algorithm is specially indicated for adaptive interference reduction in biomedical application. Ravikanth and Sanket Dessai [2] proposed a design and development of Noise Cancellation System (NCS) to cancel the background noise from the speech signal of speaker in android mobile phones. NCS system has been implemented using an adaptive filter. In [4], the authors proposed an algorithm to remove noise from speech signal in real time environment. The quality of audio signal can be improved by filtering the degraded speech signal through adaptive filters. Fengchun Wang, Jiyan Du [6], proposed the EFLMS (Error Feedback Least Mean Square) algorithm is based on Sigmoid function.

In [7], Akash Kashyap, et. al, designed an adaptive filter to remove an unwanted noise which might occur during music recordings, and echo in telephone networks. In [8], Zhenghua Ma; Lijun Yao; Jiongru Zhou; and Dan Kong introduced a design scheme of digital broadcast telephone system of coal mine based on TCP/IP protocol and TMS320DM642 digital processor. The system makes use of audio codec chip TLV320AIC23B to achieve audio codec and adaptive anti-noise algorithm so that the output audio's quality could be good to meet requirements of the voice in the mine, etc.

In this paper, the new algorithm based LMS algorithm where the step-size adjustment is controlled by the mean-square error and the norm of the gradient.
vector is proposed to enhance the convergence rate, steady-state MSE and other performance indexes of the ordinary LMS algorithm in the presence of statistically stationary or non-stationary measurement noise. As will be shown in the simulations, the proposed algorithm has improved performance as compared with existing LMS, NLMS algorithms.

Throughout the paper, the following notations are adopted, Table 1.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>E[ . ]</td>
<td>Expectation operator</td>
</tr>
<tr>
<td>Tr(.)</td>
<td>Trace of a matrix</td>
</tr>
<tr>
<td>w</td>
<td>Weight vector (Lx1)</td>
</tr>
<tr>
<td>R</td>
<td>Auto-correlation matrix (LxL)</td>
</tr>
<tr>
<td>e</td>
<td>Scalar-valued error signal</td>
</tr>
<tr>
<td>x</td>
<td>Row input vector (1xL)</td>
</tr>
<tr>
<td>(.)^T</td>
<td>Transpose of a vector</td>
</tr>
<tr>
<td>(.)^H</td>
<td>Transpose of a matrix</td>
</tr>
<tr>
<td>L</td>
<td>Number taps coefficients</td>
</tr>
<tr>
<td>N</td>
<td>Total number of samples</td>
</tr>
<tr>
<td>J</td>
<td>Function cost</td>
</tr>
<tr>
<td>µ</td>
<td>Step-size parameter</td>
</tr>
<tr>
<td>V</td>
<td>Gradient vector</td>
</tr>
<tr>
<td>P_s</td>
<td>Average power of the input x</td>
</tr>
<tr>
<td>δ^2</td>
<td>Variance</td>
</tr>
<tr>
<td>λ</td>
<td>Eigenvalue</td>
</tr>
</tbody>
</table>

Table 1. NOTATIONS

### 2 The Proposed New LMS Algorithm

In this section, the novel algorithm is proposed to remove unwanted noise in various environments such as underground coal mines. Figure 1 shows our adaptive noise cancellation for broadcast-telephone system. The adaptive filter consists of two distinct parts: a digital filter with adjustable coefficients w, and an adaptive algorithm which is used to adjust or modify the coefficients of the filter. Two input signals d_k (which is received by the microphone 1st - Mic1) and x_k (Mic2) are applied simultaneously to the adaptive filter. The signal d_k is the contaminated signal containing both the desired signal s_k - voice of mine worker - and the noise n_k - sound of devices in underground coal mine. These two signals (s_k and x_k) are assumed to be uncorrelated. The signal x_k is a measure of the contaminating signal which is correlated in some way with n_k. The noise x_k processed by the digital filter to produce an estimate n_k of n_k. An estimate of the desired signal is then obtained by subtracting the digital filter output n_k from d_k (1).

The main objective in noise cancelling is to produce an optimum estimate of the noise in the contaminated signals and hence an optimum estimate of the desired signal is received. This is achieved by using e_k in a feedback arrangement to adjust the digital filter coefficients w (weight vector) via the new adaptive algorithm, to minimize the noise in \( \hat{s}[k] \). From (1), we have:

\[
e_k = d_k - n_k
\]

From equation (4), we have the gradient of the performance surface:

\[
\frac{\partial J}{\partial w} = \frac{\delta^2}{\partial w} - \frac{\delta}{\partial w} \left( 2p^T_k w_k \right) + \frac{\partial}{\partial w} \left( w_k^2 \right)
\]

We use instantaneous estimates of cross-correlation

\[
\frac{\partial J}{\partial w} = \frac{\delta^2}{\partial w} - \frac{\delta}{\partial w} \left( 2p^T_k w_k \right)
\]

The weight vector is updated from sample to sample as follows

\[
w_{k+1} = w_k - \mu \frac{\partial J}{\partial w}
\]

Figure 1. Adaptive noise cancellation for broadcast-telephone system

\[
e_k = s_k - d_k - n_k = s_k + n_k - n_k
\]
vector \( p \) and auto-correlation matrix \( R \) by removing the expectation operator in (5). Thus
\[
\nabla_i = -2p_k + 2R_k w_k
\]
\[
\nabla_k = -2x_k d_k + x_k x_k w_k
\]
\[
\nabla_k = -2x_k (d_k - v_k w_k)
\]
\[
\nabla_k = -2x_k e_k
\]
From equations (5), (6) and (7), we have weight vector
\[
w_{k+1} = w_k + 2\mu x_k e_k \tag{8a}
\]
\[
w_{k+1} = w_k + 2/\mu_k - \mu R_k w_k \tag{8b}
\]
We consider, \( w \) is the optimum value of the weight vector. We have weight-error vector at \( k \) and \( (k+1) \) iterations:
\[
c = w_k - w_* \tag{9a}
\]
\[
c_{k+1} = w_{k+1} - w_* \tag{9b}
\]
From equations (8b), (9a), and (9b) we have
\[
c_{k+1} = e_k - 2\mu R_k e_k - 2/\mu_k w + 2\mu_k \tag{10}
\]
Because \( w \) is the optimum value of the weight vector, from (5) and (10) we get
\[
c_{k+1} = e_k - 2\mu R_k e_k \tag{11}
\]
\[
c_{k+1} = (1 - 2\mu R_k) e_k \tag{12}
\]
The auto-correlation matrix \( R \) may be diagonalized by using a unitary similarity decomposition \( R = QQ^H \), where \( A \) is the diagonal matrix of eigenvalues of \( R \), and the columns of \( Q \) contain the corresponding orthonormal eigenvectors, therefore
\[
e_{k+1} = (1 - 2\mu Q_k A_k Q_k^H)^{-1} e_k \tag{13}
\]
Using the property \( Q^H = Q^{-1} \), we get
\[
Q_k^H e_{k+1} = (1 - 2\mu_k A_k) Q_k e_k \tag{14}
\]
By defining \( v = Q_k e_k \), we may write
\[
v_{k+1} = (1 - 2\mu_k A_k) v_k \tag{15}
\]
From equation (15) we may write:
\[
\begin{bmatrix}
\nu_{k+1}(0) \\
\nu_{k+1}(1) \\
\vdots \\
\nu_{k+1}(L-2) \\
\nu_{k+1}(L-1)
\end{bmatrix}
= \begin{bmatrix}
1 - 2\mu_k & 0 & \cdots & 0 \\
0 & 1 - 2\mu_k & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 1 - 2\mu_k \\
\end{bmatrix}
\begin{bmatrix}
\nu_k(0) \\
\nu_k(1) \\
\vdots \\
\nu_k(L-2) \\
\nu_k(L-1)
\end{bmatrix}
\tag{16}
\]
where \( \lambda_i \) is the \( i \)th eigenvalue of the correlation matrix \( R_k \).
Equation (16) \( \Rightarrow \nu_{k+1}(i) = (1 - \mu_k^i) \nu_k(i) \Rightarrow \nu_{k+1}(i) = (1 - 2\mu_k^i)^{k+1} \nu_k(0) \). Since all eigenvalues of the auto-correlation matrix \( R \) are positive and real, the response \( \nu_{k+1}(i) \) will no oscillations. For stability or convergence of the algorithm, the magnitude of the geometric ratio of the above geometric series must be less than 1 for all \( i \).
\[
-1 < -2\mu_k^i < 1 \forall i \Rightarrow \mu < \frac{1}{\lambda_{\max}} \tag{17}
\]
where \( P_{sk} = \text{E}[x_k^2] \) is the average power of the input \( x_k \). Since the trace is just the sum of the L diagonal elements, \( \text{tr}(R_k) = L \mu_k \).
Note that:
(1) The norm of the gradient vector will be large initially and converge to a small value, potentially zero, at steady-state;
(2) The polarity of the gradient vector will generally be consistent during the early stage of the adaptive process and change frequently after the system converges;
(3) From equations (4), (6), (17) we have \( 0 \leq J(w_{k+1}) < J(w_*) \);
(4) In practice, if the ADC has M-bit data, we have \( J_{\max} \leq (2^M)^2 \) \( \tag{18} \)
(5) If the ADC has M-bit data, from equation (7) so we have
\[
\| J(k) \| = \left\| -2x_k e_k \right\| \leq 2^M \tag{21}
\]
Hence, from (17), (20), and (22), we have
\[
\frac{1}{J_{\max}} \leq \frac{1}{L \mu_k} \tag{23}
\]
\[
\frac{1}{\sum_{j=0}^{N-1} b_j |v_k(j) e_k(j)|} \leq \frac{1}{J_{\max}} \tag{24}
\]
\[
\frac{1}{2^M + \sum_{j=0}^{N-1} b_j |x_j(e_j(j))|} \\
\frac{1}{2^M + l \sum_{j=0}^{N-1} a_j e_j^2(j)}
\]

\[
\mu = \frac{\mu_{\text{max}}}{s_j + L^{-1}} \geq \mu_{\text{max}}
\] 

\[
\mu = \frac{\mu_{\text{max}}}{s_j + L^{-1}} \leq \mu_{\text{min}}
\] 

Hence, we may choose step-size by

\[
\mu = \frac{\mu_{\text{max}}}{s_j + L^{-1}} \geq \mu_{\text{max}}
\]

where: \( a_j \) and \( b_j \) are coefficients designed as the coefficients of the low-pass filter with cut-off frequency depends on the spectrum of input signal \( s \).

\[
\alpha = \frac{B}{2^M + l \sum_{j=0}^{N-1} b_j}
\]

and 

\[
\beta = \frac{A}{2^M + l \sum_{j=0}^{N-1} a_j}
\]

where: \( A \) and \( B \) are gain coefficients. We use the gain coefficients \( A \) and \( B \) to increase the role of \( J \) as well as the norm of the gradient vector, then we use \( \mu_{\text{max}} \) and \( \mu_{\text{min}} \) to guarantee bounded.

From the proposed algorithm (26), we have large step-size \( (\mu_{\text{max}}) \) at the early stages of the adaptive process and value of the step-size is decreasing to \( \mu_{\text{min}} \) when the system approaches convergence so that we can improve convergence speed and get good value of SNR; In the (26), \( a_j \) and \( b_j \) coefficients are designed as the coefficients of the low-pass filter with cut-off frequency depends on the spectrum of \( s \) so that the noise of \( J \) and the norm of the gradient vector are reduced.

3 Simulation Results

Simulations are performed in this subsection to show the advantages of the proposed algorithm. In simulations, the proposed algorithm (26) will be compared with the LMS, NLMS algorithms.

3.1 Simulation 1

In the first simulations, the noise \( x \) is a pseudorandom, zero-mean and unit variance Gaussian process. The noise signal \( n \) is a pseudorandom, zero-mean and unit variance Gaussian process. The signal is \( s(n)=1+0.01*\sin(2*\pi*1000*n) \). A frequency of \( s \) signal is 1kHz - in the human voice frequency range, with a length of 1000 samples.

Case 1: In this case, the parameters used in the new algorithm are: \( L = 4; N = 10; a_j = 0.1; b_j = 0.1; \alpha = 5.10^{-3}; \beta = 5.10 \), \( \mu_{\text{max}} = 0.15; \mu_{\text{min}} = 8*10^{-3} \); The parameter used in the LMS algorithm is \( \mu=0.01 \); The NLMS algorithm is \( \mu_{\text{NLMS}} = 0.01 \). The simulation results are presented in Figure 2 and Table 2. Figure 2 shows the MSE obtained from Monte Carlo simulations (100 runs), and the values of SNR are shown in the Table 2.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Before filtering</th>
<th>New algorithm</th>
<th>NLMS</th>
<th>LMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0084</td>
<td>21.1687</td>
<td>17.0503</td>
<td>15.8665</td>
</tr>
<tr>
<td>2</td>
<td>-0.0031</td>
<td>21.2747</td>
<td>20.7448</td>
<td>15.5701</td>
</tr>
<tr>
<td>3</td>
<td>0.2937</td>
<td>19.4356</td>
<td>17.1666</td>
<td>14.5539</td>
</tr>
<tr>
<td>4</td>
<td>-0.0359</td>
<td>14.3001</td>
<td>13.3513</td>
<td>11.7474</td>
</tr>
<tr>
<td>5</td>
<td>0.3259</td>
<td>19.1568</td>
<td>18.3731</td>
<td>12.6130</td>
</tr>
</tbody>
</table>

Figure 2. Comparison of MSE and convergence rate of new algorithm and LMS, NLMS in case 1

Case 2: This is the same as the previous case but the parameter used in the LMS algorithm is \( \mu=0.08 \); The NLMS algorithm is \( \mu_{\text{NLMS}} = 0.08 \). The simulation results are presented in Figure 3 and Table 3. Figure 3 shows the MSE obtained from Monte Carlo
simulations (100 runs), and the values of SNR are shown in Table 3.

Table 3. SNR (dB) in case 2

<table>
<thead>
<tr>
<th>S.No</th>
<th>Before filtering</th>
<th>New algorithm</th>
<th>NLMS</th>
<th>LMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1700</td>
<td>20.6856</td>
<td>9.4027</td>
<td>7.1442</td>
</tr>
<tr>
<td>2</td>
<td>0.4142</td>
<td>21.9611</td>
<td>9.1564</td>
<td>7.1248</td>
</tr>
<tr>
<td>3</td>
<td>-0.1061</td>
<td>17.0344</td>
<td>8.9119</td>
<td>6.5772</td>
</tr>
<tr>
<td>4</td>
<td>0.1751</td>
<td>20.5712</td>
<td>9.5833</td>
<td>7.2531</td>
</tr>
<tr>
<td>5</td>
<td>-0.3751</td>
<td>19.8617</td>
<td>8.6202</td>
<td>6.3245</td>
</tr>
</tbody>
</table>

According to the simulation results, the performance of the LMS, NLMS methods with parameters as in case 1 has a small MSE and a large SNR, but the convergence rate is very slow. In case 2, we use the large step-size for LMS and NLMS algorithms; therefore the convergence rate increases. Moreover, when we use the large step-size for LMS and NLMS, the SNR value is reduced. The results shows that the new algorithm has better SNR and converges rate than LMS, and NLMS algorithms.

Figure 3. Comparison of MSE and convergence rate of new algorithm and LMS, NLMS in case 2

3.2 Simulation 2

In this simulation, the signal and noise are non-stationary random process. The noise $x$ and $n$ are a sound of devices in coal mines. They are recorded with a sampling rate of 8,000 Hz, Figure 5. The signal $s$ is samples of a speech signal which is available from web site www.voiptroubleshooter.com, and the file name is “OSR_us_000_0030_8k.wav”, Figure 4. It has a sample rate $F_s$=8,000 Hz. Figure 6 shows that the spectrum of the noise overlaps spectrum of the human voice. The parameters used in the new algorithm are: $L = 4; N = 10; a_j = 0.1; b_j = 0.1; \alpha = 5 \times 10^{-1}; \beta = 2 \times 10^{-4};$ the parameter used in the LMS algorithm is $\mu=0.01;$ The NLMS algorithm is $\mu_{NLMS} = \frac{0.08}{\|h\|_2 + 1}$. In this simulation, we performed experiments with the SNRs (before filtering) are scaled -14.0990 dB, -8.0784dB (Figure 7), -0.1196 dB; 2.8030 dB; 5.9010 dB,... The simulation results are presented in Figure 8, 9, Figure 10 and Table 4. Figure 8 shows the behavior of the LMS algorithm, the amplitude of the noise is reduced along the horizontal axis. The Output_LMS.wav shows that the noise reduce gradually, and disappeared about 5 seconds. Figure 9 shows the behavior of the NLMS algorithm. In this figure, the amplitude of the noise is reduced quickly than that of Figure 8. However, we still hear some noise when we play Output_NLMS.wav file because the SNR value is not high. It is clear from Figure 10 and Table 4 that the new LMS algorithm performs better than the LMS, NLMS in term of the SNR as well as the convergence rate. It looks like we do not hear the noise when we play the Output_New_Algorith.wav file.

Figure 4. Original signal (OSR_us_000_0030_8k.wav) with 110.000 samples

Figure 5. Noise (sound of engine) with 110.000 samples
MINING, BUILT INFRASTRUCTURE AND HUMAN FACTORS

Figure 6. Amplitude spectrums of Original signal and Noise, in the range between 0 and 500 Hz

Figure 8. Output signal (Output_LMS.wav) with 110.000 samples, SNR=14.8680dB

Figure 9. Output signal (Output_NLMS.wav) with 110.000 samples, SNR = 11.9066dB

Figure 10. Output signal (Output_New_Alg_wav) with 110.000 samples, SNR = 18.6652 dB

<table>
<thead>
<tr>
<th>S.No</th>
<th>Before filtering</th>
<th>New algorithm</th>
<th>NLMS</th>
<th>LMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-14.099</td>
<td>15.1303</td>
<td>11.5308</td>
<td>10.3922</td>
</tr>
<tr>
<td>2</td>
<td>-8.0784</td>
<td>18.6652</td>
<td>11.9066</td>
<td>14.8680</td>
</tr>
<tr>
<td>3</td>
<td>-4.5565</td>
<td>19.8141</td>
<td>11.9787</td>
<td>16.6313</td>
</tr>
<tr>
<td>4</td>
<td>-2.0578</td>
<td>20.3001</td>
<td>12.0038</td>
<td>17.4678</td>
</tr>
<tr>
<td>5</td>
<td>-0.1196</td>
<td>20.5445</td>
<td>12.0153</td>
<td>18.6311</td>
</tr>
<tr>
<td>6</td>
<td>1.4641</td>
<td>20.6831</td>
<td>12.0215</td>
<td>19.0341</td>
</tr>
<tr>
<td>7</td>
<td>2.8030</td>
<td>20.7687</td>
<td>12.0252</td>
<td>19.3497</td>
</tr>
<tr>
<td>8</td>
<td>3.9628</td>
<td>20.8252</td>
<td>12.0275</td>
<td>19.5622</td>
</tr>
<tr>
<td>9</td>
<td>4.9859</td>
<td>20.8643</td>
<td>12.0291</td>
<td>19.8410</td>
</tr>
<tr>
<td>10</td>
<td>5.9010</td>
<td>20.8924</td>
<td>12.0302</td>
<td>19.5982</td>
</tr>
<tr>
<td>11</td>
<td>7.4847</td>
<td>20.9292</td>
<td>12.0316</td>
<td>18.6738</td>
</tr>
<tr>
<td>12</td>
<td>8.1799</td>
<td>20.9417</td>
<td>12.0321</td>
<td>18.6995</td>
</tr>
<tr>
<td>13</td>
<td>8.8236</td>
<td>20.9516</td>
<td>12.0324</td>
<td>18.7200</td>
</tr>
<tr>
<td>14</td>
<td>9.4229</td>
<td>20.9595</td>
<td>12.0327</td>
<td>18.7366</td>
</tr>
</tbody>
</table>

4 Conclusion
In this paper the process of noise removal from speech signal to improve a quality of broadcast-telephone system by using the new LMS algorithm is presented. The performance of the new LMS algorithm is better than the LMS, NLMS in term of the SNR and the convergence rate because the step-size (26) is controlled by both the smoothed mean-square error and the norm of the smoothed gradient vector. Simulations show that this algorithm can obtain not only the fast convergence rate but also the large SNR. Thus the theoretical and simulation results prove that the proposed algorithm is the promising method for removing the noise in the signal. Future work will use the new algorithm (26) for echo cancellation to improve audio quality for VoIP-Wifi-Phones use in underground coal mines.

References

Abstract -

The objective of this research is to develop, optimize and evaluate an approach for automatically generating rigging path guidance in order to provide sufficient assistance on the user interface of tele-operated crane. Besides operational safety and efficiency, this approach focuses on solving operability issues of following path guidance on a tele-operated crane system. To provide path guidance that can realistically be performed in a tele-operated rigging scenario, three factors are proposed for consideration in the development of such an approach: Visibility, Control Features, and the Dynamic Environment. A simulation platform in the virtual environment, called PathGuider, is developed for evaluating the use of the proposed path guidance considerations. Demonstrations for visibility and the dynamic environment issues have been performed and they show that the guidance path can be re-planned in real-time once the guided trajectories are being sheltered by surrounding buildings from operators’ view or dynamic objects interfered with it. To evaluate the efficiency of the developed re-planning algorithm, a benchmark has been developed. The results show that the calculation time of cases considered are all within real-time span. The proposed path guidance approach reveals its efficiency and feasibility in providing useful assistance to tele-operated crane’s operators.

Keywords -
Tele-operation; Optimization; Crane Operation; Rigging Path Planning; PRM

1 Introduction

Tele-operation, the remote control of machines, has in recent years been increasingly applied in practice and been the subject of research [1]. Tele-operation allows for a reduction of human presence at a site, manipulation on avatars to finish tasks and hence significantly reduces safety risks. Tele-operation technologies have been widely applied in diverse fields such as rescue missions, mining, and space exploration. They also show promising results on improving productivities of manufacturing fields.

With the increase in research effort focused on tele-operation applications for improving construction operations, the interface design of construction equipment remains an important topic for further study. Due to variation in operation functions and automation assistance with different construction equipment, the interface design can directly influence the efficiency and effectiveness with which operators can finish their tasks. Yokokohji et al. [2] surveyed state-of-art robots for rescue missions and proposed design guidelines for the tele-operation interface, saying that the global view which covers the remote machine itself should always be included in the user interface for operators to monitor the overall situation while robots are working in challenging environments. This has also motivated further examination of the relationship between humans and tele-operated user interfaces in this research.

During the last two decades, a significant amount of research in the area of rigging path-planning problems for construction cranes has been aimed at using computers to automatically generate the rigging path solutions. For current development of motion planning in construction, it is usually utilized in simulations at the planning stage before actual constructions. Sivakumar et al. [3] used heuristic functions to seek out collision-free paths for dual-crane cooperative rigging activities, while Ali et al. [4] targeted the same problems but instead used a genetic
algorithm and achieved more efficient results. Kang and Miranda [5] utilized the probabilistic roadmap method (PRM) and developed and implemented three different path-planning methods to handle different path-planning problems while minimizing computation cost. Zhang et al. [6] have analyzed and implemented path re-planning methods for cranes using rapidly randomized tree (RRT) methods. Chang et al. [7] developed a planning method simplified search dimension of rigging elevation to efficiently solve cooperative rigging activities.

Although extensive research has been conducted on the above-mentioned aspects, the path-planning methods for tele-operation and operational feasibility for real crane operators are rarely discussed. They are two important issues influencing rigging path guidance for remotely and manually controlled cranes. Removing the operator from the field introduces a challenge as a result of limitations in viewing the operations. Due to inevitable restrictions in the field of view of the remote camera(s) and in the quality of the visual displays fed back from the remote worksite, the operator is often unable to maintain a level of situational awareness sufficient for safe and efficient task execution [8]. Although multiple cameras mounted to view the field of operations would be of benefit to remote operation, guidance paths should also be generated that consider the field of vision of the cameras. In this research, an approach is developed to automatically generate rigging paths for operators to remotely control the cranes in real-time. The guidance paths are generated using the PRM in C-space to enable a collision-free operation field, and taking into account the three specified feasibility factors for tele-operation.

2 Automatic Rigging Path Guidance

Considering the topics of automatically generating appropriate rigging paths by computers, many research have developed and provided variable solutions, which shows great impacts on improving construction automation. However, current research usually focuses on safety and efficiency aspects rather than operational feasibility issues and tele-operated scenarios. As tele-operation guidance of this research, the generated paths can only be useful if they fit the common sense of crane operations and operators are willing to reference them. So operational feasibility of the suggested paths becomes an essential issue and we put focus on it during developing the proposed approach.

2.1 Pilot Test

In order to solve the feasibility issues of suggested rigging paths, the requirement of operable motion has been identified. Actual end users participated in a usability test to evaluate a developed user interface (UI) [9] for a tele-operated crane. In this test, a KUKA KR 16 robot arm was used as a crane prototype and four monitors displaying different angles of view (top, right-side, left-side, and global view) are also provided to remotely display conditions on the construction site model. The user interface with augmented reality (AR) technology–enhanced rigging path guidance can be seen in the top and global views. This path is generated automatically by simply using PRM in Cartesian space.

Five professional crane operators were recruited and participated in an in-depth user test to verify the usability and feasibility of the developed UIs. All the operators held professional licenses from the Industrial Safety and Health Association (ISHA) of Taiwan, and each had over ten years of experience. The testing times ranged from 63 to 92 minutes. The operators were asked to execute the simulated rigging tasks previously designed, and were then interviewed after completing the tasks.

From the observations of the tests and the subsequent interviews, it was found that one suggestion was most often provided by the operators, which was that the virtual rigging path guidance should be improved. The path taken during rigging depends on the style of each operator. However, all operators agreed that paths are not simply composed of straight lines, and must be generated according to the natural movement of the crane so that they can be easily followed. Guidance paths if possible should not be generated over the heads of field workers or existing buildings, and should be kept at low elevations. The view angle can also limit an operator’s depth perception if the guidance path is only shown on the top and global views. Considering the feedback from all test participants, it was concluded that three feasibility factors need to be considered to improve the current mechanisms of path guidance generation: visibility, control features, and the dynamic environment. The basic principle and these factors are described below.

2.2 Approach Overview

The proposed approach is basically established on a famous motion planning method called Probabilistic Roadmap (PRM) [10]. All the algorithms, as path planning considerations, developed in this research can be integrated into the process procedure of PRM and further considerations can also extend by the same way in the future. The general procedure of PRM can be seen as Figure 1.

The user should identify the starting and the end point in the space first. Then, the PRM method randomly samples geometrical points in the space and combine them with adjacent ones including the starting and end point. All the sampled points should not overlap with other objects within the space. It then becomes a graph structure $G$ that every geometrical point represents vertex $V$ and the connections between each vertex represent edge $E$. The formulation of a graph structure shows as follows:
\[ G = (V, E) \]  

After a graph structure established, it becomes an optimization problem which maximizes users’ requirements to find out edge combinations between the starting and end point by using graph traverse algorithms. In the common situation, distance between two vertices represents the cost of their edge. And the shortest path algorithms have capability to find out the shortest solution with minimum cost edges passed by, which optimally solves the demand of efficient purpose. In other words, the solution shows efficient rigging path suggestion regarding to moving distance and time spent in this research.

Figure 1. The general procedure of the Probabilistic Roadmap (PRM): (a) assign a start and an end point and randomly sampling on the space; (b) connect points as a graph; and (c) traverse the graph to find a collision-free and feasible path.

In addition, collision detection becomes an important process needed to be considered during path finding procedure. It can make sure each of path segments is not collided with surrounding obstacles in the space. The paths blocked by obstacles cannot be operated and is not correct suggestion as candidate solutions for guiding operators. The collision problems can be prevented by ray tracing method, which determines whether the two vertices on two sides of an edge can "see" each other or not. In this research, we adopted this method for solving collision-free issues.

Besides securing the potential collisions of the rigging objects, all the movement of crane components should be collision-free as well. The method to guarantee the collision-free trajectories of all the crane components is Configuration Space (C-space) transformation. In this research, as shown in Figure 2, a tower crane is used as a rigging platform to develop the proposed approach. The position of the rigging object can be represented with respect to Cartesian coordination system \((X, Y, Z)\). It can also be represented with respect to the tower crane’s configuration. A configuration of the tower crane is given by \((r, \theta, l)\), which is a 3-DOF manipulator. The symbol \(r\) denotes the radius of the boom rotation as well as the current distance between the mast and the trolley; \(\theta\) represents the rotation angle of the tower crane, and \(l\) represents its current hoisting length. All the movements of the tower crane become one single point on the C-space. And every obstacle in Cartesian space transforms to the ones in different shape in the C-space, which is called C-obstacle. The C-obstacle represents a batch of postures \((r, \theta, l)\) of crane we should avoid to perform because collisions can happen on somewhere of the crane.

\[ f(x) = g(x) + h(x) \]  

How to identify the cost of a path passing vertex \(x\), is decided by two parts. The first part is \(g(x)\), which represents the function calculating the exact cost from the starting vertex to current vertex \(x\). It is the same as conventional search principles which focus on what choices we have based on the current search status. About the second part \(h(x)\), it represents a heuristic estimate function to calculate the expected cost from current vertex \(x\) to the destination. Unlike Dijkstra’s algorithm which guarantees optimal result but requires considerable computing time, A* search algorithm provides a balance mechanism between efficiency and performance by eliminating unlikely solutions with heuristic judgments and at the same time keeping...
optimization promise. It is thus integrated as the primary search stagey in this research.

The developed path finding algorithm, named \textit{PathSuggestor} in this research, can be seen as Table 1. Assuming that the C-space transformation has been performed and all the C-obstacles have been identified at the pre-processing stage, the user initials a starting point and an end point as input parameters. Then we randomly sample feasible points in the C-space. For establishing a graph structure, every connections (edges) between each sampled adjacent point need to be examined to see whether they are connected without collided with any C-obstacles. We use \textit{OperableCheck} function to archive this objective. Afterward, we get a cost matrix representing the weight of each edge on a graph. It then is used as a parameter of \textit{AStarShortestPath} function. The function is typically an \textit{A}* search algorithm to find out an efficient path from \( P_{\text{start}} \) to \( P_{\text{end}} \), which minimizes the cost in sum of every edge it passed by.

\begin{table}[h]
\centering
\caption{PathSuggestor algorithm}
\begin{tabular}{l}
\hline
1: \( n \leftarrow \text{GridSampling()} \) \\
2: \( \text{// Randomly sampling vertices in C-space} \) \\
3: \( M \leftarrow \text{OperableCheck}(n) \) \\
4: \( \text{// Check collisions among every edge} \) \\
5: \( \text{Count} \leftarrow 0 \) \\
6: \( \text{WHILE} \ \text{Count} \neq \theta \) \\
7: \( P \leftarrow \text{AStarShortestPath}(P_{\text{start}}, P_{\text{end}}, M) \) \\
8: \( \text{// \text{A}* shortest path search} \) \\
9: \( \text{IF} \ \text{CheckVisibility}(P) = \text{FALSE} \) \\
10: \( \text{// IF P doesn’t fit, search again} \) \\
11: \( \text{Count} \leftarrow \text{Count} + 1 \) \\
12: \( \text{IF} \ \text{Count} = \alpha \) \\
13: \( \text{// IF Count equals to threshold, abort} \) \\
14: \( P \leftarrow \text{NULL} \) \\
15: \( \text{BREAK} \) \\
16: \( \text{CONTINUE} \) \\
17: \( \text{ELSE} \) \\
18: \( \text{BREAK} \) \\
19: \( \text{//Otherwise, P fits the requirement} \) \\
20: \( \text{END WHILE} \) \\
21: \( \text{RETURN} P \) \\
\hline
\end{tabular}
\end{table}

For considering operational feasibility on tele-operation scenarios, \textit{PathSuggestor} approach is developed for handling \textit{visibility}, \textit{control features} and the \textit{dynamic environment} concerns. They are described in the following sub-section individually. It should be noticed that \textit{PathSuggestor} approach or even original PRM algorithm did not consider the effects from physic-based behaviors, such as rigging cable swing, in this research.

2.3 Visibility

Considering operational feasibility of suggested rigging paths, viewing limitations of cameras on the tele-operated crane need to be addressed and eliminated from the path searching space. With the lack of visualized information and immersive senses on tele-operation scenarios, these remote cameras become operators’ “eyes” and operations are significantly depend on how these camera are set around. Without sensing clear images on the space, it is useless to guide operator passing those unknown regions. So the suggested rigging paths should not pass blind zones where the visions of cameras are blocked by surrounding buildings or objects at construction site. Even in the conventional operation scenarios, it is the rule of thumb for rigging activities as well.

To preserve the operator’s depth perception and visual information, paths that are partially or completely unmonitored or monitored by one camera only are eliminated from the tele-operation scenarios. Every geometrical points on the candidate path is examined in term of visibility using the ray tracing method to observe if any obstacles are located between the point and camera positions. If the views are blocked, the candidate path will be discarded and another path is re-planned using the same evaluation process. A path may only be generated within regions monitored by at least two cameras to ensure that the operability and safety of users following the path is not compromised.

As for \textit{PathSuggestor} algorithm mentioned in previous sub-section, a \textit{Count} variable has been added to count the times of path re-search. After calling \textit{AStarShortestPath}, the system will call \textit{CheckVisibility} to check whether the visibility of candidate path encounters limitations or not. In the \textit{CheckVisibility}, we check every combination of vertices on candidate path \( P_i \) and camera positions \( P_j \). If one of these combinations cannot see each other, discard candidate path and search another solution again with increment value of \textit{Count}. If \textit{Count} doesn’t exceed the threshold \( \alpha \), the re-search iteration will continue until the candidate path fits visibility requirements. Otherwise, it will stop and report that no solution can be found to users.

2.4 Control Features

From the observations and feedback of crane operators, it was found that operators tend to operate the crane along trajectories which allows them to maintain a low elevation and not to cross over any existing objects on the construction site. This is due to safety concerns regarding potential falling payloads and an attempt to minimize any potential damage to the surrounding environment. In addition, a general control style of the operators is that they tend to maintain the same rigging direction unless the environment changed or they encountered obstacles.

In consideration of these issues, the edges of sampled
graph in C-space were weighted as in Equation (3):
\[ C'_{ij} = (W_E + W_O + W_C + 1)C_{ij} \]  
(3)

where \( C_{ij} \) denotes the traveling cost from point \( i \) to \( j \) in the C-space sampled graph, and \( C'_{ij} \) denotes weighted cost considering the control features described above. Meanwhile, \( W_E \) represents the weight of the rigging elevation, which is directly proportional to the difference of elevation value \( z \) between point \( i \) and \( j \) in Cartesian space. The bigger difference between elevation of point \( i \) and \( j \), the higher weight influencing the travel cost from point \( i \) to \( j \) is:
\[ W_E \propto z_j - z_i \]  
(4)

The weight \( W_O \) governs crossing over existing objects on the construction site. Every object, including dynamic objects, should be examined in relation to whether or not their overhead Cartesian space overlaps the trajectory of the candidate point pair \( n_i n_j \). If so, a penalty value \( \alpha \) will be assigned between the related C-space point \( i \) and \( j \). If not, the weight will not influence the cost in this case, as:
\[ W_O = \begin{cases} \alpha, & \text{Overhead } n_i n_j = \text{true} \smallskip \\
0, & \text{Overhead } n_i n_j = \text{false} \end{cases} \]  
(5)

The weight \( W_C \) governs changing the rigging direction. Less rigging actions decrease the risk during rigging operations, fit with a human control style, and save on construction costs. The value of the weight from point \( i \) to \( j \) depends on the direction of the previous candidate point pair \( n_i n_j \). Here, \( n \) denotes the neighboring point of \( n \), which is selected previously in a candidate path. If the direction of \( n_i n_j \) does not equal that of \( n_i n_j \), a penalty value \( \beta \) will be assigned between the related C-space points \( i \) and \( j \). If not, then the weight will not influence the cost in this case, as:
\[ W_C = \begin{cases} \beta, & n_i n_j \neq n_i n_j \smallskip \\
0, & n_i n_j = n_i n_j \end{cases} \]  
(6)

The weighting mechanism is integrated into the search method and applied as a heuristic function (as OperableCheck) for evaluating the cost of candidate point pairs on the sampled graph. The weighting terms of Equation (3) can have an influence on each other. For example, if the candidate point pair along an elevation direction is selected at the beginning of the search process, the value of \( W_C \) influences the search results to continue selecting point pairs in the same direction. However, the increasing value of \( W_E \) will eventually affect the search process in such a way as to change the searching direction.

In addition to safety concerns, human control capacity must also be considered during crane operations. Crane operators can manipulate, at most, two DOFs of a crane simultaneously. For example, though manipulations like rotating the boom while hoisting allow for a decreased operation time, they are subject to the operators’ perception capacity. However, the paths generated by the PRM may not take this factor into consideration and can sometimes involve the manipulation of more than two DOFs. This may not be feasible in practice and operators need to create an alternative rigging plan by themselves. Even the suggested path is feasible to operate, the operator needs to pay more attention on extra DOFs which may exceeds operator’s perception capacity and cause potential safety issues in crane operations.

To deal with the capacity problem, it is proposed to use a grid sampling method in C-space when searching for a feasible path. As shown in Figure 3, the edges of the graph can only be linked to adjacent points. Points must be linked by creating a rectangular shape for a single DOF configuration, and points can be linked diagonally for a 2-DOF configuration. The path found on this graph is guaranteed to be within the capacity of human manipulation to perform by using combinations of the control sticks.

![Figure 3. Overview of the grid sampling method](image)

The grid sampling method is implemented as GridSampling function in the PathSuggestor algorithm. In each C-space’s dimension, we randomly sample a number of digits to become an element serial. The length of element serials is ceiling of \( \sqrt{n} \), where \( n \) denotes the desired number of samples. After all three dimensions are sampled, the combination of all three element serials becomes the result matrix of grid sampling. It then returns for furthering operable check and the search graph building procedures.

2.5. Dynamic Environment

When considering the complexity of a modern construction site, it is very likely that rigging tasks can be interfered with by the surrounding elements, such as
other construction vehicles or machines. This is
potentially a very serious problem which needs to be
specially handled in tele-operation scenarios. Without a
human presence, it is difficult to handle any unexpected
scenarios if the construction plans for the tele-operated
machines on the site are not well-established. The
participants of the test also pointed out the fact that the
re-planning ability of rigging plans is a requirement of a
tele-operation system to handle the dynamics of a
construction.

Considering the usage of PRM in this research, it is
time-consumed for handling dynamic objects. The large
computational demand is because once the status of an
object is updated, the entire C-space needs to be rebuilt in
order to identify any potential collision regions. To
improve the computational efficiency, the Octree
structure for collision detection was integrated. This can
be used to partially rebuild those regions of space that
have potentially changed due to dynamic objects. In the
redesigned C-space, each element \((r, \theta, l)\) contains the
information about the relative Cartesian position \((x, y, z)\)
and is categorized according to the geometric relationship
of the Cartesian space.

Firstly, all the elements in the C-space are input into
the tree root. The space is then split into eight octants,
segmenting the related Cartesian space, and the elements
covered in each octant are recorded into the
corresponding child structure. Secondly, each octant is
split and the elements are recorded in the same way, and
this is repeated until the desired level is achieved. The
layout of the octree structure with respect to the current
environment is built in this fashion. By noting the
positions of any dynamic objects (can be done by
localization technologies in construction) and traversing
octree structure, the groups of elements (C-space regions)
whose Cartesian positions are near the dynamic objects
can be retrieved. Next, it is attempted to rebuild these
elements with the collision detection methods. In an ideal
case, the number of elements need to be rebuilt can be
reduced from \(N\) to approximately \(N/8M\). The variable \(N\)
represents the number of elements (resolution) in the
entire C-space. And the variable \(M\) represents the level
number of octree structure. The algorithm of this
updating procedure, named \(\text{UpdatePartialSpace}\), is
described as Table 2.

The following algorithm, as shown in Table 3, is called
at every time frame of crane operation. Every time when
an operator is performing operation according to the
suggested path, the algorithm will monitor each moving
obstacles in the surrounding environment. It can be done
by attaching markers or devices and use positioning
technologies to track them, such as GPS or UWB. The
algorithm can update the cost matrix \(M\), generated by
\(\text{PathSuggestor}\) algorithm, to \(M'\) according to
\(\text{UpdatePartialSpace}\) function.

### Table 2. \text{UpdatePartialSpace} algorithm

<table>
<thead>
<tr>
<th>Algorithm: \text{UpdatePartialSpace}(\text{S}, \text{P}_{\text{obstacle}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{S}_{\text{space}}): Partial C-space needs to be updated.</td>
</tr>
<tr>
<td>(\text{S}): Partial Cartesian space related by (\text{S}_{\text{space}}).</td>
</tr>
<tr>
<td>(v): Split sub-space of (\text{S}).</td>
</tr>
<tr>
<td>(n): Sampled geometrical points inside (\text{S}_{\text{space}}).</td>
</tr>
<tr>
<td>(\text{P}_{\text{obstacle}}): The Cartesian position of the moving obstacle.</td>
</tr>
<tr>
<td>1: IF (\text{S}.\text{Covered}(\text{P}_{\text{obstacle}}) = \text{TRUE})</td>
</tr>
<tr>
<td>2: // If (\text{S}) includes the moving obstacle</td>
</tr>
<tr>
<td>3: IF (\text{S}.\text{SubSpace}() = \text{NULL})</td>
</tr>
<tr>
<td>4: // If (\text{S}) is leaf node in octree, return related (n)</td>
</tr>
<tr>
<td>5: (\text{S}_{\text{space}} \leftarrow \text{Mapping}(\text{S}))</td>
</tr>
<tr>
<td>6: (n \leftarrow \text{GetSamples}(\text{S}_{\text{space}}))</td>
</tr>
<tr>
<td>7: RETURN (n)</td>
</tr>
<tr>
<td>8: ELSE</td>
</tr>
<tr>
<td>9: // Otherwise, recursively search its sub-spaces</td>
</tr>
<tr>
<td>10: FOREACH (s) in (\text{S}.\text{SubSpace}())</td>
</tr>
<tr>
<td>11: (n' \leftarrow n + \text{UpdatePartialSpace}(s, \text{P}_{\text{obstacle}}))</td>
</tr>
<tr>
<td>12: ELSE</td>
</tr>
<tr>
<td>13: // If the moving obstacle is out of (\text{S}), return empty</td>
</tr>
<tr>
<td>14: (n \leftarrow \text{NULL})</td>
</tr>
<tr>
<td>15: RETURN (n)</td>
</tr>
</tbody>
</table>

### Table 3. \text{PathRePlanning} algorithm

<table>
<thead>
<tr>
<th>Algorithm: \text{PathRePlanning}(g), Update cost matrix according to the position of moving obstacles, and re-plan the path if it is necessary.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{P}_{\text{obstacle}}): The Cartesian position of the moving obstacle.</td>
</tr>
<tr>
<td>(L): The list of moving obstacles.</td>
</tr>
<tr>
<td>(n): The list of sampled vertices needed to be updated.</td>
</tr>
<tr>
<td>(P): The current suggested path.</td>
</tr>
<tr>
<td>1: FOREACH (\text{P}_{\text{obstacle}}) in (L)</td>
</tr>
<tr>
<td>2: // Collect sampled vertices needed to be updated</td>
</tr>
<tr>
<td>3: (n' \leftarrow n' + \text{UpdatePartialSpace}(\text{S}, \text{P}_{\text{obstacle}}))</td>
</tr>
<tr>
<td>4: (M' \leftarrow \text{PartialOperableCheck}(n', M))</td>
</tr>
<tr>
<td>5: // Update partial cost matrix</td>
</tr>
<tr>
<td>6: FOREACH (P) in (P)</td>
</tr>
<tr>
<td>7: IF (n'.\text{Contain}(P) = \text{TRUE})</td>
</tr>
<tr>
<td>8: // If a point on the path needs to be updated</td>
</tr>
<tr>
<td>9: (P \leftarrow \text{AStarShortestPath}(\text{P}<em>{\text{source}}, \text{P}</em>{\text{end}}, M'))</td>
</tr>
<tr>
<td>10: // A* shortest path search</td>
</tr>
<tr>
<td>11: BREAK</td>
</tr>
<tr>
<td>12: RETURN (P)</td>
</tr>
</tbody>
</table>

Once the list of sampled points needed to be updated
\((n')\) contains the point \((P)\) along the suggested path,
which means the moving obstacle is intersecting or near
somewhere along the suggested path, the algorithm will
re-plan another path according to current rigging position
for operator’s reference. As such, this approach makes
the re-planning process more efficient when the positions
of dynamic objects change, particularly in the case of a
complex construction site containing many obstacles.

### 3 Implementation and Evaluation

To demonstrate and evaluate the feasibility of the
proposed path guidance approach, a simulation system
was implemented namely \text{PathGuider} in a virtual
environment. The C# language was used to implement
the system on a .NET developer platform. Also included was a graphic engine, implemented with the Microsoft® XNA framework to simulate the rigging tasks and environment, and a physics engine, implemented with the PhysX library to handle the collision detection and ray tracing processes. The implemented algorithms include basic PRM method, PathSuggestor, UpdatePartialSpace and PathReplanning. As for the OperableCheck function, \( W_E \) is assigned as a function: \( W_E = 0.002h \), where \( h \) denotes elevation. \( \alpha \) equals to 0.2, and \( \beta \) equals to 0.2. Further study about how to choose optimized weighting configuration will be conducted in the future.

In the test scenario of PathGuider system (as shown in Figure 4), a 3-DOF tower crane, a mobile crane, and an under-constructed steel frame building similar to that in the pilot test were set up. The movement of the tower crane was guided by the suggested paths, and the mobile crane, which is free to move around the construction site, will be treated as a dynamic object in the environment.

For evaluating whether the proposed path guidance algorithms have achieved our design intentions, we made demonstrations and a benchmark for adopting algorithms with visibility and the dynamic environment consideration. An i7 Dual-Core 2.8 GHz computer with 4 GB memory was used in these tests. In searching for a suitable path, the sampling number of point was 500. The resolution of the C-space was 200×180×200.

By testing the algorithm considering visibility issues, we selected single camera scenario to demonstrate how it works. In the beginning of the demonstration, the path planning algorithm generates a rigging path from a specific starting point to a destination. A first perspective view is set up and the user can control the camera’s viewing angle and position freely. By moving the camera, the user is expected to see real-time path regeneration happened while the viewing direction of camera is blocked by existed building component.

To evaluate the efficiency of the developed re-planning algorithm, a benchmark for the dynamic environment has been developed. The demonstration is set up as test scenario mentioned above. Among the pre-processing stage of establishing C-space, we use five octree structures with different levels (from 3 to 7). In the virtual environment, a sphere-like object was randomly placed in the areas surrounding the guidance paths in order to observe the calculation time for the re-planning process and ensuring collision-free path generations. For each octree structure, we recorded the calculation time for 400 times of the conditions which the placed position of the sphere-like object did not trigger the re-planning action. Similarly, the calculation time for 100 times of the conditions which the placed position of the sphere-like object triggered the re-planning action has been recorded. In this benchmark, we are able to know the relationship between established octree’s level and re-planning calculation time. Also, the efficiency of the proposed re-planning algorithm can be identified.

4 Result and Discussion

The demonstration of the proposed algorithm considering the visibility issues is illustrated at Figure 5. From a conventional view of the tower crane cabin, the planned path has been partially blocked by existed steel frame structure (Figure 5a). The proposed algorithm responses in real time to re-plan another path which can be observed completely for the view (Figure 5b). This algorithm is also valid with more views in a tele-operated scenario.

By testing the algorithm considering visibility issues, we selected single camera scenario to demonstrate how it works. In the beginning of the demonstration, the path planning algorithm generates a rigging path from a specific starting point to a destination. A first perspective view is set up and the user can control the camera’s viewing angle and position freely. By moving the camera, the user is expected to see real-time path regeneration happened while the viewing direction of camera is blocked by existed building component.

To evaluate the efficiency of the developed re-planning algorithm, a benchmark for the dynamic environment has been developed. The demonstration is set up as test scenario mentioned above. Among the pre-processing stage of establishing C-space, we use five octree structures with different levels (from 3 to 7). In the virtual environment, a sphere-like object was randomly placed in the areas surrounding the guidance paths in order to observe the calculation time for the re-planning process and ensuring collision-free path generations. For each octree structure, we recorded the calculation time for 400 times of the conditions which the placed position of the sphere-like object did not trigger the re-planning action. Similarly, the calculation time for 100 times of the conditions which the placed position of the sphere-like object triggered the re-planning action has been recorded. In this benchmark, we are able to know the relationship between established octree’s level and re-planning calculation time. Also, the efficiency of the proposed re-planning algorithm can be identified.

4 Result and Discussion

The demonstration of the proposed algorithm considering the visibility issues is illustrated at Figure 5. From a conventional view of the tower crane cabin, the planned path has been partially blocked by existed steel frame structure (Figure 5a). The proposed algorithm responses in real time to re-plan another path which can be observed completely for the view (Figure 5b). This algorithm is also valid with more views in a tele-operated scenario.
5 Summary and Conclusion

A rigging path-planning approach modified from PRM for tele-operated scenarios was developed in this research. By conducting pilot test with crane operators on a prototype tele-operated crane, three operational feasibility factors were identified: visibility, control features, and the dynamic environment. In consideration of the visibility issues, the guidance path can only be generated within regions monitored by at least two cameras, in order to ensure the user’s operability when they are following the paths. A weighting mechanism for control features to adjust the travel cost between every configuration space (C-space) point pair was developed as well. By adapting the selection mechanism, the guidance path can avoid crossing over existing objects, maintain rigging directions, and maintain low elevations as far as is possible. The suggested path can also be followed with at most 2 DOFs movement under operators’ perception capacities. In consideration of the dynamics of a construction site, an octree structure for partially rebuilding the C-space was integrated to speed up the re-planning process.

To demonstrate and evaluate the developed approach for rigging path guidance, a simulation system, named PathGuider, was implemented with a test scenario similar to that of the pilot test for crane operators. Two demonstrations, showing the real-time path re-planning ability and securing visibility of the proposed algorithms, were performed and presented promising results. A benchmark was developed for measuring computation time of re-planning with respect to the octree level numbers. The results showed that the re-planning process using 500 sampled points to ensure a collision-free guidance could be completed within 200 ms. This shows that real-time rigging path planning for tele-operated cranes is feasible and workable. By adapting the selection mechanism, the guidance path can avoid crossing over existing objects, maintain rigging directions, and maintain low elevations as far as is possible. The suggested path can also be followed with at most 2 DOFs movement under operators’ perception capacities. In consideration of the dynamics of a construction site, an octree structure for partially rebuilding the C-space was integrated to speed up the re-planning process.

To demonstrate and evaluate the developed approach for rigging path guidance, a simulation system, named PathGuider, was implemented with a test scenario similar to that of the pilot test for crane operators. Two demonstrations, showing the real-time path re-planning ability and securing visibility of the proposed algorithms, were performed and presented promising results. A benchmark was developed for measuring computation time of re-planning with respect to the octree level numbers. The results showed that the re-planning process using 500 sampled points to ensure a collision-free guidance could be completed within 200 ms. This shows that real-time rigging path planning for tele-operated cranes is feasible and workable. By adapting the selection mechanism, the guidance path can avoid crossing over existing objects, maintain rigging directions, and maintain low elevations as far as is possible. The suggested path can also be followed with at most 2 DOFs movement under operators’ perception capacities. In consideration of the dynamics of a construction site, an octree structure for partially rebuilding the C-space was integrated to speed up the re-planning process.

To demonstrate and evaluate the developed approach for rigging path guidance, a simulation system, named PathGuider, was implemented with a test scenario similar to that of the pilot test for crane operators. Two demonstrations, showing the real-time path re-planning ability and securing visibility of the proposed algorithms, were performed and presented promising results. A benchmark was developed for measuring computation time of re-planning with respect to the octree level numbers. The results showed that the re-planning process using 500 sampled points to ensure a collision-free guidance could be completed within 200 ms. This shows that real-time rigging path planning for tele-operated cranes is feasible and workable. By adapting the selection mechanism, the guidance path can avoid crossing over existing objects, maintain rigging directions, and maintain low elevations as far as is possible. The suggested path can also be followed with at most 2 DOFs movement under operators’ perception capacities. In consideration of the dynamics of a construction site, an octree structure for partially rebuilding the C-space was integrated to speed up the re-planning process.

6 Acknowledgement

The authors would like to thank the National Science Council of Taiwan for the financial support they have provided to us under the project NSC 98-2221-E-002-185-MY3.

References

Visualization Requirements of Engineers for Risk Assessment of Embankment Dams

Varun Kasireddy\textsuperscript{a}, Semih Ergan\textsuperscript{a} and Burcu Akinci\textsuperscript{a}

\textsuperscript{a} Department of Civil and Env. Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213, USA
Corresponding Author E-mail: varunkasi@cmu.edu

Abstract -
Aging infrastructure in the US has gained quite a bit of attention in the past decade. Being one type of a critical infrastructure, embankment dams in the US require significant investment to upgrade the deteriorated parts. Due to limited budgets, understanding the behavior of structures over time through risk assessment is essential to prioritize dams. During the risk assessment for embankment dams, engineers utilize current and historical data from the design, construction, and operation phases of these structures. The challenge is that during risk assessment, various engineers from different disciplines (e.g., geotechnical, hydraulics) come together and how they would like to visualize the available data sets changes based on the discipline-specific analyses they need to perform. The objective of this research study is to understand the discipline specific visualization needs of engineers from US Army Corps of Engineers (USACE) who are involved in risk assessment of embankment dams when they deal with large set of data accumulated since the inception of dams. The requirements were identified through a three-phased research approach including interviews with engineers who are regularly involved in risk assessment processes, a card game and review of standards and published work on risk assessment of embankment dams. The findings suggest that visualization of the dam layout, components and geometry within 3D settings overlaid with sensor data (which could be queried based on engineers’ discipline specific needs) and data analytics results provide a better flexibility to engineers to understand the risk associated with potential failure modes.

Keywords -
Mining; Built Infrastructure; Human Factors; Embankment Dams; Visualization; 3D modelling; Risk Assessment

1 Introduction

Embankment dams, particularly, the aging ones are prone to failure with progressing time. Various types of failures, including internal erosion, sliding due to loading and overtopping, exist for an embankment dam. Many dams have already received a “poor” rating as per the grade card released by ASCE recently [1]. Most importantly, these dams are an integral part of a prospering economy, and directly concern the lives of a large percentage of population living nearby. To repair and rehabilitate all of those dams are simply not possible due to budget constraints, and hence dams that require immediate remedial actions need to be identified and prioritized. One practical approach for this prioritization is through risk assessment, which includes the assessment of these dams periodically for the level of risk of failure and the magnitude of economic and life causalities associated with such a failure, and act accordingly.

Risk assessment process is an interdisciplinary process and involves engineers of various disciplines like Geotechnical Engineering (GT), Geology (GE), Hydraulic Engineering and Hydrology (H&H), and Structural Engineering (SE). Also, risk assessment activities are typically carried out in different frequencies and granularities. Examples include daily monitoring, which is performed on the daily data collected on the dam to detect changes in readings overtime; periodic inspection (PI), which is conducted every five years in a detailed manner including historical data, and periodic assessment (PA), which is conducted every ten years with interdisciplinary parties. Currently, during these sessions, the multi-disciplinary team of engineers has access to different types of information, such as design, construction and operation information and accesses them through digital or hard copy documents. Collecting the required information and processing/analysing the document based information are resource and time intensive [2].

Unique challenges that engineers face during risk assessment include (a) bringing a spatial context to the
sensed data from piezometers, inclinometers, survey monuments and weirs, (b) understanding the behaviour of dams over time by correlating several parameters about dams (e.g., evaluating pool elevations with respect to piezometer readings, piezometer readings with respect to their station locations, piezometer tip elevations with respect to soil layers etc.). While data collection and processing efforts are preliminary data stages, it is the data visualization stage that plays a vital role in understanding the valuable information concealed inside the data. As data can be represented in different forms, and stored in multiple formats, it is important to understand which form is the most useful for the end users of the data, i.e., dam engineers in this case, to aid in the risk assessment process. For this purpose, it is necessary to identify the engineer’s visualization requirements.

Engineers develop various artefacts to keep track of the correlations in mind, such as correlation plots, cross section layouts, piezometer locations on a plan view, lithology plans showing bore-hole locations and properties. Current tools and artefacts used by engineers do not enable them to perceive the data and correlations between them through views that can be generated flexibly based on how the engineers would like to look at the data. The artefacts are static and are not always capable of correlating the parameters at a glance [3]. Likewise, our initial interactions with engineers during a risk assessment session showed that the visualization requirements and corresponding modes of visualization vary as per the background of an engineer. For example, geotechnical engineers require to look at how different rock types are spatially distributed over the dam site and laboratory rock tests reflected as such. On the other hand, geologists intend to look at the same data in a layer-wise manner, and prefer to be able to turn on/off different rock-type layers within the same 2D/3D visualization window. Consequently, this mandates the requirement of a flexible visualization paradigm to ensure effective and efficient perception and comprehension of the data.

Within the context of this paper, the authors provide the details of the findings on identification of discipline specific visualization requirements of engineers needed during risk assessment of embankment dams. The authors describe the related background research (Section 2), detail the three-pronged research methodology adopted in this study (Section 3) and give details of the findings (Section 4). The paper concludes with recommendations and possible future directions.

2 Background Research

Several studies in the literature have been done in relation to usage of various forms of visualization to aid the dam risk assessment process. Harnessing different modes of visualization, i.e. 3D and 4D, to present different types of information from disparate sources enhances the ability to absorb the content, as well as the ease of its access, when required [4].

In relation to 3D and 4D visualization, researchers typically represented dam body and its features in 3D, while some features which varied over time were simulated in a 4D environment. Studies focused on highlighting different parts of the dams to be visualized. Such studies include visualization of surface and groundwater features for hydraulic erosion for various types of dams and levees [5], and visualization of geometric surfaces, lithological and hydraulic level properties [6]. Such information visualization has also been performed over web-based platforms to facilitate quick feedback and information dissemination during multi-disciplinary meetings with participants from disparate locations [7].

Besides these, some researchers used GIS paradigm to model, simulate and visualize dam-specific features, for example, Serre et al. [8] modelled levee performance to help in planning inspections, maintenance and repair work; and Qi and Altinakar [9] for floods. Likewise, GIS and geo-databases were integrated to represent rich contextual information, with Shumilov & Breuing[10] specifically working on the integration of GIS and geo physical 3D modeling tools. Apart from 2D-3D visualization of behavior of dam and site features and characteristics over time, engineers also prefer to easily access past construction photos and reports, in order to understand what features of the dam have changed over different phases of its life cycle.

One of the main differences between the previous studies on 2D-3D-4D visualization applied to dam risk assessment and the study presented here is the way visualization is utilized. The previous studies did not focus on developing a holistic understanding of the ways engineers would like to look at the data given their engineering discipline and developing visual forms to enable those. The study presented in this paper focuses on characterization of such visualization needs to better serve engineers during their decision making processes while assessing risk levels of dams.

3 Research Objective and Methodology

The main objective of this study is to understand the requirements of the engineers with regard to their
preferences in visualizing information while performing embankment dam risk assessment activities for a dominant failure mode. This paper provides findings in relation to internal erosion. Internal erosion, in particular, is complex to understand, and can even be triggered by normal day-to-day operations without a high intensity event like frequent high pool elevations. Internal erosion is also a major cause of failure of embankment dams [11], and hence was the reason to focus on internal erosion in this study. Previous literature on requirements elicitation (Wiegers [12]; Gould & Lewis [13]) suggests that the most productive approach to accumulate and analyze requirements for a specific task is to determine use cases and build prototypes with varying levels of details while utilizing user feedback at each stage of the prototype development process. The research team used a similar approach that incorporated a multi-phased requirement elicitation and case analysis to interact with engineers and document their visualization requirements during risk assessment process.

A three phased approach is used in this study to identify and validate the visualization requirements of engineers drawn from different disciplines. These phases are described in details in subsequent sections:

### 3.1 Phase 1: Requirements Elicitation through Systems Investigation and Interviews

In this phase, the research team conducted face-to-face unstructured interviews with engineers involved in risk assessment processes, and investigated the information systems used by the engineers to understand different views/figures currently generated with these systems. The larger goal of this phase is to compile a preliminary list of visualization requirements which would constitute an initial list of use-cases for a more-structured elicitation and validation of requirements. 15 engineers from different disciplines, as detailed in Table 1, participated in this study. Majority of these engineers were experienced engineers who have been involved in risk assessment processes for several embankment dams.

Several systems are currently used by engineers to store, access, and visualize the collected sensor data. They gave integrated plotting, reporting and GIS-linking capabilities, based on predetermined templates. During the study, these systems have been evaluated as part of the preliminary analysis so that the preliminary list of visualization requirements could be enumerated and that they could be communicated and discussed during the Phase I interviews with the engineers.

The primary focus of the interviews during Phase I was to capture discipline specific visualization requirements without delving too much into the process of extracting only those requirements which are relevant to the particular failure mode being assessed in this study.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Number of participants</th>
<th>Years of Experience</th>
<th>Discipline(s)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>7</td>
<td>10-32</td>
<td>H&amp;H, GT, SE, GE, CE</td>
</tr>
<tr>
<td>II</td>
<td>5</td>
<td>13-37</td>
<td>GT, H&amp;H, GE, CNSTR, SE</td>
</tr>
<tr>
<td>III</td>
<td>3</td>
<td>4-16</td>
<td>H&amp;H, CE</td>
</tr>
</tbody>
</table>

*H&H: Hydraulic Engineering and Hydrology; SE: Structural Engineering; GE: Geology; GT: Geotechnical Engineering; CE: Civil Engineering; CNSTR: Construction Engineering

These preliminary findings were also useful to determine how engineers would like to visualize different dam features, and also to remove the ambiguity, if any, in the meaning of the terms from the perspectives of each engineering discipline.

### 3.2 Phase 2: Requirements Elicitation through a Card Game, Examination of Standards/Guidelines and Case Documentation

Unlike the previous phase, wherein the requirements were collected in a generic sense, in this phase, the focus was particularly on assessment of internal erosion problems. In this regard, a card game was designed to expand the initial findings of the Phase I. Additionally, the team investigated standards and publications related to internal erosion assessment; and other risk assessment documentation available for three selected dams. The main strategy here is to corroborate the visualization requirements based on the analysis of multiple sources of information, i.e., through triangulation. Triangulation ensures the generality of the findings.

To approach capturing the discipline specific visualization requirements of engineers, a card game was designed to be used with accompanying scenarios. The card-game included pile-of cards, and each card represented an information item that an engineer might be interested in knowing to understand the behaviour of a dam. Piles included several categories such as information about instrumentation, embankment features, historic reports, field tests, and drawings. Among each pile of cards, blank note cards were placed to accommodate the situation in which a participant asked for information that was not already represented in the pile of cards. Given a scenario, engineers requested information to assess the risk level for internal erosion and define how they would like to visualize that...
information.

As part of the triangulation efforts, the research team examined various engineering guidelines/manuals like engineering manuals (EM), engineering regulations (ER). In addition, for three selected embankment dams, the research team examined the plots and visualization approaches used to depict or highlight identified facts about the dams in previous risk assessment reports.

3.3 Phase 3: Requirements Validation through Prototype Development and Face Validation

The main tasks carried out in this phase to validate the requirements identified in the above two phases included development of a functional prototype integrating all visualization requirements, and taking user feedback regularly through showing each identified and implemented view. The prototype was developed using an object-oriented language and enabling renderings of rich 2D-3D graphics. With this prototype, it was possible to do face validation with the users in terms of pinpointing any discrepancies between what the research team interpreted vs what the users actually asked for.

4 Research Findings

The findings are presented in terms of what has been identified as visualization requirements through the requirements elicitation approaches and then how the findings were implemented in the functional prototype.

4.1 Identified Visualization Requirements

The research team identified a total of 42 unique visualization requirements based on the research methodology outlined in the previous section. They have been tabulated in Table 2 based on the engineering disciplines and the overarching categories of visualization. Observations from Table 2 reveal that some of these discipline-specific requirements overlap with those of other disciplines, and the details of the same are discussed in the subsequent paragraphs. For the convenience of the reader, the authors have highlighted the overlapping requirements across different disciplines in bold font in Table 2. As a whole, Table 2 gives an idea of how visualization requirements vary with engineers from different backgrounds for the case of internal erosion risk assessment.

The distribution of the findings with respect to the engineering disciplines is not equally distributed. We can clearly understand from Table 2 that 78% of the total unique visualization requirements were provided by geotechnical engineers and 35% of the requirements were provided by geologists, with overlapping requirements between groups. They were followed by Hydraulic engineers/Hydrologists (H&H group), who contributed to 16% of the total requirements. Similar in scale to the H&H group, structural engineers contributed only 14% of the total. The reason for having a wider set of requirements stated by geotechnical engineers and geologists is due to the scope of the problem being internal erosion, which falls more to the domain of geotechnical engineers. Also, since the scope of this study was limited to embankment dams in which structural features are minimal in comparison to other dam types such as the concrete dams, having a less number of requirements defined by structural engineers is expected.

When Table 2 is analysed in terms of commonalities of visualization requirements based on engineering disciplines, it was observed that only 14% of the total 42 requirements such as geometrical information about dams; pre-existing structures; and reservoir pool and tail water elevations; were of interest to the engineers to look at collectively from all disciplines. There was a consensus among engineers regardless of their disciplines regarding certain visualization requirements. For example, all engineers preferred to have site plans for pre-existing features, which are important to know about for internal erosion assessment, around the dam site in 2D views. Similarly, the opinion was unanimous as far as the representation of dam geometry and information related to it in a 3D view. They also would like to have additional tools to be able to export different cross sections and plan views, and to turn on and off different layers (e.g., instrumentation, zoning, soil layers, pre-existing site plan, etc.). All disciplines also underscored the importance of visualizing the zoning within the dam (e.g., cross-hatching, colour, etc.) as well as the reservoir and tail water information. Here, all the engineers prefer to access the raw reservoir pool and tail water elevations and look at the related plots in a single view. In the same context, engineers would also like to be able to visualize water levels and flows over time (i.e., a 4D simulation of the water level on 3D dam geometry). In addition to these, the research team studied and identified that some of the requirements i.e., instrumentation information and readings provided within 3D settings and geotechnical and geologic information provided in plan views were common to at least three engineering disciplines.

Though there are overlaps in the visualization requirements among engineering disciplines, the percentage of overlap varies with the discipline specific visualization requirements. For instance, from Table 2, it is evident that most of the 3D visualization requirements of geotechnical engineers overlapped with the requirements of the engineers from other disciplines.
The overlapped features include turning on/off various layers of the information on the 3D model as well as visualization of instrumentation information (e.g., location, tip elevation etc,) and instrumentation readings within the 3D settings. In contrast to that, the requirements of geologists do not have many overlaps with engineers from other disciplines.

Specific to the H&H group, hydraulic engineers were interested in the features enabling the visualization of regional rainfall inundation map, Possible Maximum Flood regional map, Hydro Meteorological Report-51 i.e. a probable maximum precipitation document, and the 3D view of the dam geometry. Furthermore, they also expressed interest in accessing tail water, pool elevation and reservoir inflow characteristics in a tabular form. Besides that, they also wanted to look at the hydrologic loading data for coincident pools for seismic PMFs, hydrologic loading data for flood events, inflow-volume-duration-frequency curve [1-7] day computed probability, pool-frequency, and pool-duration curves.

Incidentally, the visualization requirements of the structural engineers have a good overlap with those of the H&H group as far as the H&H tabular data is concerned. They have additional requirements for 3D visualization of the dam instrumentation and the site plan. On the other hand, the interests of civil engineers lie in the availability of instrumentation data - in the form of tables, and 3D geometry of the dam.

### 4.2 Implementation of Visualization Requirements in the Functional Prototype

The prototype was developed in an iterative and a participative manner, in which the opinion and feedback of the end users regarding the functionalities incorporated in the prototype, visual requirements implemented, and usability aspects, were regularly taken to customize existing features and also add new features if necessary. Initially, a view for accessing and displaying instrumentation meta-data was implemented along with a 2D data viewer for static 2D plots (i.e., requirements 8, 10, 16 and 17 in Table 2). A 3D model viewer was built in to the model and integrated with several required data to display contextual information about dam features and instrumentation data were added (i.e., requirements 25, 26, 29, 30, 33 in Table 2). In the next phase, querying capabilities for instrumentation data were incorporated (i.e., requirement 34 in Table 2). 2D data viewer was augmented with a dynamic time slider to visualize variation of readings over time (i.e. requirement 9 in Table 2), based on the feedback of engineers. In the following phases, views for bore-hole test results (i.e., requirements 19-20 and 36-37 in Table 2), document/photo access panels and image display capabilities were added to the prototype (i.e., requirements 1-7 and 11-15 in Table 2).

Discussing all the features implemented in the prototype is out of scope of this publication; simply due to their sheer number and the space restrictions. However, some of them are detailed below:

#### Implementation for visualization of piezometer meta-data and time-series readings

In relation to instrumentation data visualization, piezometers were the commonly referred instrument type to know about for internal erosion assessment. Engineers wanted to select different piezometric zones of influence within the 3D dam body and select the desired piezometers within them to examine their meta information. Meta-data and additional information to be specified for each piezometer included tabular and plotted piezometer data over time with respect to pool elevations, instrument location and tip elevation with respect to soil layers and stations in the dam, as well as piezometer influence zone in 3D phreatic surface (i.e., requirements 9, 10, 33 in Table 2). In addition to this, engineers would like to compare different piezometers using the querying functionality and plotting their readings over time along with the pool elevation variation using the time slider; and in the form of time series data were implemented– as shown in Figure 2.

![Figure 1. A snapshot showing that different instruments can be selected from 3D model interface](image)

#### Implementation for visualization of testing data such as boring logs and rock tests

Within testing data, “boring logs” is one of the frequently used words in the interviews with most of the geotechnical engineers and geologists (i.e., requirements 20 and 37 in Table 2). Important features implemented, concerning boring log information, are meta-data display of any selected bore hole inside a data panel; and display of different soil strata within each boring log. As engineers also showed tremendous interest in the ability to query for different bore holes based on a certain criteria, advanced query docking frame has been implemented for customized comparison, and here, users are able to put different bore-holes side-by-side and view...
their strata properties, and meta information and other related information (as shown in Figure 3).

Figure 2. A snapshot showing that a time series of selected piezometers

Figure 3. A snapshot showing that different bore holes can be compared (we can see different strata layers of each bore hole in this figure)

Figure 4. A snapshot is showing various documents and photos can be stored and accessed from the integrated prototype

**Implementation for visualization of documentation and construction history photos**

Most of the dams have been constructed many years ago and they have lot of paper documentation concerning its construction history, repairs, site instrumentation, standards etc. With time, it becomes very difficult to retrieve particular old documents, say, if needed for a risk assessment process, or even for the perusal of the project engineers. Hence, engineers wanted an internal document indexing system within the prototype to drag and drop digitized files and photos and to be able to retrieve these indexed files quickly within the same interface, whenever needed. The implementation of this feature is shown in Figure 4, wherein a user selected a photo from the file index panel, and it is being displayed in the adjacent docking panel.

Figure 5. A snapshot showing 3D model panel

5 Conclusions

Visualization empowers engineers to conveniently visualize, integrate and accurately interpret the data from disparate sources. For internal erosion risk assessment in embankment dams, engineers from several disciplines require dam information to be viewed from different perspectives. This study provides the findings of visualization requirements of engineers involved in risk assessment processes while looking at historical dam information.

While the engineers would like to be able to use the current methodologies they are using to visualize static data related to embankment dams, they desire for an advanced 3D visualization paradigm that allows the end users to at least import different cross sections and plan views; turn on and off different information layers concerning instrumentation and other site plans; and simultaneous comparison through querying and visualization of multiple boring logs, piezometers, and monuments.

The findings from this study suggest that engineers would like to visualize the dam layout, components and geometry within 3D settings overlaid with sensor data, and querying capabilities in order to get a better flexibility to understand the risk associated with potential failure modes. Armed with this flexibility, they can be more effective and efficient during risk assessment sessions, and can contribute to better dam maintenance.
decisions.

Future work can include putting efforts to quantify the value of using such visualization tools with engineers through scenarios from a specific dam for assessment of internal erosion. Among the instruments mainly used in the data collection tasks at the dam location, the research team focused mainly on the piezometers in the risk assessment process for the current study. In the future, other available instrumentation and their readings could be investigated to understand internal erosion risk and risk due to other failure modes in a holistic manner.

Acknowledgements

The research presented in this paper is supported by USACE grant. The authors would like to acknowledge the support of Chris Kelly, Meghann Wygonik and other engineers from USACE who participated in this research at various stages. Also, we developed the prototype using the research version of toolbox provided by IFC Tools Project [14].

References


<table>
<thead>
<tr>
<th>Discipline</th>
<th>2D visualization mode</th>
<th>3D visualization mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geotechnical</td>
<td>1. Site Plans for pre-existing structures</td>
<td>25. Dam geometry</td>
</tr>
<tr>
<td></td>
<td>2. Drawings showing the geometry to compare the as-constructed drawings with dam body</td>
<td>26. 3D model with renderings of the pre-existing structures</td>
</tr>
<tr>
<td></td>
<td>3. Drawings with zoning, thickness and geometrical properties within the dam body</td>
<td>27. Animation for reservoir and tailwater</td>
</tr>
<tr>
<td></td>
<td>4. Cross-section: showing the geometry of the spillway</td>
<td>28. 3D Immersive foundation model</td>
</tr>
<tr>
<td></td>
<td>5. Contour diagrams of the channel and topography</td>
<td>29. 3D model where can export cross sections or play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td>6. Current photos and change of vegetation over time</td>
<td>30. Full rock testing report shown spatially by rock type</td>
</tr>
<tr>
<td></td>
<td>7. Site Plans</td>
<td>31. Simulations of different scenarios as consequence information</td>
</tr>
<tr>
<td></td>
<td>8. Tabular data of reservoir pool and tail water elevations</td>
<td>32. 3D view of instrumentation and readings</td>
</tr>
<tr>
<td></td>
<td>9. Top 10 events performance and ratings for the reservoir and tail water</td>
<td>33. Piezometer in 3D planar surface showing the influence zone for the user defined period</td>
</tr>
<tr>
<td></td>
<td>10. Time history plots of reservoir level per cross section</td>
<td>34. Comparison of the piezometer cross sections according to their elevation and material type</td>
</tr>
<tr>
<td></td>
<td>11. Geologic Maps</td>
<td>35. Profile of dam with lines to show how it has settled</td>
</tr>
<tr>
<td></td>
<td>12. Geo technical and geologic drawings</td>
<td>36. Visualization of total site boring plan</td>
</tr>
<tr>
<td></td>
<td>13. All drawings related to spillways</td>
<td>37. 3D visualization of side by side boring logs adjusted for elevation enables engineer to click and obtain the boring log report</td>
</tr>
<tr>
<td></td>
<td>14. Construction alstoric data reports</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15. GIS map</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16. Instrumentation tabular data and plots</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17. Tabular data and plots showing the deflection in the instrumentation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18. Full laboratory soil testing report including values</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19. Soil test value portrayed by soil horizon and depth which has its own hatch style</td>
<td></td>
</tr>
<tr>
<td>Geologist</td>
<td>• Site Plans for pre-existing structures</td>
<td>• Dam geometry</td>
</tr>
<tr>
<td></td>
<td>• Tabular data of reservoir pool and tail water elevations</td>
<td>38. 3D model with pre-existing layer which enables to cut the cross section and super impose dam</td>
</tr>
<tr>
<td></td>
<td>• Geo technical and geologic drawings</td>
<td>39. Water level max/min and piezometer readings per cross-section</td>
</tr>
<tr>
<td></td>
<td>• Construction alstoric data photos</td>
<td>40. 3D Rose diagram showing the fracture pattern</td>
</tr>
<tr>
<td></td>
<td>• GIS map</td>
<td>41. 3D model where can export cross sections or play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td>• Instrumentation tabular data and plots</td>
<td>42. Ability to move between the rock testing reports and images and play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td>20. Boring logs sections with explanatory drilling apertures</td>
<td>• Simulations of different scenarios as consequence information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D view of instrumentation and readings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D visualization of side by side boring logs adjusted for elevation enables engineer to click and obtain the boring log report</td>
</tr>
<tr>
<td>Hydrology and Hydraulics</td>
<td>• Site Plans for pre-existing structures</td>
<td>• Dam geometry</td>
</tr>
<tr>
<td></td>
<td>21. Breach parameters of the embankment</td>
<td>38. 3D model with pre-existing layer which enables to cut the cross section and super impose dam</td>
</tr>
<tr>
<td></td>
<td>• Tabular data of reservoir pool and tail water elevations</td>
<td>39. Water level max/min and piezometer readings per cross-section</td>
</tr>
<tr>
<td></td>
<td>22. Tabular data of reservoir inflow characteristics and volume-duration-frequency</td>
<td>40. 3D Rose diagram showing the fracture pattern</td>
</tr>
<tr>
<td></td>
<td>curves</td>
<td>41. 3D model where can export cross sections or play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td>23. Hydrologic loading data for coincident pools, for seismic PAF’s and flood events</td>
<td>42. Ability to move between the rock testing reports and images and play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td>24. Regional rainfall, inundation, PAF regional and HMR-51 maps</td>
<td>• Simulations of different scenarios as consequence information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D view of instrumentation and readings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D visualization of side by side boring logs adjusted for elevation enables engineer to click and obtain the boring log report</td>
</tr>
<tr>
<td>Structural</td>
<td>• Site Plans for pre-existing structures</td>
<td>• Dam geometry</td>
</tr>
<tr>
<td></td>
<td>• Tabular data of reservoir pool and tail water elevations</td>
<td>38. 3D model with pre-existing layer which enables to cut the cross section and super impose dam</td>
</tr>
<tr>
<td></td>
<td>• Hydrologic loading data for coincident pools, for seismic PAF’s and flood events</td>
<td>39. Water level max/min and piezometer readings per cross-section</td>
</tr>
<tr>
<td></td>
<td>• Geo technical and geologic drawings</td>
<td>40. 3D Rose diagram showing the fracture pattern</td>
</tr>
<tr>
<td></td>
<td></td>
<td>41. 3D model where can export cross sections or play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>42. Ability to move between the rock testing reports and images and play views with layers turned on or off</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Simulations of different scenarios as consequence information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D view of instrumentation and readings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3D visualization of side by side boring logs adjusted for elevation enables engineer to click and obtain the boring log report</td>
</tr>
</tbody>
</table>

Note:
1) Only the requirements which are repeating, i.e. overlapping over two or more disciplines' requirements are marked in bold text.
2) The first instance of occurrence of each requirement is numbered, and the repeating instances are listed in bullet format.
Ineffective Rock Breaking and its Impacts on Pick Failures

Y. Sun and X.S. Li

CSIRO Earth Science and Resource Engineering, PO Box 883, Kenmore QLD 4069 Australia
E-mail: yong.sun@csiro.au, xing.li@csiro.au

Abstract -
Picks are critical components in excavation machines used in civil and mining industries. Any failures of picks in an excavator directly affect the reliability and productivity of the machine. Pick tip wear, pick tip crack, pick body wear and pick body bending are common failure modes of picks in production. Among various failure causes, ineffective rock breaking during a cutting process is a major cause which can result in most of these failure modes. In this paper, the mechanism of ineffective rock breaking is investigated and the influences of ineffective rock breaking on pick failures are studied. It shows that apart from pick tip geometry, pick installation angles on a drum and rock properties, pick lacing space is a major factor that affects the effectiveness of rock breaking. More importantly, it is a more controllable parameter in the drum design stage. The results of this research can be used to reduce pick failure probability, and automate optimal drum design.

Keywords -
Pick failures; Rock cutting; Reliability, Drum design; Rock breaking; Cutting interaction

1 Introduction

Excavation machines equipped with picks such as continuous miners, longwall shearsers and roadheaders have been widely used to cut rocks in civil and mining industries. As critical components in these machines, any failures of picks during production is of major concern because pick failures can pose significant negative impacts on machine operators and owners. On the one hand, picks are not cheap. Consuming a large number of picks can be a big cost to the operators or owners. In fact, pick usage per ton of coal is one of KPIs (key performance indicators) in coal mining industry; on the other hand, pick failures will normally cause machine shutdown and production interruption, often resulting in a significant production loss. Therefore, it is important to understand the failure mechanism of picks and reduce pick failures.

However, research on pick failures is so far inadequate. While most existing researches about picks either focused on analyzing and predicting these forces [1-8], some studies have also been carried out on determination of attack angle and tilt angle of a cutting pick [9], cut interactions [2], new tip materials [10-12], and frictional ignition risk [13].

Regarding pick failures, Sun and Li developed a model to analyze the impact of the property variation of pick tip material on the failure probability of the tips [1]. Li and his colleagues investigated and compared the wear characteristics of cutting tips made of TSDC (thermally stable diamond composite) and WC (tungsten carbide) in abrasive cutting operation [12, 14]. McNider et al suggested that pick life could be prolonged through using capped tips [13]. A significant gap in existing pick failure research is that the interaction between rock breaking and pick failures has not been investigated yet, although rock breaking or cutting patterns have attracted the attention of researchers for a long time [15, 16].

During rock cutting, a pick indents into rock body in a certain angle and causes a certain amount of rock to break out and then be removed from the rock body as rock chips. Given that, for any excavation machine, a number of picks are installed on a drum and work together to cut rocks, investigation of rock breaking of any individual pick should consider the collective influences of its adjacent picks. Hurt and MacAndrew [16] pointed out that if the ratio of line spacing to depth of cut into a pre-existing groove is too large, over-deepening (also called groove-deepening) may occur. While most existing research on rock breaking patterns focused on cutting efficiency and forces, this study focuses on its impacts on pick failures.

2 Pick Failure Modes and Causes

Picks have various failure modes. Some examples are demonstrated in Figure 1. Typical pick failure modes include pick body bending (Figure 1-a), pick body wear (Figure 1-b), pick tip fracture (Figure 1-c), pick tip wear (Figure 1-d) and a combination of these (Figures 1-e and 1-f). Different failure modes could be...
caused by different mechanisms. For example, tip wear could be caused by the high frictional force during cutting hard and abrasive rock and insufficient wear resistance of the tip material; pick body wear could be caused by the rubbing between pick body and hard rock surface, flying rock chips against body and body’s contact with corrosive materials; tip cracks could be caused by excessive impact force exerting on the tip, thermal fatigue and/or defects in tip material; pick body bending could be caused by excessive bending force due to inappropriate attack angle, inappropriate tilt angle, inappropriate lacing space and/or incorrect pick body design. However, there is a common cause which can result in most of these failure modes. This common cause is ineffective rock breaking during a cutting process.

Figure 1. Examples of pick failure modes
3 Phenomenon, Causes and Effects of Ineffective Rock Breaking

Effective rock breaking and ineffective rock breaking are relative concepts. Previously, they are mainly differentiated in terms of cutting efficiency [16]. However, this criterion is still vague and hard to apply in practice. To address this issue, here they are differentiated in terms of the cutting perimeters in cutting patterns. A cutting perimeter is an envelope of all picks tips on the same sequence cutting into rock in one cutting cycle (corresponding to that a drum rotates one revolution, refer to Figures 2 and 3). This means if a drum has n sequences, there are n perimeters in one cutting cycle. If the whole rock before the cutting perimeter of the previous cutting sequence has been effectively removed by the cuts in the current cutting sequence, the rock breaking of the current cutting sequence is said effective. If any rock areas before the cutting perimeter of the previous cutting sequence are still left after the cutting of the current cutting sequence, the rock breaking of the current cutting sequence is not effective, or more specifically, the cuts which are supposed to remove these rock areas in the current cutting sequence have an ineffective rock breaking. Figures 2 and 3 illustrate an example of effective rock breaking and an example of ineffective rock breaking respectively. In both examples, the drum is assumed to have one sequence only.

![Figure 2. An example of effective rock breakout](image-url)
In Figures 2 and 3, $S_{12}$ and $S_{23}$ are the line spacing between Picks 1 and 2 as well as between Picks 2 and 3 (mm) respectively, $\gamma_1$, $\gamma_2$ and $\gamma_3$ are the breakout angles of Picks 1, 2 and 3 (deg) respectively. All the parameters in Figure 3 are the same as those in Figure 2 except that line spacing between two adjacent picks is much larger than that in Figure 2.

From Figure 2, it can be seen that the whole rock before the cutting perimeter of the first cutting cycle has been effectively removed by the cuts in the second cutting cycle. Hence, the rock breaking of the cuts in the second cutting cycle is effective. On the other hand, from Figure 3, it can be found that if the rock could break out along the broken lines, then the cuts in all cutting cycles would also make effective rock breaking. However, because the spacing is too large, the cuts are actually not able to breakout so much rock. The breakout angle of each cut becomes much smaller, leading to the actual breakout lines as shown by the continuous lines. As a result, after the second cycle, some rock areas before the cutting perimeter of the first cutting cycle are still left, i.e., the cuts given by Picks 1 to 3 in the second cutting cycle all make ineffective rock breaking. Similarly, rock breaking of Picks 1 to 3 in the third cutting cycles is also ineffective.

Linear rock cutting experiments in CSIRO’s Rock Cutting Laboratory have validated the above theory. Some experimental results are shown in Figure 4. Technical details of the experiments will be published in due course. Figure 4-a shows three cutting lines (labelled 1, 2 and 3 respectively) which were made to simulate an initial rock cutting condition by three adjacent picks in a drum. Cut 3 was 10mm deeper than cuts 1 and 2. In Figure 4-b, cut 4 was cut 20mm deeper on the groove made by cut 1 and in Figure 4-c, cut 6 was cut 20mm deeper on the groove made by cut 3. It can be seen that the breakout angle of cut 4 was much smaller than that of cut 1 and the breakout angle of cut 6 was much smaller than that of cut 2. After cut 6, most original rock surface left by cuts 1 to 3 still remained,
indicating that the rock breaking of either cut 4 or cut 6 was ineffective and groove-deepening has occurred.

Ineffective rock breaking will not only reduce cutting efficiency, but also increase pick failure rate.

1. Narrowed breakout angle will cause rubbing between pick body and rock surface (Figure 5), resulting in severe pick body wear. As evidence, obvious rubbing marks have been observed inside the cutting groove and the pick body both after cut 4 (Figure 5).

2. Groove-deepening will exert larger frictional force on pick tips and cause them to wear and/or fracture. From Figure 5, it can be found that some white powder was left on the pick tips. This powder was produced because the tips ground the rock during the cutting process.

3. When ineffective rock breaking happens, the cuts in some cutting cycles or sequences cannot remove the rock they aim to remove. As a consequence, some cuts after these cycles or sequences have to remove more rock than they are designed to. This means much larger forces will be applied to the picks that make these cuts and may cause the cracking of their tips and/or the bending of their bodies.

4. Ineffective rock breaking will significantly increase the temperature between pick cutting elements and rock face because of the increased frictional force. It will also generate large sparks due to the rubbing of pick steel body against rock surface as shown in Figure 6. High temperature will often weaken the mechanical strength of cutting elements and pick bodies.

4 Discussions

Ineffective rock breaking is very harmful to pick reliability and thus should be avoided. Comparing Figure 2 with Figure 3, it can be seen that ineffective rock breaking also means that the picks in a drum cannot form effective cut interactions. When designing a drum, one has to take into account rock breaking property to ensure effective rock breaking for optimal pick usage and production efficiency.

In drum design, cutting patterns are often used to illustrate cutting actions [16] and optimize pick lacing arrangements. However, in the current practice, cutting patterns are often drawn based on an assumption that breakout angle is unchanged and breakout line can extend indefinitely. Figure 7 shows a part of a cutting pattern developed based on this assumption. Nevertheless, from Section 3, it is known that this assumption is not always valid. When some parameters such as advance speed, rotational speed and line spacing
change, ineffective rock breaking could happen. In this case, breakout angles and breakout direction no longer remain the same. As a result, the cutting pattern based on the currently used assumption becomes invalid. For example, in Figure 7, the cutting lines made by picks 33 and 38 are very long. This may not reflect the reality. The actual cutting lines made by these two picks are very likely to be different from what are shown in Figure 7. If this is the case, failure assessment results of picks 33 and 38, and other related picks could also be different from what are estimated based on this pattern. Cutting pattern can be used to investigate cutting actions such as interactions between cuts only when it can reflect both effective rock breaking and ineffective rock breaking correctly.
A lot of factors including pick geometry, cutting tip material, rock property, drum design and operational parameters can affect rock breaking effectiveness. In the current, the most effective approach to identifying ineffective rock breaking is still laboratory rock cutting tests. It is desirable to develop theoretical models for determining rock breaking pattern because rock cutting tests are costly and time-consuming.

5 Conclusions

In rock cutting, if the line spacing between picks on a drum is too large, cuts of picks cannot interact with each other effectively. As a result, ineffective rock breaking will happen. In this case, pick failure probability can significantly increase due to narrowed breakout angle and increased work loading. Ineffective rock breaking can cause various pick failure modes including pick body wear, pick body bending, pick tip wear and pick tip fracture. Consuming a large number of picks can be a big cost to the operators or owners. In addition, higher pick failure rate usually means lower machine availability and productivity. Therefore, ineffective rock breaking has to be avoided in production. To achieve this, the characteristics of rock cutting with given picks should be well examined and considered in the drum design and pick failure assessment. In the future, the authors will study how to automate cutting pattern design to reflect effective and ineffective rock breaking correctly.

Acknowledgement

The authors wish to thank Peter Clark and Craig Harbers of CSIRO Earth Science and Resource Engineering for their technical supports in the rock cutting tests.

References

AUTOMATION, CONSTRUCTION AND ENVIRONMENT

ROBOTICS AND MECHATRONICS
Control of Hovering Altitude of a Quadrotor with Shifted Centre of Gravity for Inspection of High-rise Structures

Alexej Bulgakov a, Sergey Emelianov a, Thomas Bock b and Daher Sayfeddine c

a South West University, Kursk, Russia
b Thomas Bock of TU Munich, Munich, Germany
c South Russian State Polytechnic University, Novocherkassk, Russia
E-mail: a.bulgakow@gmx.de, ems@mail.ru, Thomas.Bock@bri.arch.tu-muenchen.de, daher@live.ru

Abstract-
Unmanned aerial vehicles are specialized robots. Recently they were used in many civil applications such as patrolling, firefighting, rescuing tasks and being the shadow cinematographer in many movie-making companies. The new developments in material technology led to new sophisticated, miniature range of sensors and actuators, which are able to substitute gyroscopes and accelerometers. In light of this, new miniature aerial vehicles appeared in the market. Micro-UAV (MUAV) is able to infiltrate sites, where the bigger aerial robots could not. It is more maneuverable and recently is becoming able to fly to new altitudes. This allow the MUAV to be used for site scanning, to perform or assist in photogrammetry procedures in light to create 3D building models and maps. The future of the MUAV seems to be existing: In relation to medical field, the aerial robot is being used to perform radiation leakage test that, is done manually in most of the hospitals. It is also planned to be used to deliver lightweight goods within the city.

The versatility and adaptability of the miniature aerial robot creates new requirements for stability, maneuverability and speed. In this paper, we offer to study the quadrotor while inspecting high-rise structures using simulations. As the quadrotor is vertical take-off and landing vehicle, it is very important to analyze the effect of the altitude on the stability of in order to create control algorithm able to maintain the necessary hovering position of the quadrotor. A better stabilization leads to better scanning results and high quality captured images.

Keywords-
Inspecting structures, Quadrotor, Optimized Control System, Simulation.

1 Background

The success of implementation of unmanned aerial vehicles to perform military tasks has made an opening in civil markets for such specialized robots. While we cannot compare the evolution of UAVs with piloted planes, were aircrafts flew “alone” in the sky, the UAV designers are challenged to achieve better stability, offer sophisticated navigation options and honor several safety terms. These requirements have created a need to develop new control algorithm, recognition methods, path planning and generation solutions. In this regard, many researchers and hobbyists put lots of effort to create the best platform to simulate UAVs, study their aerodynamics, design new controllers and trajectory generators. The quadrotor consists of one of the favorite research platform, as it is reasonably affordable and technically easy to build and maintain.

2 Purpose

Quadrotors are take-off and landing rotorcraft. They are known for their high maneuverability. The quadrotor, like other aircrafts that belong to the rotorcraft class, flies by changing the rotational speed of one or several rotors rigidly fixed to the frame. These rotors are the only flying mechanism and lift force generator for such aerial vehicles. Although being described as a six degree of freedom object, three-dimensional movement of the quadrotor is realized by four flight regimes only: hover, roll, pitch and yaw. The first being the only linear position to be controlled directly, the rest represents the rotational movement of the quadrotor. Recently, many approaches were
adopted to simulate the quadrotor dynamics and control. These approaches consider the nonlinearity of the simulated system [1,2,3,4,10]. Most of the control schemes found refer to ideal position of center of gravity [5,6,7]. Using the ideal center of gravity in real control systems leads sometimes to tremendous errors in positioning. This is due to the following reason: the shifted center of mass generates additional accelerations and velocities that are registered by the inertial sensors meant to collect data on positioning. Figure 1 illustrates possible real position of the center of gravity with reference to the quadrotor axis.

Figure 1. Illustration of shifted position of centre of gravity.

Referring to the aforementioned analysis, we dedicate this paper to highlight the difference in considering an ideal and a real center of gravity in simulated control system. By achieving this, we will build our case study which consists of designing an optimized fuzzy logic controller to track a desired altitude, hereafter, to be implemented for scanning high-rise structures.

The quadrotor belongs to miniature rotorcraft class, which generally weighs more or less 1kg and has an average length of 1m. We will shift the center of gravity position by 0.1%, 1% and 10% from the ideal position and check what influence it has on the performance of the autopilot [4,9,10,11]. The results (table.1) are obtained from simulation of 3D flight (XOYZ- earth axis) with 0.1%, 1% and 10% shifted center of gravity.

<table>
<thead>
<tr>
<th>Shifted Centre of Gravity</th>
<th>X [m]</th>
<th>Y [m]</th>
<th>Z [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>Desired</td>
<td>Obtained</td>
<td>Obtained</td>
</tr>
</tbody>
</table>

To design an optimal control, it is mandatory to consider the nonlinearity of the quadrotor as well as the real position of the center of gravity. Quadrotor dynamics are described using the following equations [5,7,8].

\[
\begin{align*}
\dot{x} &= -c(\psi) s(\theta) c(\phi) + s(\psi) s(\phi) \frac{T}{m} \\
\dot{y} &= -s(\psi) s(\theta) c(\phi) - c(\phi) c(\psi) \frac{T}{m} \\
\ddot{z} &= -c(\theta) c(\phi) \frac{T}{m} + g \\
\dot{\phi} &= \frac{T_\phi}{I_x} \\
\dot{\theta} &= \frac{T_\theta}{I_y} \\
\dot{\psi} &= \frac{T_\psi}{I_z}
\end{align*}
\]

Where, \( s \) is \textit{Sine} function, \( c \) is \textit{Cosine} function, \( \ddot{x}, \ddot{y} \) and \( \ddot{z} \) are the second derivative (acceleration) of the quadrotor position along earth axis OX, OY and OZ respectively and \( \dot{\phi}, \dot{\theta} \) and \( \dot{\psi} \) are the second derivative of the roll, pitch and yaw angles, \( I_x, I_y \) and \( I_z \) are the terms of inertia of the quadrotor while performing a rotational movement, \( m \) is the mass of the quadrotor or term of inertia in linear movement and \( T, T_\phi \) and \( T_\psi \) are the total thrust generated per flight regime and its projections on the roll, pitch and yaw axis.

Inspection of buildings for cracks and damages is a long process, especially after natural disasters. It sometimes takes months to assess damages and compensate property owners. By automating this process using specialized robots, time is cut down tremendously. A swarm of quadrotors equipped with cameras can be used to achieve this quicker than doing it manually. Therefore, this specialized robot should be able to shoulder tasks optimally in positioning and time.

3 Method
It is obvious from the results that by moving the center of gravity by 10%, an error of 122% can occur by referring to one meter as desired position. This is due to the fact the control gains are calculated per different norms and parameters of the quadrotor. Considering this fact, with a different size and mass of quadrotor, the error in positioning can reach new records. Another aspect that can be noticed, it is the stability factor in altitude control, it is clearly seen from figure 2 that by increasing the variation coefficient by 1%, the quadrotor tends to shake while hovering. In light of this, performing patrolling tasks such as scanning high-rise buildings, the quadrotor will be unstable due to windy conditions. Hence, it is necessary to design a control system able to keep the quadrotor in its position and track a desired trajectory.

To control the hovering altitude, we suggest a hybrid method consisting of fuzzy logic controller optimized using particle swarm algorithm. Although in many cases, artificial intelligence algorithms were ruled out from selection for such tracking and control tasks, due to the time consumption, the particle swarm method has pushed the ordinary fuzzy logic controller for new benchmark in stability and speed, as it will be proven in the listed simulation results. The fuzzy rules were obtained analytically by finding the resulting graph of the altitude error $de(t)$ function with its derivative in time $f(e(t))$. The two inputs are represented using five triangular membership functions and the output of the regulator is represented using five triangular membership functions as shown in figure 3a and b (for inputs) and figure 4 (for output) respectively.

<table>
<thead>
<tr>
<th>Error (%)</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>1</td>
<td>1</td>
<td>1.053</td>
<td>1</td>
</tr>
<tr>
<td>1%</td>
<td>1</td>
<td>1.009</td>
<td>1.12</td>
<td>1</td>
</tr>
<tr>
<td>10%</td>
<td>1</td>
<td>1.996</td>
<td>2.218</td>
<td>1.003</td>
</tr>
</tbody>
</table>

Figure 3: a- membership function for input “$e(t)$”.

Figure 3: b- membership function for input “$de(t)$”.

Figure 4: Membership function for fuzzy regulator output.
The linguistic rules for the ordinary PD fuzzy regulator are listed in table (3).

<table>
<thead>
<tr>
<th>de</th>
<th>NB</th>
<th>NS</th>
<th>Z</th>
<th>PS</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>(-1,1)</td>
<td>(-1,-0.5)</td>
<td>(-1,0)</td>
<td>(-1,0.5)</td>
<td>(-1,1)</td>
</tr>
<tr>
<td>NB</td>
<td>(-0.5,-)</td>
<td>(-0.5,0)</td>
<td>0.5,0</td>
<td>0.5,0.5</td>
<td>0.5,1</td>
</tr>
<tr>
<td>NS</td>
<td>(0,-1)</td>
<td>(0,-0.5)</td>
<td>(0,0)</td>
<td>(-0.5,0)</td>
<td>(0,1)</td>
</tr>
<tr>
<td>Z</td>
<td>(0.5,-1)</td>
<td>(0.5,-0.5)</td>
<td>(0.5,0)</td>
<td>(0.5,1)</td>
<td></td>
</tr>
<tr>
<td>PS</td>
<td>(1,-1)</td>
<td>(1,-0.5)</td>
<td>(1,0)</td>
<td>(1,0.5)</td>
<td>(1,1)</td>
</tr>
</tbody>
</table>

Figure 5 represent generated plot of the output surface of a the fuzzy inference system using the first two inputs and the first output.

The particle swarm optimization will be used to tune the gain of the PD fuzzy controller.

The optimization method is a Runge-Kutta solver of differential equations. In particular, it uses six functions to estimate and calculate the fourth and fifth tolerance order. The difference between these solutions is taken as an error. This error estimate is very convenient to design adaptive algorithms, such as the use of fuzzy logic to control the process.

The global best value updating the swarm positions to the target. Equation (7) describes mathematically used PSO model

\[ V_{0i}^{k+1} = wV_{i}^{k} + c_{1}(k+1)(pbest_{i} - S_{i}^{k}) + c_{2}(k+1)(gbest - S_{i}^{k}) \]  

Where, \( V_{0i}^{k+1} \) is next value speed, \( w \)-weight function, \( V_{i}^{k} \)-current moving speed, \( c_{1}, c_{2} \) are weights, \( pbest_{i} \) is the personal best value for the particle, \( S_{i}^{k} \) is the current position of the i-th particle and \( gbest \) is the global best position or the target [11].

4 Results

By implementing the obtained control scheme in altitude loop, the autopilot was able to maintain the quadrotor in hovering position. Figure 3 shows simulation results of the altitude control using Fuzzy PSO method.

The comparison between the obtained results with those found in [12, 13] shows that the control approach in [13] (figure 7) lead to a satisfactory result by holding the quadrotor to hovering altitude of one meter with overshoot equal to 0.6 m for 4 seconds. The time required for full stabilization is about 7 seconds. While the first curve (figure 6) shows that, the hybrid
controller consisting of fuzzy logic regulator optimized using PSO has reduced the necessary time to reach one-meter altitude to 4 seconds and without overshoot. Using the suggested algorithm, the quadrotor is more stable.

Figure 6. Simulation results of altitude control of a quadrotor with shifted centre of gravity using Fuzzy PSO. Horizontal axis- time [s], vertical axis- altitude [m].

Figure 7. Altitude control results as by [13]. Horizontal axis- time [s], vertical axis- altitude [m].

5 Implementation

Building on the aforementioned results, we have changed the control task from hovering on an altitude of one meter, to scanning a sequence of building with different heights. Taking in consideration that the average height of one floor is about 3 meters; the desired altitude signal is designed to simulate 10-storeyed building as minimum and 30 storied as maximum including the safety distance to be honored between the building and the UAV. This altitude is enough to challenge any miniature quadrotor (up to 1-meter length) with drag effect, gravitational forces and peripheral wind force. The drag factor will consist a major competition to lift the quadrotor. The overall wind load on the quadrotor is calculated as follows

$$ F = S \times P \times C_d \times K_g \times G $$

Where, F is the load of the wind on the quadrotor, S is the exposed area, P is the wind pressure, is the drag coefficient, is the exposure coefficient proportional to the elevation of the quadrotor from the ground and G is the gust response factor. Assuming that S, P and are constants, and G will definitely changes in parallel with the altitude variation H. The exposure coefficient ad the gust factor are calculated as follows. The exposure coefficient ad the gust factor are calculated as follows

$$ K_g = \left(\frac{H}{33}\right)^{2/7} \quad G = 0.65 + \frac{0.6}{\left(\frac{H}{33}\right)^{1/7}} $$

Fig. 8. shows how the exposure coefficient and gust factor changes with altitude. Applying the aforementioned condition in the quadrotor model, we obtain the results in figure 9. It clearly shows the suggested Fuzzy PSO controller was able to lead the quadrotor to scan high rise structures with ideal stability (no overshoot registered theoretically) conserving the fact of shifted center of gravity of the UAV and changes in surrounding circumstances such as wind.
6 References


Bulldozer as a Mechatronics System with the Intelligent Control

Alexej Bulgakov a Thomas Bock b Georgy Tokmakov c

a South West State University, Russia
b Technical University Munich, Germany
c South Russian State Polytechnic University, Russia

E-mail: a.bulgakow@gmx.de, thomas.bock@bri.arch.tu-muenchen.de, tokmakov.tun@gmail.com

Abstract-Background. Improvement of quality, decrease of terms and cost of construction are inseparably linked with problems of effective use of bulldozer equipment.

Purpose. The most important problem of control tractions modes of the bulldozer is the fullest use of traction opportunities of the machine at the expense of management of work tool. Automatic maintenance of the maximum traction power or resistance preset value on work tool is complicated by a large number of the random factors operating on the bulldozer. In this regard the system of automatic control has to possess possibility of self-adjustment [1].

Method. In this paper with applying of analytical simulation method and neural network technologies, been decomposed model bulldozers workflow as mechatronic system realized [2,3].

Results & Discussion. For those sub-processes, where is possible analytical modeling based on knowledge on the links between the parameters of the bulldozer, analytical dependences are obtained. Models of these sub-processes included in the overall structure of the simulation model bulldozers workflow and are designed for both individual research bulldozers units using analytical relationships between the parameters of the workflow and simulation bulldozers workflow in general. For another thing are method of identification and modeling bulldozers base workflow based on represented (Fig. 5.).

Keywords- Robotics and mechatronics, Automation and control, Bulldozer, Neural network technologies.

1 Introduction

Bulldozers equipped with modern navigation and information systems are mobile mechatronic objects, and they can be integrated into general process of intellectual construction. The integration will provide optimal efficiency of the construction cycle and will ensure lean production process.

On the basis of bulldozer’s workflow dynamics modeling and analyses described in a variety of works, we have concluded that the models to describe kinematics and dynamics of its working equipment, hydraulic and transmission features tend to be analytical formulas derived from well-known laws of physics and from information on bulldozer’s structure and mechanisms. If some parameters of the workflow are unknown or constantly changing, the models are either statistical tables or empiric dependences summarizing experimental data. The models depict interaction of end-effectors, engines and environment as well as statistic features of bulldozer’s complex units.

Application of regulators based on classical control theory is difficult due to the frequent changes in workflow conditions. Thus, it is necessary to develop adapted control systems to eliminate the difficulties described. The system includes both the bulldozer’s dynamics modeling and bulldozer’s workflow control method to take into consideration the complex non-linear dependencies between workflow parameters and incomplete information on its working conditions changes.

Having reviewed adaptive and intellectual control methods [4, 5], we propose to create an adaptive control system for technological processes to increase efficiency of bulldozer’s control in comparison with traditional control methods.

2 Bulldozer’s Workflow Modelling

The main goals for analytic simulation modeling of bulldozer workflow are:
- Bulldozer simulation as a controlled object to realize bulldozer’s workflow parameters for using them at workflow neural network identification;
- Efficient traction modes parameters definition to be supported by the control system;

Simulation tasks:
- To single out the main sub-systems in bulldozer’s structure and interrelations between the sub-systems;
- To develop analytic and simulation models for workflow elements and to include them into the general structure of the model.

General structure of the workflow model for automated bulldozers is developed (Fig. 1). The structure meets the goals of workflow control. When moving soil by the bulldozer, it is necessary to utilize bulldozer’s traction capacity in full keeping the nominal traction value \( N \); when surfacing, the altitudes of the right and left side of the blade \( y = (y_r, y_l) \) are to correspond the design marks. The key element at the scheme (Figure 1.) shows the choice for the first or the second operational mode.

At developing the models, we use mathematical apparatus of the random processes theory, transfer functions, table interpolation, numerical solution of algebraic equations and ordinary differential equations in the Cauchy form.

Random changes in the coordinates of untreated soil surface \( f \), as well as normalized fluctuations in the resistance forces on the working organ \( P_f \), caused by the heterogeneity of the soil are highlighted among the disturbing effects on the working organ of the bulldozer from soil conditions. Disturbance \( f \) cause unwanted vertical movement of the working organ that affects both the \( y \) coordinates and the change in the digging depth \( h \). Dependence of the blade position and dig depth from disturbances \( f \) reflects the intricate relationship between the geometric parameters of the bulldozer in space.

Loading conditions on the working organ are due to random variation in the dig depth and heterogeneity of soil properties. Soil digging process with bulldozer working organ is studied on the base of the finite element model of the soil mass, a mathematical model of random forces of resistance on the working organ \( P \) being developed.

The actual bulldozer velocity \( v \) depends on the strength \( P \) and the properties of the mover, transmission and the power unit. In its turn, disturbance parameters, movement of the working organ and the formation of stress depend on the velocity \( v \). Bulldozer drive model and mover interaction with the soil include engine model, mechanical and hydro mechanical transmission, as well as slipping.

Control system regulator depending on the objectives, control algorithm and the incoming data from the bulldozer as a control object produces electrical signals \( c \) to the electro- hydraulic distributors being part of the working organ hydro drive. Lifting or burying the blade is done to control either the pulling power \( N \), or the blade coordinates \( y \). The following describes the models of the bulldozer workflow elements.

A formation model of the random forces of resistance on the working organ being developed as follows\(^{[5]}\):

\[
P = P_d (1 + P_f);
\]

(1)
where $P_{tr}$ is the trend of resistance forces depending on the dig depth $h$; $P_f$ are the normalized random fluctuations caused by the heterogeneity of the soil (Figure 2).

Auto correlated random signal $f$ is generated based on the specified values of the autocorrelation function $\alpha_f$, $\beta_f$, parameters, and the standard deviation of the coordinates $\sigma_f$, as well as speed $v$. Soil cutting depth $h$ associated with $f$, geometric parameters of the bulldozer, its speed $v$ and extension rods of hydraulic cylinders of the working organ $l$. Normalized fluctuations $P_f$ dependent on heterogeneity of physical and mechanical properties of the soil are a random signal with a standard deviation $\sigma_{P_f}$ generated by the given parameter values of the autocorrelation function $\alpha_f$, $\beta_f$ and depending on the speed of the bulldozer $v$.

Component of the random process $P$ is due to heterogeneity of the soil and equals $P_f P_r$ while the standard deviation of the $P_r$ process equals the coefficient of fluctuations variation $\sigma_{P_r} = \psi_f$. Autocorrelation functions $R_f(l)$ of micro profile $f$ coordinates can be approximated by the expression:

$$R_f(l) = \sigma_f^2 e^{-\alpha l} \cos \beta l; \quad (2)$$

Where $l$ is the waypoint coordinate; $\sigma_f^2$ is the variance of the random process; $\alpha$, $\beta$ – are coefficients of the autocorrelation function.

The corresponding expression of the spectral density of disturbance at a bulldozer constant speed:

$$S_f(\omega) = 2\alpha \sigma_f^2 \frac{\alpha^2 + \beta^2 + \omega^2}{(\alpha^2 + \beta^2 + \omega^2)^2 - 4\beta^2 \omega^2}; \quad (3)$$

Generating a random signal $f$ is performed by filtering white noise $Q$ with specially created shaping filter. A discrete transfer function of the shaping filter, corresponding to (3) being generated as follows:

$$W_f(z^{-1}) = \frac{a_0 + a_1 z^{-1}}{1 + b_1 z^{-1} + b_2 z^{-2}}; \quad (4)$$
Discrete shaping filter for generating random auto correlated signal \( f(n) \) is also represented with the recurrence relation:

\[
f(n) = a_0Q(n) + a_1Q(n-1) - b_1f(n-1) - b_2f(n-2);
\]

(5)

Where \( n \) – is the current number of element sequence \( f \) or \( Q \); \( a_0, a_1, b_1, b_2 \) – are the shaping filter coefficients.

A continuous transfer function of the shaping filter disturbances from the ground conditions being generated as follows:

\[
f(t) = \frac{\sqrt{2\alpha}}{v^2} \frac{dQ(t)}{dt} + \frac{\sqrt{2\alpha}}{v^2} \sqrt{\alpha^2 + \beta^2} Q(t) - \frac{1}{v^2} \frac{d^2f(t)}{dt^2} - \frac{2\alpha}{v} \frac{df(t)}{dt};
\]

(6)

The coefficients of the transfer function (6) are dependent on the speed of the bulldozer. For convenience of the shaping filter implementation in MATLAB, a second order differential equation that relates the white noise \( Q(t) \) in the shaping filter input with disturbance \( f(t) \) at the output has been obtained:

For simulation disturbances caused by ground conditions, the differential equation (7) is implemented as a subsystem of MATLAB / Simulink. This subsystem is applicable to both continuous and discrete models for bulldozer workflows. Modeling disturbances from soil heterogeneity, i.e. fluctuations in the resistance force on the working organ \( P_f \), is accomplished similarly to (7).

Mathematic model is developed to describe the influence of soil surface micro profile coordinates derivation on the end-effectors coordinates as well as on the digging depth, taking into consideration bulldozer’s major geometrical parameters and its velocity. Average digging depth is also influenced by the distance between the blade side shift and the turning table \( L_{vi} \) as follows:

\[
h_{sr} = h_n + \frac{h_l - h_n}{G} \left( \frac{G}{2} - L_{vi} \right) = h_n + \left( h_l - h_n \right) \left( 0.5 - \frac{L_{vi}}{G} \right);
\]

(8)

Simulation model realization to show (Figure 3) correlations between geometrical parameters and velocity \( v \) allows to estimate the influence of perturbation actions (stochastic changes in surface altitudes of the right \( f_r(t) \) and left \( f_l(t) \) tracks) on the end-effectors altitude \( y_r(t) \) and \( y_l(t) \), as well as on the average digging depth \( h_{sr}(t) \).

Dynamic model is developed to form traction prism and to describe the dependence of prism volume \( V_{sr} \) and digging depth variable \( h \) and bulldozer’s moving velocity variable \( v \). Analytical expression for prism volume at the given moment of time \( t \) is obtained:

\[
V_{sr}(t) = B \sin \alpha \int_{t_0}^{t} v(t)h(t)dt - \frac{\sin \alpha \cos(\alpha + \rho)}{\cos \rho} \int_{t_0}^{t} h(t)v^2(t) \left[ 1 - \exp \left( -p \frac{B \cos \rho}{v(t)\cos(\alpha + \rho)} \right) \right] dt dt;
\]

(9)
where \( B \) – is the blade width; \( \alpha \) – is the entrance angle; \( \rho \) – is the soil inner friction angle; \( p \) – is Laplace operator.

The developed models for bulldozer workflow elements are to be used for separate bulldozer units study with the help of analytical dependences between workflow parameters as well as for bulldozer general workflow simulation. Elements models of bulldozer workflows being developed are intended both for the research of individual bulldozer units using analytical relationships between the parameters of the workflows and simulation of bulldozer workflows in general. When constructing a discrete simulation model, the following assumptions are taken:
- the linear motion of the machine is investigated;
- the design is considered to be rigid;
- backlash and friction between the elements of the working equipment are not considered;
- the elastic-damping properties of movers are not considered;
- the dynamic characteristics of a diesel engine with fuel regulator and hydro mechanical transmission torque converter are replaced with static;
- coordinates of the treated soil surface are completely determined by the coordinates of the cutting edge of working organ;
- engine power selection to the drive of the working organ and auxiliaries are neglected;
- rate of motion of hydraulic cylinders rods for lifting and burial of the working organ is identical and does not depend on the applied load;
- mover rolling resistance is constant.

A simulation model is implemented in MATLAB / Simulink (Figure 4).
Figure 4. Simulation Model for Bulldozer Workflow.
3 Neural Network Model of Bulldozer Workflow

The Autoregressive model structure with external inputs (Figure 5) is a dynamic two-layer recurrent neural network. It is found from the autocorrelation signal functions that the autocorrelation coefficient is greater than 0.8 in the time interval 0.1 sec. for speed $v(t)$ of 0.5 sec. for digging depth $h(t)$ and 0.2 sec for the resistance force $P(t)$. Length of delay lines TDL taking into account the sampling frequency of 10 Hz are up to 1, 5 and 2 accordingly (Figure 5).

Vector for adaptive model adjustable parameters comprising weights and displacements of neural network,

$$\mathbf{X} = [\mathbf{b}^1, \mathbf{b}^2; \mathbf{IW}^{11}, \mathbf{IW}^{12}, \mathbf{LW}^{12}, \mathbf{LW}^{21}].$$  \hspace{1cm} (10)

Criterion for neural network model optimal tuning, i.e. current learning error at a given moment of time we take as follows:

$$F(\mathbf{X}) = e(t) = P(t) - a^2(t) \to 0; \hspace{1cm} (11)$$

The network learning task is the task of multiple non-linear optimisation

$$\mathbf{X} = \arg \min_\mathbf{X} \| F \|; \hspace{1cm} (12)$$

The author propose the bulldozer workflow neural network model adaptive learning algorithm based on the recurrent least square method (exponential forgetfullness method) and on the algorithm of Forward Perturbation or dynamic back propagation[6,7]. In the process of learning the neural network accumulates information on workflow dynamics, new tendencies of process development prevail on the earlier ones at that. Degree of importance for the previously learned information is considered with forgetfullness parameter $\lambda$. Network optimal learning criterion gradient comprises frequent derived
learning errors based on neural network model adjusted parameters:

$$\nabla F = \frac{\partial F}{\partial X} = \left[ \frac{\partial F}{\partial b^1}, \frac{\partial F}{\partial b^2}, \ldots, \frac{\partial F}{\partial w_1^{11}}, \ldots, \frac{\partial F}{\partial w_2^{13}}, \frac{\partial F}{\partial w_1^{22}}, \ldots, \frac{\partial F}{\partial w_2^{23}} \right] = -\nabla^2 a^2 = \left[ \frac{\partial a^2}{\partial b^1}, \frac{\partial a^2}{\partial b^2}, \ldots, \frac{\partial a^2}{\partial w_1^{11}}, \ldots, \frac{\partial a^2}{\partial w_2^{12}}, \frac{\partial a^2}{\partial w_1^{22}}, \ldots, \frac{\partial a^2}{\partial w_2^{23}} \right].$$

(13)

Software algorithm of adaptive learning for neural network model of bulldozer workflow has been designed and implemented. The weight vector and bias network $X(t)$ are adjusted in accordance with the recursive expressions at each time step:

$$\dot{X}(t) = X(t) - \frac{P(t-\Delta t) \times \nabla F(t) \times e(t)}{\lambda + [\nabla F(t)]^T \times P(t-\Delta t) \times [\nabla F(t)]}.$$  

(14)

Covariance matrix of the vector $X(t)$ of neural network parameters used in the algorithm:

$$\dot{P}(t) = P(t-\Delta t) - \frac{P(t-\Delta t) \times \nabla F(t) \times P(t-\Delta t) \times [\nabla F(t)]^T \times [\nabla F(t)]}{\lambda}.$$  

(15)

4 Conclusions and Results

Adaptive neural network model of digging allows you to simulate and predict the dependence of the resistance strain of gauge bogie displacement depending on the dig depth and trolley speed in dynamics. The accuracy of the prediction $P(t)$ being estimated, the average relative error after learning the network is 4.5%.

A neural network model of bulldozer workflow has been developed, allowing to model the dependence of pulling power from the blade penetration. Input model signal, used for training, simulation and verification is presented in Figure 6. Adaptive learning for the model is stopped at time $t = 9.5$ sec. Receiving at this moment a neural network model parameter values, modeled digging resistance force and speed of the machine (Figure 7, 8) are accomplished, as well as the forecast for another 0.5 seconds is developed.

Figure 9 shows the output of neural network models-pulling power of the bulldozer. In modeling and prediction of the neural network output is close to the experimental data only in the time interval of 7-10 sec. This is due to a change in unmeasurable chip thickness, as well as the rapidly changing conditions of the mover clutch with the ground. Therefore, the parameters of the adaptive neural network model must be adjusted in real time. The accuracy of prediction of pulling power $N(t)$ has been estimated; the average relative error being 14.7% on an interval from 7 to 10 s.
Identification Technique of bulldozer workflows and models obtained on its basis, are designed for use in the development of adaptive systems of automatic workflow management of bulldozer.

The development methodology of the adaptive control systems of bulldozer workflows is based on the application of neural network technology. For the formation of the control actions influencing the bulldozer, particularly electrical signals actuating control valves of hydraulic cylinders lifting and lowering the working organ, the structure and algorithms of adaptive neural network controller have been designed.

References


Generation the 3D Model Building by Using the Quadcopter

T.T. Bertram\textsuperscript{a}, T.T. Bock\textsuperscript{b}, A.G. Bulgakov\textsuperscript{c} and A.A. Evgenov\textsuperscript{d}

\textsuperscript{a}Technical University of Dortmund, Germany
\textsuperscript{b}Technical University of Munich, Germany
\textsuperscript{c}South West State University, Kursk, Russia
\textsuperscript{d}South Russian State Technical University, Novocherkassk, Russia

E-mail: torsten.bertram@tu-dortmund.de, Thomas.Bock@bri.arch.tu-muenchen.de, a.bulgakov@gmx.de, alexeyevgenov@gmail.com

Abstract -
Last decade saw a considerable growth in number of methods to create 3D models for buildings. Some companies have developed their own online maps systems. One of the most well-known of them is Google Maps. The maps are infrequently updated because the process is costly and labor-consuming, and, as a rule, the update is not necessary. The process of 3D model creation for a separate building using such equipment as LIDAR and UltraCamD is very expensive; and land-based aerophotogrammetry has its limitations based on the height of the building and its location that should be suitable to place the camera to get an accurate photos.

The article presents results of using 3D reconstruction technique based on 2D pictures of a building taken by a quadcopter. The latter is an unmanned aerial vehicle (UAV) being used to survey buildings of various heights without interfering with public transport. Moreover, the quadcopter can reach any position required to take the necessary angle for a favourable photos.

The building under consideration has been surveyed using Autodesk 123D package to reconstruct 3D geometry of it. The results prove that the quality of regenerated pictures taken by quadcopter is comparable to those received using UltraCamD and Photogrammetry equipment, facade pictures being even of better quality. The modeling results show good image quality and sufficient area of overlap which makes the model more real.

Keywords -
3D-model; 3D-reconstruction; Building; Aerial photogrammetry; Quadcopter; UAV; Texturing

1 Introduction
Last decade saw a considerable growth in number of surveys concerning methods to create 3D models of buildings on the basis of photographs. Authors in [2], for example, combined aerial photographs and ground mapping pictures to create the 3D model for a building of complex architecture. They used a HexaCopter, an unmanned aerial vehicle (UAV), and special software to create a map based on the photographs. The software program triangulated the pictures and then generated orthophotos to obtain the footprint of the building. After that 3D model and the footprint of the building were combined into one coordinate system. The method described is cost-saving comparing to 3D model plotting, but it is more time consuming as the process takes up several days.

<table>
<thead>
<tr>
<th>Method</th>
<th>Tacheometry method</th>
<th>Manned aircraft</th>
<th>UAV technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field-work</td>
<td>Traversing – 1 day</td>
<td>Traversing, pre-marking and flight planning – 96 hours</td>
<td>UAV setup – 20 min and Flight – 30 min</td>
</tr>
<tr>
<td>Processing</td>
<td>Generate the topographic plan – 1 day</td>
<td>Image scanning and processing – 1 month</td>
<td>Image processing until map production – 4 hours</td>
</tr>
<tr>
<td>Man power</td>
<td>4 person</td>
<td>2 pilots and 1 pilot and 2 person on the ground</td>
<td></td>
</tr>
</tbody>
</table>

Another example of unmanned aerial vehicle application is shown in [3]. Cropcam, a drone with fixed wings, proved its efficiency to create topographic maps as well as showed some advantages such as time consumption, in comparison with other wide-spread methods. The comparison is presented in Table 1. Visual and plotting analysis shows the accuracy of 1 meter for the maps created with the method described.
Highly realistic pictures of facades and roofs are obtained by the authors in [12] using pictometry and photographs taken by UltraCamD camera. Combination of vertical and oblique pictures provides good results to create high-quality 3D models of urban areas. Maximum discrepancy when using UltraCamD reached 0.042 m, maximum discrepancy when using pictometry was no more than 0.396 m. But there are several problems that occur while using the method. The first problem is the difficulty to generate the textures hidden by trees, cars and other objects in front of the façade. You can see the example of a façade blocked by vegetation in Fig. 1.

![Figure 1. Façade textures hided by vegetation around the building [8]](image1.png)

3D reconstruction method for facades that uses two parallel 1D high resolution cameras providing stereo-images of textures is presented in [5]. The experiment is carried out by an automobile equipped with cameras moving along the road at the speed of 40 km/h. The authors of the methods are able to overcome the problem of textures blocked by trees and telegraph posts. But the method has its own limitations: the automobile can move only along the roads and cannot overtake some big blocking obstacles in front of a building.

![Figure 2. LiDAR-equipped aircraft during the process of 3D landscape survey](image2.png)

The importance of oblique images when texturing the building is described in [1]. They are used to provide better quality of obtained textures.

Considering the abovementioned methods of 3D models creation it is necessary to note that it is unreasonable to use commercial aircraft to generate a model of a single building. The ability of ground photography is also limited. It is worth mentioning that quality increase results in computational expenditure increase, while acceleration of photo-processing speed results in quality decrease, the problem to reduce graphic information processing time is not, though, among the tasks of the presented survey.

## 2 Aerial Photographs Means

Great majority of all topographic maps and 3D models of buildings are obtained while using commercial aircrafts, as described in [12]. The following equipment, as a rule, is mounted on the aircraft:

- High-resolution camera with IMU sensors to register metrical values, such as Ultracam D with picture accuracy up to 2 mkm [9]. The camera comprises IMUs and a GPS-receiver to register picture location coordinates and to filter noises.

- LiDAR-scanner of high-resolution mode with 18 cm accuracy for vertical pictures and 30 cm accuracy for horizontal pictures [6].

Commercial aircraft application efficiency is proved by high quality of photos taken and by multitude of topographic maps and 3D models created. Fig. 2 shows the cabin of the aircraft equipped with Ultracam D camera. Aerial pictures taken at high altitude from an aircraft are presented in [4], [7] and [8]. The method provides high-resolution pictures and simplifies the process of 3D geometry model creation. But the aircraft application is very expensive especially if a 3D model is needed only for one building.

At present UAV are used for various purposes, both military and civilian ones such as monitoring, payload transportation, medical purposes and others. Their application field is widening daily.

UAVs are robotic transport systems with remote control, sensor sets and sensor control system. They have imbedded function to follow the route of a pre-assigned trajectory. UAVs can be equipped with additional devices. Two types of UAVs are available: with fixed wing and rotor-type ones. The main advantage of the latter is a special flying mode of hovering which is extremely useful for photo data collection. Fig. 4 shows a rotor-type drone to create 3D models of buildings.

## 3 Methodology

The process of 3D model creation for a building depicted in Fig. 3 begins with working process planning, equipment preparation and its tuning. Quadcopter is a
transportation means for GoPro camera which is calibrated in advance to avoid self-calibration.

Then there is the photography stage. The photographs are taken with overlapping to simplify the later search of correspondence points in the pictures, to simplify pictures matching and to increase the quality of model’s textures. The working process ends with photographs processing and 3D model creation.

3.1 Equipment

To ensure success of 3D model creation we use four-rotor UAV equipped with high-resolution camera as shown in Fig. 4. UAV simplifies the process of taking pictures because operator can set the trajectory of the vehicle individually. We use 3DR IRIS quadcopter on the basis of Paxhawk autopilot. The vehicle has the embedded function to store the route waypoints to follow them during the flight. To obtain high-resolution photos GoPro Hero 3 camera is used. The camera is of HD resolution with 1920x1080 pixels for HD-video at 60fps and 12MP CMOS-pictures (complementary metal-oxide-semiconductor). Another advantage of the camera is remote control of the shutter as well as preview mode for images. The camera weighs 77 g.

GoPro is a wide-angle camera with high yield of distortion. That is why it is necessary to calibrate it and to take pictures very close to each other with the overlapping as wide as possible. This will diminish the effect of short focus and wide angle of GoPro [10]. The aim of camera coefficients determination when calibrating is to state the exact size of pixel, the optical axel deviation of the lens, the length of focus and lens distortion. The results of calibration process are show in Table 2. But even with all these calibrating efforts the quality is not as high as that of a digital camera. Additional tuning can help reduce errors when radial distortion coefficient is added.

Figure 4. Image acquisition equipment: a) remote control quadcopter and b) GoPro 3 camera

<table>
<thead>
<tr>
<th>Table 2 Camera calibration parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changing parameters</td>
</tr>
<tr>
<td>Focal LengthX</td>
</tr>
<tr>
<td>Focal LengthY</td>
</tr>
<tr>
<td>ImageX Center</td>
</tr>
<tr>
<td>ImageY Center</td>
</tr>
<tr>
<td>Radial Distort Param1</td>
</tr>
<tr>
<td>Radial Distort Param2</td>
</tr>
<tr>
<td>Tangential Distort Param1</td>
</tr>
<tr>
<td>Tangential Distort Param2</td>
</tr>
</tbody>
</table>

3.2 Photo Data Collection Process

Before the flight it is necessary to set the trajectory of the UAV movement. At this stage a set of waypoints, direction and camera resolution is assigned to the UAV by a special software program (Figure 5). When control points are determined and saved in the autopilot system of the UAV, the vehicle takes pictures automatically. During the flight the operator can preview the taken pictures by means of Wi-Fi interface of the GoPro camera and control the shutter.

Majority of UAV flying parameters can be controlled in real time, such as its location, position, distance, altitude, speed, and battery charge level.

3.3 Photo Processing

When pictures have been taken, they are processed in several stages. High-resolution picture processing requires a lot of computation expenditures. Each stage of 3D model creation algorithm includes finding correspondence for pictures, precisimg them, creating the points cloud for object’s surface, defining the surface
1. Preliminary map creation. Initial analysis of the pictures is carried out to identify correlation between pictures. Preliminary map is created either using GPS information or by AtiPE (Automate Tie Points Extraction) [1] procedure for extremely low-resolution pictures.

2. Features Determination. SIFT operator allows deriving features in the pictures. SIFT Algorithm includes four stages: scale-space extreme detection, keypoint localization, orientation assignment, keypoint descriptor creation.

3. Corresponding Features Determination for Pairs of Pictures. Corresponding feature are to be found at comparison of features description acquired at the previous stage. To search for correspondence, kd-tree procedure is used, the latter being based on Approximate Nearest Neighbours (ANN) and Fast Library for Approximate Nearest Neighbours (FLANN) libraries. The method provides a sufficient number of correspondences, though some outliers can be found among the results. To eliminate the outliers the epipolar constraints found in significant E matrix and in fundamental F matrix are used. RANSAC reliable algorithm finds derived E and F values to identify outliers.

4. Integration of Images. When all correspondences between pictures are identified, the method allows integration of all images together. A sequence of images is divided into n-2 groups of three, and correspondences of each three images are determined. Then, image coordinates of the next group is compared with the previous one. The method bears linear computational cost depending on the number of pictures.

5. Image Coordinates refinement. Matched points image coordinates acquired with the help of SIFT descriptor are not accurately oriented. To increase the accuracy of image coordinates least squares matching (LSM) algorithm is used.

6. 3D Model Creation. Image coordinates for all matched points are the result of the method applied. They are later used to orient the image and to create sparse geometry of the object under consideration. Scene reconstruction is then performed using bundle adjustment method in correlation with the scene in the pictures and the object.

The acquired point clouds are converted into a structures network. The object model is then textured using the known camera and pictures parameters.

3.4 Test Site and Data

The new laboratory building of Electromechanical Department of Dortmund Technical University located at the Northern Campus at Dortmund, Germany, is chosen as the surveyed object.

The created 3D model of the building is of a simple form of rectangular parallelepiped as shown in Fig. 6. The building is a good practical example to study as the surface of its walls is flat and the façade is not complicated. The present article does not focus on roof texturing, only on building’s façade texturing.

Partial model of the building has been acquired by the time the present article is written. The reconstructed 3D model part comprises western and southern parts of the building.

Vertical images overlapping varies from 40% to 50%, horizontal overlapping is from 20% to 30%.
4 Results and Analysis

Pictures, having been taken by the quadcopter, are than processed applying the algorithm of 3D model reconstruction from 2D photographs, Figure 8.

Deviation values are measured by comparison of the known geometrical parameters of the building and the acquired photogrammetry results. Coordinates of 30 points onto an object were used and their 3D coordinates were computed only as intersections of homologous rays, after the bundle adjustment. Check points on the façade were assumed as check points. Their 3D coordinates were derived with a total station.

![Figure 7. Surface reconstruction for the model](image)

Surface reconstruction for the model takes up one hour; camera position adjustment takes 15 minutes, number of 1920x1080 pictures is 12 (Figure 7).

The results of accuracy calculation are presented in Table 3. Notably, the standard deviation of the model is less than 10 mm, and the depth accuracy deviation is 15 mm. The deviation along the X-axis and the Z-axis lies in the range between 2 mm and 58 mm and between 2 and 45 mm, respectively. The depth accuracy lies in the range up to 115 m.

![Figure 8. Surface reconstruction for the model](image)

Table 3 Instrumental errors for 30 points of the model

<table>
<thead>
<tr>
<th>Deviation</th>
<th>X</th>
<th>Depth</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation, mm</td>
<td>8</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Maximum deviation, mm</td>
<td>58</td>
<td>115</td>
<td>45</td>
</tr>
<tr>
<td>Minimum deviation, mm</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3 shows the model deviations. Judging by the derived accuracy values for the model it is necessary to note that depth accuracy of some image areas is quite low (up to 1 m.). The problem can be solved by taking more pictures of the required areas from different angles. In general, the model presents a realistic picture of the object’s geometric properties and its major parameters with accuracy of up to 8 mm.

5 Conclusion and Future Work

The results of the survey prove that high-resolution aerial pictures can be acquired using quadcopter with a mounted digital camera. The technique simplifies the building facades texturing process. Deviations of the described models are the results of insufficient number of pictures taken. Unmanned aerial vehicle application for photogrammetry is far more cost-effective than usage of commercial aircraft.

Further work will be devoted to creation of full 3D models of buildings with roof texture. Buildings with complicated façade structures are also in the scope of the author’s interest.

References

[1] Barazzetti L., Scaioni M. and Remondino F. Orientation and 3d modelling from markerless terres-

782


Performance Evaluation of Mecanum Wheeled Omni-directional Mobile Robot

B. Chu

Department of Intelligent Mechanical Engineering, Kumoh National Institute of Technology, Yangho-dong, Gumi-si, Gyeongbuk, South Korea, 730-701
E-mail: cbs0816@gmail.com

Abstract - Mobile robots with omni-directional wheels can generate instantaneous omni-directional motion including lateral motion without any extra space for changing the direction of the body. So, they are capable of traveling in every direction under any orientation to approach their destinations even in narrow aisles or tight areas. Especially, if a construction tool is combined to the mobile robot, it can be a mobile construction robot to be able to move from one position to another. In this research the Mecanum wheel, which is most frequently utilized in industrial fields, is selected to achieve omni-directionality of the mobile robot. Through intensive experiments, performance evaluation of the developed omni-directional mobile robot was conducted to confirm the feasibility for industrial purposes. Velocity performance and straightness for each directional motion were selected as performance indices to assess the omni-directional mobile robot. Ultrasonic sensors installed on the frontal and lateral sides were employed to measure the real-time distance between the mobile robot and the side wall of workspace. The linear position, angular position and velocity of the mobile robot were calculated with the distance information.

Keywords - Mobile robot; Mecanum wheel; Omni-directionality; Ultrasonic sensor; Performance evaluation

1 Introduction

Conventional mobile platforms with differential-drive system such as automobiles can generate 3-DOF motion like translational or rotational motion on 2-dimentional plane. However they cannot control each DOF independently, which is called non-holonomic characteristics [1]. For example, a differential drive mobile platform with a steering system is able to make forward and backward motion, but it cannot generate instantaneous lateral motion without rotation of the body. So, in order to generate lateral motion, quite complicated motion plan and dexterous operation are required such as combination of several times of forward and backward motion and rotation of the body using the steering system. This property of conventional mobile platforms gives a bad effect on development of industrial mobile robots in aspect of easy control system realization as well as simple mechanical design.

To develop a mobile robot system for automation process, this research suggests a mobile platform based on omni-directional wheels [2]. This type of mobile platform is capable of overcome shortcomings of conventional mobile robots. In order to move sideways, conventional systems require some extra space to rotate the orientation of the body. The size of the space is usually much larger than that of the system itself. However, the omni-directional wheels enable immediate lateral motion without rotating the body platform. Moreover, since the mobile platform with omni-directional wheels doesn’t need any extra space to change the direction of the body, it can generate very flexible motion trajectory to approach its destinations even in narrow aisles or tight areas.

It is noteworthy that those working environments often occur in various industrial fields such as manufacturing facility, warehouse, hospital, exhibition space and so on [3]. This mobile platform wheel is commonly used in robotic applications requiring a high degree of maneuverability, such as those experienced by NASA for hazardous environment exploration [4]. The objective of the OmniBot project (Figure 1. (a)) is to develop a hazardous duty mobile base as an advanced development test bed to research alternate technical approaches for remotely controlled operations in hazardous areas. Airtrax ATX-3000 industrial forklifts (Figure 1. (b)) excel in applications requiring tight manoeuvring or transporting long loads sideways through standard sized doors or narrow aisle ways. The ATX’s unique, omni-directional movement allows it to travel in all directions thus making it an ideal vehicle to work in tight spaces where turns are not possible and finite control is a necessity [5]. Uranus (Figure 1. (c)) was the first mobile robot with Mecanum wheels, designed and constructed in Carnegie Melon University [6]. It was built to provide a general purpose mobile base to support research in to indoor robot navigation.
In this research, an omni-directional mobile robot based on Mecanum wheels was designed and manufactured. In this paper, the developed mobile robot prototype is briefly introduced and a few basic performances were evaluated by intensive experiments. Since it is a ‘mobile’ robot, the mobility was selected as the first performance index. After four-directional translation of forward, backward, left and right motion was generated, each directional velocity profile was obtained to observe it linear velocity and acceleration properties under a given distance. The second performance index is omni-directionality or omni-directional motion accuracy. It is measured by observing angular error between designated trajectory and measured trajectory for four-directional translation. The experimental results of the two performance indices are discussed in later sections.

2 Basic Properties of the Omni-directional Mobile Robot

2.1 Properties of the Omni-directional Wheels

Figure 2. shows various types of omni-directional wheels. Especially the Mecanum wheel and orthogonal wheel designs are based on a concept that activates traction in one direction and allow passive motion in another, thus allowing greater flexibility in congested environments. In this research the Mecanum wheel, which is most frequently utilized in industrial fields, is selected to achieve omni-directionality of the mobile robot (Figure 2. (a)) [7]. Due to holonomic characteristics to generate instantaneous omni-directional motion, the mobile platform with omni-directional wheels can make flexible 3-DOF motion including instantaneous forward, backward, lateral and rotational movement in planar space and realize a simple control system through intuitive operating method. The Mecanum wheel is commonly composed of several sub-rollers which are mounted around the rim wheel circumference at a specific angle to the wheel axis (Figure 3.). Figure 3. (d) is the lateral view of the Mecanum wheel where sub-rollers are attached around the rim wheel with a specific angle, \( \theta = 45^\circ \) in this study [8].

2.2 Mobile Robot Motion Mechanism Based on Mecanum Wheels

The suggested mobile robot has a squared platform with four Mecanum wheels at each edge in order to simplify the mathematical model and the motion control.
Using four of these wheels provides omni-directional movement for a vehicle without needing a conventional steering system. The sub-roller angled at 45° about the rim wheel divides the force driven by the wheel rotation into one portion in the rotational direction of the sub-roller and the other portion in the axial direction of the sub-roller. The rotational force portion is dissipated by rolling of the sub-roller. Combination of the individual direction and velocity of each Mecanum wheel generates the total resultant force vector in any desired direction. Then, the mobile robot moves to the direction of the force vector without changing the direction of the mobile platform. Figure 4. shows the overview of the motion mechanism of the 4-wheeled omni-directional mobile robot.

3 Control System Design of the Omnidirectional Mobile Robot

3.1 Electrical Circuit Design

In this research, CompactRIO (cRIO) and C Series motion & I/O modules of National Instrument Co. are utilized for the control system [9]. CompactRIO, which is a stand-alone embedded real-time processor, can equip various conditioning and I/O modules for sensors and actuators, so that it is often used as a control or data acquisition system. Table 1 presents the CompactRIO and C Series motion & I/O modules used in this research. NI cRIO-9082 is the embedded real-time controller, which has a chassis with 8 slots for inserting C Series motion & I/O modules. In order to coordinate four Mecanum wheels independently, four NI 9512 motion modules are employed. Analog and digital I/O signals are acquired or generated by NI 9205 and NI 9403 I/O modules. Four 800W servo motors and Sigma Series servo packs of Yaskawa Co. drive four Mecanum wheels. Figure 5. represents the control system hardware of the developed omni-directional mobile robot.

Table 1 CompactRIO (cRIO) controller and C series I/O modules

<table>
<thead>
<tr>
<th>Controller or Module name</th>
<th>Purpose</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>NI cRIO-9082</td>
<td>Embedded real-time controller</td>
<td>8 slots for I/O modules</td>
</tr>
<tr>
<td>NI 9512</td>
<td>Motion generation</td>
<td>Stepper drive</td>
</tr>
<tr>
<td>NI 9205</td>
<td>Analog input</td>
<td>8 channels</td>
</tr>
<tr>
<td>NI 9403</td>
<td>Digital input/output</td>
<td>32 channels</td>
</tr>
</tbody>
</table>

3.2 Electrical Components

To generate planar motion of the mobile robot, an operator adjusts a joystick. Figure 6. (a) shows the joystick (HFX-44R10, CH Products) adopted in this research. Forward/backward, lateral and rotational motion without changing the direction of the mobile platform or an additional space can be driven by tilting the joystick to forward/backward, lateral and rotational direction. An analog voltage signal from the joystick is given to the NI cRIO-9082 real-time controller via NI 9205 analog input module. Then an appropriate motor command signal made by the controller is sent to the motors, which are connected to the Mecanum wheels, via NI 9512 motion module. Three ultrasonic sensors are used to measure the distance between the mobile robot and the side wall. Figure 6 (b) shows the ultrasonic sensor (PID616110, SensComp) used in this research. The specification of the sensor is given in Table 2. It makes an analog voltage signal proportional...
to the distance, which is transported to the NI cRIO-9082 real-time controller via NI 9205 analog input module. For designated linear motion, if the operator inputs a desired position on the user-interface screen, cRIO-9082 real-time controller generates motion command according to a designated maximum velocity and acceleration rate. Then the mobile platform follows the motion command automatically. The angular position feedback information is provided by Yaskawa servo packs including incremental encoders.

![Joystick and Ultrasonic sensor](image)

*Figure 6. Joystick (HFX-44R10, CH Products) and ultrasonic sensor (PID616110, SensComp)*

<table>
<thead>
<tr>
<th>Table 2 Specifications of the ultrasonic distance sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model type</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Distance range</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
</tr>
</tbody>
</table>

4 Experimental Performance Evaluation

Ultrasonic sensors in stalled on the frontal and lateral sides were employed to measure the real-time distance information between the mobile robot and the side wall of workspace. The linear position, angular position and velocity of the mobile robot were calculated with the distance information. Figure 7. shows the laboratory environment where artificial side walls were installed for reflecting the ultrasonic signals.

![Laboratory view for experiments](image)

*Figure 7. Laboratory view for experiments*

4.1 Mobility

Since the developed system is a ‘mobile’ robot, the mobility was selected as the first performance index. After four-directional translation of forward, backward, left and right motion was generated, each directional velocity profile was obtained to observe its linear velocity and acceleration/deceleration properties under a given distance. Figure 8. and Figure 9. show the real experiment views for forward/backward and lateral translation.

![Initial and final position](image)

*Figure 8. Experiments for forward/backward translation*

![Velocity performance information](image)

*Figure 10. shows the velocity performance information for forward, backward, left and right translation. In these figures, while signal 1 is measured from the frontal side sensor, signal 2 and 3 are measured from the lateral side sensors. Therefore, signal 1 presents the forward and backward motion in Figure 10. (a) and (b), and signal 2 and 3 do the lateral motion in Figure 10. (c) and (d). Table 3 summarizes linear velocities of the four cases. In forward and backward motion, the linear velocities were 0.72m/s and 0.73m/s, and in the lateral motion, they were 0.64m/s when the desired travel distance and the angular velocity of the wheels were fixed.*

![Linear velocities summary](image)
It is considered that the linear velocity difference between the forward/backward motion and lateral motion is caused by the structural characteristic of the Mecanum wheel. In the case of forward/backward motion, the Mecanum wheels act like conventional wheels without using passive rotation of the sub-rollers. However, in the case of lateral motion, passive rotation of the sub-rollers and frictional slip on the ground happen and the rotational velocity component of the driving rim-wheel is converted to the axial directional velocity component which makes the lateral motion of the mobile platform. In this converting process, the frictional slip on the ground gives rise to velocity loss.

Table 3 Linear velocities for four-directional motions

<table>
<thead>
<tr>
<th>Direction</th>
<th>Linear velocity (m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>0.72</td>
</tr>
<tr>
<td>Backward</td>
<td>0.73</td>
</tr>
<tr>
<td>Left</td>
<td>0.64</td>
</tr>
<tr>
<td>Right</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Figure 9. Experiments for lateral translation

Figure 10. Velocity performance information for forward, backward, left and right translation

Figure 11. Angular error calculation for straightness
Table 4 Angular errors for four-directional translation

<table>
<thead>
<tr>
<th>Direction</th>
<th>Angular error(°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>0.86</td>
</tr>
<tr>
<td>Backward</td>
<td>1.45</td>
</tr>
<tr>
<td>Left</td>
<td>1.32</td>
</tr>
<tr>
<td>Right</td>
<td>3.13</td>
</tr>
</tbody>
</table>

4.2 Omni-directionality

The second performance index of the developed mobile robot is the omni-directionality or omni-directional motion accuracy. It is measured by observing angular error between designated trajectory and measured trajectory for four translation cases, forward, backward, left, right motion. In order to find the angular error for translation, distance information from the frontal and lateral ultrasonic sensors is measured. With the distance information, the final position coordinate is calculated, then the relation between the initial and final position gives the angular error. Figure 11. presents how to calculate the angular error for straightness where \((x, y)\) is the final position coordinate and \(\theta\) is the angular error. Table 4 shows the angular errors for four-directional translation. Since the accumulated angular errors are not compensated, there exist angular errors between reference and actual trajectory. For mobile application, some amount of slip between wheels and bottom always happens. Therefore, these kinds of mobile platforms are usually used in quite moderate situations where precise position accuracy is not required. Or an additional position control system should be employed.

5 Concluding Remarks

In this research, basic properties of the omni-directional mobile robot based on Mecanum wheels, control system design, and experimental performance evaluation were treated. The suggested mobile robot has a squared mobile platform and four Mecanum wheels at each corner. By harmoniously coordinating the four Mecanum wheels, immediate forward/backward, lateral and rotational motion, in other words, omni-directional motion is guaranteed. In electrical design aspect, NI CompactRIO embedded real-time controller and C Series motion & I/O modules were employed. The operator can give driving signal of the mobile robot on the LabVIEW front panel. Ultrasonic sensors installed around the mobile robot can measure environmental situation such as distance between the mobile robot and the side walls. Based on this mechanical and electrical design, a real prototype was manufactured in this research project and performance evaluation for mobility and omni-directionality was conducted. After intensive experiments, it was confirmed that the developed mobile robot guarantees quite decent performances.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A040562).

References

Cooperative Control of a Single-user Multi-robot Teleoperated System for Maintenance in Offshore Plants

Sunghoon Eom*, Seungyeol Lee*, Daejin Kim*, Dongbin Shin* and Jeon Il Moon*

*Division of Robotic System Research, DGIST, Republic of Korea
E-mail: sheom@dgist.ac.kr, syl@dgist.ac.kr, djkim1018@dgist.ac.kr, dbs@dgist.ac.kr, jimoon@dgist.ac.kr

Abstract -
In this paper, we propose a cooperative control method of a single-user multi-robot (SUMR) teleoperated system for maintenance in offshore plants that is designed to perform in a 1:N mode (here, “1” refers to the number of user, and “N” denotes the number of slave robots), in which a single user teleoperates a number of slave robots directly to conduct a specific operation or in an autonomous cooperation mode between slave robots in order to overcome the limitation of the 1:1 teleoperation mode. This paper is also designed to extend compatibility in the SUMR teleoperated system’s controller. A haptic device, which is one of the master devices, is used for duplex transmission of control and status data between a robotic manipulator and user. In this paper, a new haptic library is also designed to connect between a haptic device (PHANToM premium, Sensable) and a controller based on LabVIEW. The designed new haptic library (DLL, dynamic linking library) that is created using C++ is called in LabVIEW, which is a GUI (graphical user interface) based software development tool.

Keywords -
Cooperative control; Single-user multi-robot; Teleoperation; Maintenance; Haptic; Offshore Plants

1 Introduction

In general, a teleoperated robotic system consists of a master device that collects task commands from users and a slave robot that follows the collected task commands run in a 1:1 teleoperation mode. Such a 1:1 teleoperation mode becomes limited in real sites if the target work piece exceeds the payload of the slave robot, a required workspace exceeds a the slave robot’s workspace, or the target work piece is blocked by obstacles such that a user cannot perform teleoperation via visual information [1]. Furthermore, in this mode, if a user performs simple repetitive tasks continuously, the user can easily feel fatigue because of the repeated task commands generated using the master device, thereby negatively affecting the task quality.

This study proposes cooperative control method of a SUMR teleoperated system for maintenance in offshore plants that is designed to perform in a 1:N mode (here, ‘1’ and ‘N’ refer to the number of a user and slave robots), in which a user teleoperates a number of slave robots directly to conduct a specific task, or in an autonomous cooperation mode between slave robots in order to overcome the limitations of the aforementioned 1:1 teleoperation mode [2], [3]. The proposed control method is responsible for the role sharing and integrated management of slave robots to divide the operation mode of the slave robots into various types according to the user’s intervention level and the characteristics of the target task beforehand and to perform the target task using the robot operation mode selected by the user. The control method also includes technologies for the following: (1) a role change between slave robots that are controlled remotely and slave robots that perform the autonomous operation, (2) path generation for slave robots that perform the autonomous operation, and (3) path compensation when slave robots contact the obstacles while performing the task under various environments where one of the slave robots simply follows the task command given by a user via the software and the rest of the slave robots follow the operation of a slave robot that is remotely controlled, or where all slave robots perform an autonomous operation, to perform the target task set by the user [4]. Further, when the target task is a simple repetitive task that is performed continuously, the software includes a remote storing and playback technology using which the slave robots edit and store the task command created via remote storing from a remote user, thereby following the stored task commands repeatedly without any task intervention for the subsequent repetitive tasks performed to achieve the target task.

The cooperative control method proposed in this paper is also designed to extend compatibility in the SUMR teleoperated system’s controller. A haptic device, which is one of the master devices, is used for duplex transmission of control and status data between a
robotic manipulator and user. In this paper, a haptic device (PHANToM premium, Sensable) and a haptic device’s controller based on LabVIEW are connected using proposed haptic library. The proposed haptic library (DLL, dynamic linking library) that is created using C++ is called in LabVIEW, which is a GUI (graphical user interface) based software development tool.

2 Cooperative Control of a Single-User Multi-Robot (SUMR) Teleoperation System

The SUMR teleoperation system allows a user to remotely control multi-robots and aims to perform tasks that cannot be done with a single robot (e.g. bulky or heavy objects handling) through the cooperation between slave robots [3], [5]. In the above system, each robot has a single robotic manipulator mounted in an upper part of a single mobile platform, as well as various sensors, including a number of cameras [6]. The development goal of the SUMR teleoperation system is to improve work efficiency by performing particular tasks with multiple robots where a single robot would be either inefficient at or incapable of performing specific tasks [7], [8].

![Image](image_url)

Figure 1. Single-User Multi-Robot (SUMR) teleoperated system with Integrated Robot management (IRM) software

Figure 1 shows the schematic diagram of the SUMR teleoperation system proposed in this paper, in which the number of users and robots can be changed according to the characteristics of the specific tasks. As shown in Figure 1, a user determines the number of users and robots required for performing the specific tasks, the cooperation mode of the robots, and the authority needed for carrying out the task between users and robots first via the GUI in the master device connected to the IRM software. A user then selects from four modes the behaviour most suited to the task required: general teleoperation with 1:1 systems, cooperation between a user and robots, and cooperation between robots with (or without) a user’s task commands in the GUI. Once the cooperation mode determined for each process is delivered to the remote IRM device as various control signal types, the IRM software performs functions such as the switching of input/output signals between user and robots and the generation (or calibration) of each robotic path in order to conduct cooperative motion between robots according to the cooperation mode chosen.

3 Integrated Robot Management (IRM) Software for SUMR Teleoperation System

The proposed IRM software as shown in Figure 1 is integrated management software that includes role sharing (or change) between robots and the generation (calibration) of each robotic path to allow cooperation between two or more slave robots for tasks that are difficult for a single teleoperated robot [9].

3.1 IRM Software

The IRM software classifies maintenance tasks in offshore plants into four categories primarily by the degree of user intervention required during the operation of the robots, based on the characteristics of the maintenance task (difficulty or risk level). Later on, it then classifies each of the above categories into four detailed task groups secondarily according to the physical motion characteristics of each maintenance task. Thus, the various tasks that can be performed are classified into 16 types of task. Also proposed is the mode in the SUMR teleoperation system that is applicable to each task group, so that the proposed SUMR teleoperation system can secure universal generality and efficiency at the same time.

The proposed IRM software is mainly responsible for the bilateral data transmission function between a user and robots, the management function of transmitted data, and the operational function for cooperation between robots for specified tasks [10]. In more detail, the proposed IRM software includes, firstly, a function of work-sharing (or change) between robots to ensure cooperation between a user and robots, and cooperation between robots with (or without) user’s task commands; secondly, a function of control of user intervention according to the characteristics of the tasks (difficulty or risk level) during cooperation between a user and robots; and thirdly, a function of generation, following, and calibration of robotic paths for performing autonomous motion (user’s intervention is 0%). Figure 1 describes the main functions of the proposed IRM software.

Your manuscript must be prepared for A4-size (210 x 297 mm) paper. Use the margin settings specified in Table 1 and do not number the pages of the paper.
3.2 Structure of IRM Software

The IRM software structure proposed in this paper is shown in Figure 2. As shown in the figure, whether slave robots perform teleoperated motion or autonomous motion is determined according to the cooperation mode (#1 ~ 4) selected through the GUI in the master device by a user. The cooperation mode can also be divided into linear or rotary motion according to the physical motion characteristics of the slave robot that performs the specific task. Thus, commands relating to the robotic path created via the master device by a user are transferred to a specific slave robot, which performs the teleoperated motion (Mode #1), while robotic paths for autonomous motion for other slave robots are created to assist the motion of that specific slave robot (Mode #2). In addition, a user can also create task commands (from the viewpoint of motions for target work piece) via the master device. These task commands are transformed into robotic paths for autonomous motion for each slave robot via the IRM software (Mode #3). Finally, in those instances where a task is repeatedly performed, the robotic path of each slave robot that corresponds to a single iteration can be stored and repeated as necessary (Mode #4). In Figure 2,'Mode #1' is same structure as general 1:1 teleoperation systems.

The IRM software includes a function that allows a user to switch a robot for control in the style of the 1:1 teleoperation system, even within the SUMR teleoperation system. A user can easily switch the target slave robot in the GUI of the master device if he/she wants to control robot #2 immediately after teleoperation robot #1 during ‘Mode #1’. Figure 2 (b) shows the switching structure between teleoperation target slave robots in the IRM software.

![Figure 2. Structure of (a) IRM software (Cooperation Mode #1~#4), and (b) Switching between target slave robots](image)

![Figure 3. Structure of (a) Mode #2, and (b) Mode #3 in IRM software](image)

Figure 2. Structure of (a) IRM software (Cooperation Mode #1~#4), and (b) Switching between target slave robots

Figure 3. Structure of (a) Mode #2, and (b) Mode #3 in IRM software

The ‘Mode #2’, as shown in Figure 3 (a), is a method of allowing cooperation between robots by having the secondary robots assist the primary teleoperated robot’s motion. Once a user creates a robotic path for the target slave robot teleoperated motion through the master device, the IRM software transmits the path to that robot while it also creates (including calibration) robotic paths for other slave
robots autonomous motion based on the motion information of the teleoperated robot. And ‘Mode #3’ is a method in which all robots perform autonomous motions based on commands for cooperative operation between robots to perform particular tasks. Once a user creates commands for cooperative operation between robots (motion commands about the target work piece) through the master device, the IRM software creates (including calibration) robotic paths for autonomous motion for each robot (Figure 3 (b)). Furthermore, ‘Mode #4’ is a method that is used when tasks are repeated iteratively. The IRM software stores the path of each slave robot in a specific subtask and has each slave robot perform its tasks by following the path repeatedly as needed. In particular, ‘mode #4’ has a structure that can store and playback the robotic paths stored in modes #1~3 selectively as needed.

4 LIBRARY FOR COMPATIBILITY BETWEEN HAPTIC DEVICE AND CONTROLLERS

In a teleoperated robot system, a master device is necessary to control a robotic manipulator as much as the slave robot system [11]. A haptic device, which is one of the master devices, is used for duplex transmission of control and status data between a robotic manipulator and user [12]. Recently, haptic devices are being used in various applications, such as industrial robots, medical or surgery robots, assistance robots, and laboratory based robotic platforms. Systems implementing haptic devices effectively replicate wrist motion and have been proven to be effective for manipulation of teleoperated robotic systems. Although haptic devices have been widely used in teleoperated robot applications, its functionality is unsatisfactory for teleoperated applications that require compatibility. This is because the software subsystems of these haptic devices have been created using packaged libraries [12], [13]. To assign compatibility, new haptic library is proposed to interface between a haptic device and a SUMR teleoperated system’s controller.

According to this study, a DLL (dynamic linking library) was generated in C++ for linking the haptic device with the embedded controller in the LabVIEW environment. Figure 4 shows the configuration of the generated DLL used for the haptic device. The generated DLL consists of a ‘Haptic Device Start’ feature for connecting with the haptic device, ‘Parameters’ of the haptic device, and a ‘Haptic Device End’ feature for releasing the haptic device. Here, the functions responsible for the start and termination of “Haptic Device Start” and “Haptic Device End” include functions used for initiating the haptic device and managing memory and threads. ‘Parameters’ include device information, location, direction, linear velocity, and angular velocity of the haptic device.

The DLL of a haptic device is generated by extracting based on each function (Solid rectangle of Figure 4). The DLL of the haptic device, written in C++, uses a library call function provided in LabVIEW, which is generated by the lower layer VI (SubVI) based on each function, as shown in Figure 5. The block diagram and the front panel, shown in Figure 6 (a), (b), respectively, for driving the embedded controller-based haptic device using LabVIEW were designed using the subVI.
5 Conclusion

In this paper, a SUMR teleoperated system with a haptic device using a new haptic library and cooperative control method, which adopts a strategy of cooperation between a user and multi-robots, were proposed to improve work efficiency. Since the proposed control method was initially developed with the aim of improving the interaction technology linking a user with robots, future research will continue to refine the system for use in a variety of applications, to industries and fields where high-dimensional and sophisticated task is required by combining systematically an experienced worker’s skill and intelligence with a robot’s accuracy and physical power. More specifically, both human-robot interface technologies that can perceive the various and complex intentions of workers accurately, and robot-based safety technologies that can prevent accidents occurring due to user’s error will be studied in the future.

References


Home Environment Interaction via Service Robots and the Leap Motion Controller

C. Georgoulas*, A. Raza, J. Güttler, T. Linner, and T. Bock

*Chair of Building Realization and Robotics, Technical University Munich, Germany
E-mail: christos.georgoulas@br2.ar.tum.de

Abstract -
Ageing society and individuals with disabilities faces numerous challenges in performing simple tasks in Activities of Daily Living (ADLs), [1]. ADLs represent the everyday tasks people usually need to be able to independently accomplish. Nowadays caring of elderly people becomes more and more important. Performance of ADLs in the long term view can be considered as a serious concern, especially when dealing with individuals who require extreme caregiver assistance. The objective of this paper is to introduce an Ambient Intelligence [2], real-scale home environment implementation, embedded with sensors and actuators, which enhances the independence and autonomy of the individuals upon performing ADLs. A real 1:1 scale experimental flat has been design and developed (Figure 1), in the authors experimental laboratory. A study was made to be able to determine the actual needs, required services, and functionality of the proposed augmented environment. In order to allow a high quality service delivery, a mobile autonomous rover [3] was introduced to act as the main human-machine interface between the user and the distributed robotic systems and actuators. The mobile rover was wirelessly interfaced with the distributed intelligence to allow efficient interaction with the user, either passively, by vocal commands issued by the user, or adaptively, by autonomously navigating into the home environment. Moreover the Leap Motion hand gesture driven controller [4] is used as an intuitive user interface, which allows sub-millimeter accuracy capabilities, interfaced to a Jaco 6-Degrees of freedom robotic arm [5]. Many robotic systems have been designed and produced for assisting ADLs, in order to compensate this loss of mobility. However most of these solution are operated via a keyboard or joystick, which according to the complexity of the robotic system, require a series of configurations and mode selection routines in order to allow a specific trajectory path to be implemented. The authors developed new intuitive interfaces to operate such devices, in order to reduce the involved operation complexity. The user is able with simple gestures to operate and control complex robotic manipulators and mobile robots, without requiring the use of a joystick or a keyboard. The resulting accuracy and evaluated performance of the implemented interfaces, allow error free, continuous, and adaptive interaction between user queries and actuated environment responses.

Keywords-
Robotics and Mechatronics; Sensor Fusion; Ambient Assisted Living; Autonomous Navigation; Leap Motion Controller; Jaco Robotic Arm

1 Introduction

Ambient Intelligence (AmI) is an emerging discipline that incorporates a degree of intelligence into human daily living. AmI is generally associated with systems that are receptive to individuals, reactive to the presence of persons, nonintrusive, ubiquitous, and mainly embedded. AmI is usually identified with intelligent sensors and software embedded in our daily environments [6]. Human-robot interaction and cooperation has become a topic of increasing importance, especially tasks unifying the workspace of humans and robots, which are most common in service applications.

Today, robots and distributed robotic sub-systems start to permeate our every day surrounding, enhancing it with services and additional features. At the same time, this permeation is on the way to transform our perception of what robots are, robot technology, robots’ possibilities and the environment they are merged with. This transformation which has to be understood as a natural part of the evolution of robotics, will especially become visible when robots enter the field of service and assistance [7].

Ageing society requires novel approaches for placing mechatronics and robotic service technologies in living environments for assisting in daily activities. Already in early development phases, knowledge at least from the architectural, medical and robotic field is necessary, and subsequent product development even requires further fields (e.g. appropriate groups within the healthcare system) to be also involved. By entering
the third stage of life (70 - 85) health problems and limitations occur more frequently. Therefore elderly people face huge difficulties to manage their ADLs independently. In most cases the straightforward solution in an intensive care retirement home. As retirement homes are running out of places (also hospitals) because of the demographic change problem, novel solutions for elderly people that offer independent living in their own home environment by the use of assistive robots should be designed and developed.

In this paper the use of a mobile rover as a communicator between the user and the distributed intelligence as well as novel intuitive interfaces for controlling robotic assistive systems based on human gestures and vocal commands, in a 1:1 scale experimental flat are described. The main idea is the realization of an AmI robotic assisted home environment, in order to increase the independence, autonomy and quality of living of ageing population, upon performing their ADLs.

The structure of paper is as follows. Related work is presented in section 2. In section 3 the experimental environment and the implemented interfaces for controlling the assistive robotic systems are explained. Conclusions based on the developed application which will serve the need of simple interfaces for human-machine cooperation and their interconnection with robotics and distributed intelligence are presented in section 4.

![Figure 1. Interactive Home Environment](image)

2 Literature Review

Some researchers already proposed integrated solutions for intelligent and assistive robotic environments as e.g. Robotic Rooms [8], Wabot House [9], or Robot Town [10]. The aim of those approaches was to distribute sensors and actuators in the environment which can communicate with the intended robot system, allowing simpler and robust robot designs. Since the 1980s several research groups have created environments and prototype buildings for so-called smart buildings. Based on Ken Sakamura's T-Engine Hardware and a complementary operating system, the Tron House 1, 2 and 3 have been built [11]. The US AwareHome [12] and PlaceLab [13] follow a similar approach and MIT's House n [14] includes even modular intelligent furniture that can be equipped with various sensor systems. Recently designed German prototypes of assistive homes, such as “Haus der Gegenwart” (house of presence) [15] and “Haus der Zukunft” (house of the future) [16], are exemplarily equipped with a variety of networked pervasive technologies integrated by modern design. Similar to our approach Smart Buildings and Robotic Rooms try to integrate sensor-actuator systems with architectonic elements. However, these approaches integrate mainly sensors, actuators and robots on an informational level. Furthermore, they are presenting implementations that are realized in a controlled experimental environment, and cannot be straightforwardly applied into a regular medium sized apartment to serve as an integrated assistive system for ADLs.

In the last few years, different optical sensors have been developed, which allow the mapping and acquisition of 3-D information. Various applications also have been introduced, which exploit the increasing accuracy and robustness, and the decreasing cost over time of 3-D sensors [17]. The applications range from industrial use, object tracking, motion detection and analysis, to 3-D scene reconstruction and gesture-based human-machine interfaces [18]. These applications have different requirements in terms of resolution, frame-rate throughput, and operating distance. Especially for gesture-based user interfaces, the accuracy of the sensor is greatly considered a challenging task [17, 19]. The Leap Motion controller introduces a new novel gesture and position tracking device with sub-millimeter accuracy [20]. The controller operation is based on infrared optics and cameras instead of depth sensors. Its motion sensing precision is unmatched by any depth camera currently available, to the best of the authors knowledge so far. It can track all 10 of the human fingers simultaneously. As stated by the manufacturer, the accuracy in the detection of each fingertip position is approximately 0,01mm, with a frame rate of up to 300 fps.

Speech-based applications primarily require uninterrupted, real-time speech recognition. An intelligent, interactive personal information assistant using natural speech is an example for replacing the slow stylus input and cramped graphical user interface of a PDA. Several recent applications, like voice control...
of GPS navigation systems and voice-controlled automotive infotainment services [21] crave a stable vocal interface. This certainly means that sophisticated natural language applications like handheld speech-to-speech translation require agile, low complexity, and real-time performance speech recognition algorithms. Various technological problems have interrupted the stationing of vocal applications on embedded devices. The most challenging of these problems is the computational necessities of continuous speech recognition for an application having extensive glossary scheme. The demand to decrease the capacity, size and power consumption for these devices compromises in their hardware and operating system software that restrict their capabilities. The raw CPU speed is one of the underlying compromises. Embedded CPUs generally lack hardware support for floating-point arithmetic functions. Furthermore, embedded devices also have very limited memory, storage capacity and bandwidth [22]. For these consequences, previous research [23, 24], has focused on simple tasks by strictly limiting vocabulary for an application. In the proposed implementation, PocketSphinx [22, 25], an open-source embedded speech recognition system capable of real-time, medium-vocabulary continuous speech recognition, was used by the authors, which was accordingly modified regarding the followed hardware infrastructure.

3 Proposed System architecture

A 1:1 scale experimental home environment was implemented in the authors laboratory in order to allow the implementation, development and testing of the human-machine interaction and robotic assistive systems.

3.1 Turtlebot Autonomous Navigation and Tele-operation

The Turtlebot mobile rover platform was used [3] to act as the human-machine communication interface (Figure 2). TurtleBot is able to acquire visual information using the Microsoft Kinect Sensor [18]. The depth sensor consists of an infrared laser projector combined with a CMOS sensor, which capture 3-D real-time data. The sensing range of the depth sensor is adjustable, and the Kinect embedded software is capable of automatically calibrating the sensor based on the user physical environment arrangement, accommodating for the presence of furniture or other obstacles. Due to the fact that Kinect sensor uses an infrared sensor, it can also provide with night vision abilities. Thus, in the proposed architecture, elderly people can be assisted by TurtleBot, in low lighting conditions and even if lights are switched off, i.e. during the night. Once TurtleBot is introduced into the home environment, the exact interior configuration must be known, in order to enable autonomous operation. A widely known technique which provides this kind of information to autonomous vehicles is the Simultaneous Localization and Mapping (SLAM) [26, 27]. SLAM is a technique used by robots and autonomous vehicles to build up a map within an unknown environment (without a priori knowledge), or to update a map within a known environment (with a priori knowledge from a given map), while at the same time keeping track of their exact current location and orientation within the environment. Such a technique is used in the proposed system, in order to get all necessary details concerning the working space of TurtleBot. Once the map of the flat is composed, the mobile rover can efficiently autonomously navigate within the environment space. Figure 3 depicts the recorded interior mapping.
Additionally the slave workstation enables the user to tele-operate the robot using any user input device connected to it. Additionally a communication interface was implemented to be able to tele-operate the Turtlebot using a standard Android/based phone or tablet device. The ICT connectivity is depicted in Figure 4.

3.2 Speech-driven Navigation and Tele-operation

As earlier mentioned in section 2, the PocketSphinx speech recognition application was used for the implementation of the vocal command driven human-machine communication interface. A special dictionary was created and vocabulary files were generated. The voice application compares the user vocal queries with the precompiled dictionary words and then publishes a matching word via the recognizer application. This allows the vocal-driven navigation of the Turtlebot throughout different locations in the flat. A set of five keyword phrases were used: “entrance”, “kitchen”, “bedroom”, “bathroom”, “living room”. When these words are detected by the Turtlebot onboard microphone input device, the requested destination point and orientation of the robot are sent as coordinates to the mobile rover. It operates autonomously dynamically avoiding any obstacles in its path, by reconfiguring in real-time its trajectory. The implemented algorithm flow chart for the speech driven operation is presented in Figure 5.

3.3 Leap Motion controller Navigation and Tele-operation

A new sensor which introduces a novel gesture and position tracking system with sub-millimeter accuracy was used to navigate the mobile rover by user issued hand gesture patterns. The Leap Motion controller, (Figure 6), tracks the user’s palm posture, and all the accompanied “roll”, “pitch” and “yaw” orientation angles, thus all information regarding the user palm Cartesian position and orientation is retrieved from the sensor.
Once the user palm is introduced into the operating range of the Leap Motion controller (a hemispherical area of approximately 1 meter radius), the roll and pitch values of the user’s palm posture are recorded with a rate of up to 300 times per second. Data streams for roll and pitch angles are generated by accordingly rotating the user hand towards the roll and pitch angles (Figure 7). The recorded roll angle values are interpreted into left/right rotation of the Turtlebot, and the recorded pitch angle values into forward and backward movement of it (Figure 8). An efficient algorithm using sequence of checks and conditions was implemented to enable a smooth translation between the recorded user palm posture, and the navigation commands sent to the Turtlebot. An intuitive “virtual joystick” function was thus implemented to navigate the robot using the Leap Motion controller – user palm interaction.

3.4 Leap Motion controller – Jaco Robotic arm interface implementation

In the proposed implementation the authors introduce a novel human-machine interface which offers intuitive and adaptive manipulation in ADLs, using the Leap Motion controller and the Jaco arm [29]. Some researchers already evaluated the efficiency of robotic systems, specifically robotic arms, used by disabled individuals in performing ADLs [30, 31]. An important parameter when concerning the efficiency of assistive devices for disabled individuals is the economic benefit in terms of comparing the robotic system cost with the total cost required for a caregiver, in a long term scheme. The Jaco robotic arm, can efficiently substitute caregivers as a cost saving alternative [5, 32].

Most manipulators are operated via a keyboard or joystick, which according to the complexity of the robotic arm, require a series of configurations and mode selection routines by pressing a series of buttons, in order to select the desired operating mode, or to perform a specific trajectory path translation. The authors research on new ways to operate such devices, in order to reduce the involved operation complexity, since elderly people tent to face difficulties upon operating a complex interface, or refrain from further on a daily basis since they consider the interrelated controlling scheme complicated. Instead of using the original joystick “Kinova Joystick” of the Jaco arm and thus having to switch between different modes of control, the joystick is replaced with the Leap Motion controller which allows for an intuitive human-machine interface realization. The Leap Motion controller monitors the user’s hand/hands, fingers, and all the accompanied positions and angles. All information regarding the user palm Cartesian position is retrieved from the controller and fed to the algorithm. The algorithm uses the current
and previous information supplied by the controller and achieves an optimum realistic mapping between the user’s real arm and the Jaco arm. Additionally the arm’s angular features such as roll, pitch, and yaw angles are considered to the mapping procedure, enabling a realistic translation of the human arm. The Jaco arm fingers were also programmed to follow all grasp and release operations performed by the user fingers. To address safety requirements between the user and the arm, safe zones were considered according to the operating workspace of the robot, in order to ensure safety in case of unintentional user hand movements within the operating range of the Leap Motion controller.

The Leap Motion - Jaco arm operation was adapted into various locations of the home environment, i.e. the bed, entrance, and also to a wheelchair within the flat (Figure 9). The main contribution of developing such interfaces is to introduce the assistive services and functions that the Jaco arm offers in performing ADLs. Ageing society faces numerous challenges in performing simple tasks in ADLs. Nowadays caring of elderly people becomes more and more important. Individuals with upper limb impairments, also face difficulties to perform ADLs, especially in cases where the impairments have resulted from spinal cord injuries, neuromuscular diseases, etc. Many technical aids have been developed to assist in impairments in the home environment. However these assistive devices provide limited functionality and cannot address in an efficient way independence and autonomy [33, 34]. By introducing the Jaco arm within the proposed interactive home environment, and by focusing on assisting the user on various daily living activities by an intuitive, easy to operate, gesture driven compact size sensor, a new human-machine communication scheme can be established.

3.5 Overall system architecture

The proposed system architecture is presented in Figure 10. The Leap Motion controller is connected via USB to the ROS Master PC, which is comprised by the TurtleBot onboard laptop. Additionally the Android-based mobile device is wirelessly connected to the ROS Master PC. Both the Leap Motion controller and the Android-based platform comprise the user “Input Level”. The related user input queries are:

a) “roll” and “pitch” angles of the user palm recorded data required to navigate the mobile platform (Leap Motion controller),
b) user palm posture in terms of position and orientation required for the manipulation of the Jaco arm trajectory planning and Jaco arm fingers grasp-release operations (Leap Motion controller),
c) key-based graphical user interface to navigate the Turtlebot in the proposed home environment (Android-based device).

The remaining of the system architecture comprises the Jaco arm operation (“Output Level”), where a dedicated miniPC according to the home environment setting (bedroom miniPC, Entrance miniPC, Wheelchair miniPC) is used to provide the recorded input data from the Leap Motion controller. The data exchange communication between to the Turtlebot Master PC and the dedicated miniPCs is implemented using standard TCP/IP sockets, where raw data recorded from the Leap Motion controller are wireless, via a WiFi local area network, transferred to the Jaco arm, which is connected via USB to the dedicated miniPCs. For the experiments conducted under the proposed work, a single Jaco robotic arm was used, manually exchanged between the bedroom, entrance and wheelchair settings.

Figure 9. Leap Motion Controller Jaco Robotic arm operation under various arrangements
Results and Further Development

The proposed robotic home environment interaction scheme is still under development in order to define final specifications and offered services, and to enhance its functionality and operation towards the user. Currently the vocal user queries for the voice guided navigation are limited to a specified dictionary list of queries. Tests implemented to evaluate the system performance presented a real-time response between user vocal queries and resulting signaling commands to the mobile robot motor drives. The observed real-time response efficiently provides a responsive system with minimal error rate. Errors are currently introduced by the voice recognition module, in case ambient noises are dominant. This error rate can be minimized by an appropriate voice recognition training stage, corresponding to a new set of vocal commands that would allow efficient differentiation of user queries and ambient noises false interpretation.

The Leap Motion controller undoubtedly comprises a very efficient and real-time response module. Due to its enhanced sampling frequency (up to 300 frames per second) and accuracy (submillimeter), it appropriately addresses the need for short time delays between user queries and system response. Fusing such a responsive controller device with the various robotic subsystems (Turtlebot and Jaco Arm), extended functionality can be enabled to the user, once many features and service dealing with navigating the mobile rover and actuating the robotic arm can be dealt with a highly intuitive manner, and more importantly via the same terminal.

Further development of the proposed human machine interface comprises user friendly operation of such complex robotic systems, since elderly people tend to face difficulties upon operating a complex interface, or refrain from further on a daily basis since they consider the interrelated controlling scheme complicated.

5 Conclusions

Performance of ADLs in the long term view can be considered as a serious concern, especially when dealing with individuals who require extreme caregiver assistance. The objective of this paper is to introduce an Ambient Intelligence real-scale home environment implementation, embedded with sensors and actuators, in order to enhance the independence and autonomy of the individuals upon performing ADLs. Thus, a 1:1 scale home environment has been design and developed in the authors’ experimental laboratory. The home environment was constructed according to the actual needs, required services, and functionality of the proposed augmented environment. In order to allow a high quality service delivery, a mobile autonomous rover was introduced to act as the main human-machine interface between the user and the distributed robotic systems and actuators. The mobile rover was wirelessly interfaced with the distributed intelligence to allow efficient interaction with the user, by user issued vocal commands and by autonomous navigation into the home environment. Moreover a novel high precision gesture driven 3-D scanning controller is used to develop an intuitive user control interface. By wirelessly interfacing this compact size gesture driven controller with the mobile platform and also a lightweight robotic arm, a novel human-machine control interface is implemented. Mobile rovers and robotic arms, require some standard user input devices, such as keyboards, joysticks, etc. By using this new compact gesture driven controller the authors propose new ways of operating such platforms, in order to reduce the involved operation complexity. Such an approach enables the straightforward incorporation of robotic systems into the proposed home environment, to enhance the independence and autonomy of individuals upon performing ADLs.
References


SENSING AND COMMUNICATION
Experimental Study of Wireless Sensor Networks for Indoor Construction Operations

Magdy Ibrahim and Osama Moselhi

Concordia University, Building Engineering Department, Montreal, Quebec, Canada

E-mail: magdy.omar@yahoo.com, moselhi@encs.concordia.ca

Abstract -
Emerging wireless sensor networks (WSN) technology offers a great potential in supporting current project management practices. Deploying wireless sensor networks on construction sites can lead to significant time and cost savings by providing accurate and near-real-time data to project management personnel. Continuous monitoring of labor usage, materials placement and equipment performance provides valuable data for assessing progress of construction operations and assists in improving safety and security on job sites. Construction activities take place in outdoor and indoor environments, while Global Positioning System (GPS) is ideal solution for tracking outdoor activities; it is not applicable for indoor application due to the lack of line-of-sight to satellites signals. Therefore, GPS-less means of tracking is required in indoor environments. While several research efforts had been attempted to develop indoor positioning systems utilizing various wireless technologies, there is no clear understanding of which wireless technology performs better in indoor construction environment. This research aims to experiment and test wireless technologies to aid the selection of wireless sensor networks configuration in support of current practice of progress tracking at construction on job sites. This paper describes experimental study conducted to determine the effectiveness of wireless technologies for dynamic indoor resource position tracking. The experiments investigate the challenges of wireless radio signals propagation in indoor environments.

1 Introduction

Accurate and frequent project progress tracking is critical for effective project control and on-time project delivery. Presently, GPS has been widely used for tracking of outdoor construction operations. Its theory of operation is based on measuring times of arrival (TOA) of radio signals travelling between orbiting satellites and a mobile GPS unit. The GPS location is then calculated using a triangulation algorithm based on measured times and satellites position. The major advantages of the GPS are its reliability, availability and practical accuracy, however it is not suitable for indoor applications due to the lack of signal coverage particular inside buildings. [1].

Indoor localization research has been going on for decades in the robotics field [2,3]. The fact that indoor localization research is to date a very active research area indicates that there are still many challenges left to resolve. The challenges depend on the required accuracy and reliability dictated by the application. The fundamental challenge indoors is that the radio frequency environment is characterized by limited coverage, severe multipath signal fading and non line of sight (NLOS) conditions, which severely impact wireless signals propagation. This paper is dedicated to experiment and investigate the challenges of wireless radio signals propagation in indoor environments.

2 Literature Review

Manually monitoring progress of construction projects is not only expensive, subject to human error, and approximate but also is delivered with a time lag. Field supervisory personnel on construction site spend between 30-50% of their time recording and analyzing field data [4] and 2% of the work on construction sites is devoted to manual tracking and recording of progress data [5]. In addition, since most data items are not captured digitally, data transfer from a site to a field office requires additional time. When the required data
is not captured accurately or completely, extra communication is needed between the site office and field personnel [6].

The construction industry lags behind other industries in adopting innovative new technologies. The need to accelerate the rate of technological adoption in the construction industry has been well documented in the literature [7]. The rapid advances in sensing technologies motivated researchers to study the feasibility of using such technologies to automate and integrate individual technologies for tracking and monitoring in the construction industry.

Recent research demonstrated that, data collection technologies and sensors coupled with mobile computers can provide cost-effective, scalable, and easy-to-implement progress tracking at construction sites [8,9,10,11,12,13,14]. Several data collection technologies had been utilized for tracking of construction activities, such as 3D imaging, Global Positioning System (GPS), Radio Frequency Identification (RFID), Ultra Wide Band (UWB), handheld computers, voice recognition and wireless technologies.

Cavanaugh [15] presented a system that uses radio frequency and radar technology to locate, in three dimensions, workers indoors. Another study by Teizer et al. [16] demonstrated that the use of remote sensing and actuating technologies such as RF, Ultra Wideband (UWB), and imaging technologies can improve construction safety by warning or alerting workers on foot and/or equipment operators in real-time when a too-close proximity to unknown or other construction resources can cause hazards. The studies by Teizer [16] have shown promising results in outdoor construction experiences as well as potential for indoor safety improvement. In this respect, Zhang et al. [17] used a multi-agent system to detect possible collisions or conflicts associated with operations of equipment on construction sites.

Due to limitations of the previously discussed technologies, the usage of WSN has been expanding in recent construction research efforts. A WSN is a self-organizing network composed of a large number of sensor nodes, closely interacting with the physical world. It features low-cost nodes, extensive network capability allowing deployment of large quantities of nodes so as to increase the network coverage, stability and reliability in wireless communication.

A new tracking architecture was implemented using wireless sensor modules by combining radio frequency signals and Ultrasound; the results showed accurate position estimations with enhanced net-work flexibility [18]. However, traditional ultrasound positioning has some disadvantages including line-of-sight transmission, multipath, high cost and power consumption which may hinder the possible applications in complicated construction environments [19]. Various combinations of RFID and Zigbee-based sensor networks have also been applied for materials tracking and supply chain management [20,21]. RFID tags were used to identify various kinds of construction materials, and the ZigBee communication technology was used to wirelessly transfer this information. These studies confirmed that WSN can improve the wireless communication and network flexibility but their primary use was only data transmission, and not positioning.

The construction environment is characterized as a spatially expansive, object-cluttered, fast-changing, and harsh environment. The adoption of data acquisition technologies for progress tracking on construction sites would require simultaneous tracking of items under challenging conditions. These conditions characterized by the presence of moving resources and by metallic environments and extreme weather events, which could impact the communications, which largely depend on the surroundings [22]. Thus, the operational ability of a technology-based tracking solution must be enhanced to survive in such environment. Recent advances in computing and communication have caused a significant shift in wireless data acquisition research. However, its deployment in buildings construction sites is still challenging, due to poor signal propagation in indoor environment. Wireless network connectivity is limited indoors by physical obstacles and structural barriers such as walls, and by interference in the frequency spectrum. This research was motivated by the increasing need for understanding the behavior of various wireless networks in indoor construction environment.

3 Indoor RF Propagation

The electromagnetic theories define radio wave propagation in free space, however predicting radio wave propagation in complex jobsites is very difficult due to the effects of wave reflection and scattering. These effects lead to multiple waves traveling through varies paths, which is known as multipath propagation. The resultant interface can be constructive of destructive in respect to the received power [23].

Received signal strength (RSSI) is used by wireless networking community to measure signal strength. The signal path-loss model is used to convert the measured RSSI into distance between a transmitter and a receiver. However the RSSI value is highly dependent on the multipath and shadow fading interferences as shown in Figure 1. The signal propagation depends heavily on surrounding environment. The difference in signal propagation can be noticed through comparison presented in figure 1(a) & (b). If a mobile node communicates in corridor environment the link
characteristics will differ from an anechoic chamber environment with pronounced multipath fading. The anechoic chamber is a room designed to minimize reflections of radio waves and to shield an experiment from external interference [24].

Several localization techniques had been proposed in literature, but most of them are based on ideal radio signal propagation. However in real construction environment and in the presence of shadow fading and multipath problems, such localization techniques are not applicable and produce huge errors. In the following sections, real signal propagation scenarios are analyzed in order to provide solutions for WSN deployment in indoor construction environment.

4 Test Bed Setup

In order to experiment and investigate indoor propagation of different wireless networks, 21 experiments are conducted and 1752 data sets are recorded for more than 876 minutes (grand total of all experiments). The experiments took place in laboratory environment at Building Engineering department at Concordia University. Two areas were used for testing, a 25 meter long corridor for the straight line testing and 20 m x 20 m open area for the grid test. All these experiments are performed in different scenarios either in terms of number of nodes, distance between the nodes, line of sight and finally, in terms of topology i.e. straight-line/grid (Figure 2).

A Waspmote platform is used to build the mobile nodes for the experimentations, which includes a microcontroller operating at 14MHz, 128Kof ROM, 8K of RAM, a wireless transceiver interface socket, and a USB interface for device programming and logging. Each device operates on rechargeable batteries. Its wireless interface socket is compatible with different communication protocols (WLAN, Bluetooth, Zigbee and Synapse SNAP) and frequencies (2.4GHz, 868MHz, 900MHz) as shown in figure 3.

Four wireless technologies are used in the experiments, in particular, Wireless Local Area Networks (WLAN), Bluetooth, Zigbee and Synapse SNAP. Their technical details with respect to frequency, output power, range, sensitivity and cost are summarized in Table 1.

<table>
<thead>
<tr>
<th>Wireless Network</th>
<th>Bluetooth</th>
<th>Zigbee</th>
<th>WLAN</th>
<th>Synapse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware Module</td>
<td>Roving Network (RN-41)</td>
<td>Xbee 802.15.4</td>
<td>Roving Network (RN-171)</td>
<td>RF300</td>
</tr>
<tr>
<td>Freq (GHz)</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>915</td>
</tr>
<tr>
<td>Data Rate (Kbps)</td>
<td>3x1024</td>
<td>250</td>
<td>921</td>
<td>150</td>
</tr>
<tr>
<td>Power Range (dBm)</td>
<td>15</td>
<td>10</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Range (m)</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>250</td>
</tr>
<tr>
<td>Sensitivity (dBm)</td>
<td>-80</td>
<td>-92</td>
<td>-83</td>
<td>-99</td>
</tr>
<tr>
<td>Tx current (mA)</td>
<td>65</td>
<td>35</td>
<td>120</td>
<td>85</td>
</tr>
<tr>
<td>Rx current (mA)</td>
<td>35</td>
<td>50</td>
<td>38</td>
<td>18.5</td>
</tr>
<tr>
<td>Cost ($)</td>
<td>$24.95</td>
<td>$22.95</td>
<td>$29.95</td>
<td>$34.64</td>
</tr>
</tbody>
</table>

All experiments are summarized in table 2. The setup for straight line experiments is shown in Figure 4. The path is 20 m long, straight track with 20 waypoints with a distance of 1 m between two consecutive waypoints. Two stationary (transmitter) sensor nodes are placed next to the track at 0 m and 21 m. A mobile
unit (receiver) is placed at the 1 meter mark and recorded the RSSI from each of the transmitters for 5 mins. Then the mobile unit was advanced to the next waypoint. Each experiment is repeated for each of the four wireless networks (WLAN, Bluetooth, Zigbee and Synapse SNAP).

Table 2 Experiments Scenarios

<table>
<thead>
<tr>
<th>Exp. Nodes #</th>
<th>Dist. (m)</th>
<th>TX Level</th>
<th>LOS</th>
<th>Topology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1-20</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1-20</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1-20</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1-20</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2-0</td>
<td>Y</td>
<td>S.L.</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>1-20</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2-0</td>
<td>Y</td>
<td>S.L.</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>2-0</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2-0</td>
<td>Y</td>
<td>S.L.</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>2-0</td>
<td>Y</td>
<td>S.L.</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>2-0</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>1-20</td>
<td>N</td>
<td>S.L.</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>1-20</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>1-20</td>
<td>6</td>
<td>N</td>
</tr>
<tr>
<td>15</td>
<td>4X4</td>
<td>3</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>16</td>
<td>4X4</td>
<td>6</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>17</td>
<td>6X6</td>
<td>3</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>18</td>
<td>6X6</td>
<td>6</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>19</td>
<td>3X4</td>
<td>3</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>20</td>
<td>3X6</td>
<td>3</td>
<td>6</td>
<td>Y</td>
</tr>
<tr>
<td>21</td>
<td>4X6</td>
<td>3</td>
<td>6</td>
<td>Y</td>
</tr>
</tbody>
</table>

Figure 4. Experimental setup for evaluation of wireless networks in a straight line setting

Figure 5. Experimental setup for grid setting

The setup for grid setting experiments is shown in figure 5. The mobile nodes are set as transmitters, and the station unit is set as receiver. Then the setup was reversed, the mobile nodes where receivers while the station unit was transmitter. This grid test is used to explore the effect of multiple broadcasting on the same bandwidth and frequency, in order to understand the effect on the RSSI. The grid size changes from 3m x 3m to 6m x 6m, with a distance of 1 m between two consecutive nodes. One stationary sensor nodes is placed next to the grid at 1 m and center of the grid. Each experiment is repeated for each of the four wireless networks (WLAN, Bluetooth, Zigbee and Synapse SNAP).

RSSI of data packets received by the mobile node is measured using a program written using C++ and installed on the fixed node. In this program, a stationary node sends a PING packet to the mobile node and the mobile node responds with another data packet that contains the RSSI value with which the PING packet was received. The two stationary nodes send PING packets in a strict round-robin fashion to avoid packet collisions. Each node sends 3 packets per second, which results in 3 RSSI samples per second per link.

5 Results analysis

The received signal strength (RSSI) of a given wireless network is dynamic, which is affected by the periodic/random changes in the physical properties of the surrounding environment or even a group of people passing around the transmitter or receiver. To make an overall comparison between the RSSI measurements, the raw RSSI data from real-time measurements is filtered to remove the small oscillations using a moving average with a window size of 10 samples and an even
sample weight as shown in figure 7.

![Figure 7. Real-Time RSSI VS Filtered RSSI](image)

**5.1 Impact of base station distance**

The collected data presented in figure 8, clearly show that the received signal strength at each point is declining as expected. Even though, the declining rate is inconsistent with all wireless technologies. For example the node at 4 meters receives a weaker signal on the incoming packets with Bluetooth and WLAN than the node at 5 meters. Also the node at 7 meters receives a weaker signal on the incoming packets with Zigbee than the node at 8 meters. In all measurements Synapse hardware showed more consistency in returning RSSI values in declining order.

![Figure 8. RSSI measurement straight line formation](image)

A major source of error when measuring RSSI is due to multipath effects caused by objects in the environment. In the office environment, where the tests were performed, the radio environment is likely to change between every measurement point as the room contains quite many things that could cause multipath effects.

**5.2 Impact of multipath**

The multipath results for signal reflection from objects and such as walls, ceilings in indoor environment. The received signal arrives as multiple reflections or direct signal as shown in figure 9. The measured RSSI is directly affected in a constructive or destructive way with respect to the original transmitted signal.

![Figure 9. Multipath interference](image)

The multipath interference can have many forms such as reflection, scattering, refraction or diffraction. Signal Reflection occurs when the signal is reflected back towards the transmitter. Signal scattering creates multiple new signals after striking an object. Signal refraction occurs when the signal is bent while it is passing through an object and signal diffraction is the change in the signal direction when it passes around and object.

**5.3 Impact of Attenuation**

As a radio signal propagates from the transmitter to the receiver, the signal attenuation takes place. This is mainly due to the transmission medium properties. The Path loss describes this attenuation as a function of the wavelength of the carrier frequency and the distance between the transmitter and receiver. Path loss is derived from the Friis transmission equation (Eq. 1) and is defined as:

\[
Path Loss = 20 \log \left( \frac{\lambda}{d} \right) \text{ dB} \quad \text{Eq.1}
\]

Where \( r \) is the distance between the transmitter and receiver, and \( \lambda \) is the wavelength.

The signal attenuation is directly related to the frequency of the RF signal and the transmission medium materials type and density. The lower the Signal frequency the higher distance the signal will travel through air and through objects. Using frequencies below 900 MHz significantly improves the connectivity in indoor environments, as shown in figure 10.

Each material has a different attenuation coefficient, which is used to quantify the amount of signal strength reduction for each material type. Drywalls have a relatively low attenuation, about 2 db, while concrete and brick walls have higher attenuation levels. The
number of walls that a RF signal has to pass through also affects the signal strength and must be taken into consideration when designing indoor wireless system.

Figure 10. Attenuation interference

5.4 Impact of Antenna Orientation

The transmitter and receiver nodes antenna orientation can affect the radio signal strength due to that fact that the antenna radiation pattern is not uniform. In order to measure the impact of the antenna orientation impact a standard procedure is applied to measure the average RSSI at 24 different degrees with a fixed receiver node and a rotating transmitter node at a 1m distance. This test was performed using Wasp mote microcontroller and the output power of the radio signal was set to -10dBm in a relatively obstacle-free environment.

As shown in figure 11(a), the radiation pattern of the Synapse antenna is unsymmetrical and suffers for distortion with a difference in the measured RSSI of up to 10dBm. Figure 11(b) illustrates the coverage range, which is calculated based on the radiation pattern. Synapse has shown a wider and higher coverage range than the other three wireless technologies.

6 Conclusion

This paper described experimental study conducted to determine the effectiveness of wireless technologies for dynamic indoor tracking of construction operations. The challenges in RF-based localization technologies were analyzed by conducting a total of 21 experiments. Four wireless technologies were investigated, in particular, Wireless Local Area Networks (WLAN), Bluetooth, Zigbee and Synapse SNAP. The results showed that Synapse SNAP out-performed all other technologies. The findings of this study can be summarized as following:

- The RSSI fluctuation is a phenomenon which is caused by multipath interference and signal reflections. A weighted average filter can be used to elevate this problem and smooth the raw data.

Figure 11. Antenna orientation and range

- Heavy congested environment suffers from higher multipath interference, which can be reduced by continually calibrate the pass loss model parameters.

- The field measurements confirmed the Faiis equation that signal attenuation is directly proportional with the operating frequency. The RF 900 MHz frequency has better performance in indoor environment than 2.4 GHz.

- The antenna radiation pattern is not uniform and there for the antenna orientation impacts the received signal strength and reduce coverage range. It had been confirmed during these experiments that WSN hardware had electromagnetic effect, and by leaving a vertical distance between the hardware and the antenna, this effect was reduced.

- The commonly used empirical model for indoor propagation cannot be fixed, and it needs continuous
calibration to reflect the continually changing surrounding environment.

References


Controlling Slab Flatness Automatically using Laser Scanning and BIM

F. Bosché and E. Guenet
School of the Built Environment, Heriot-Watt University, UK
E-mail: f.n.bosche@hw.ac.uk

Abstract -
Developments of Terrestrial Laser Scanning (TLS) and Building Information Modeling (BIM) offer great opportunities to achieve a leap forward in the efficiency and completeness of dimensional control operations. This paper presents an approach that demonstrates the value of this integration for slab flatness control. The approach first employs the Scan-vs-BIM principle of [5] to segment TLS point clouds and match each point to the corresponding object in the BIM model. It then automatically applies the Straightedge technique to the TLS points associated with each floor slab, and concludes with regard to their compliance with given tolerances. The approach is tested using data from two real concrete slabs. Results validate the performance of the proposed system when compared with traditional measurement methods. Furthermore, a novel variation of the straightedge measurement technique is presented that enables more complete flatness controls for negligible additional processing time.

Keywords -
laser scanning; BIM; quality control; surface; slab; flatness; regularity

1 Introduction
Terrestrial Laser Scanning (TLS) is a modern technology that is revolutionizing surveying works. As highlighted in numerous previous research works (e.g. [1, 7]), TLS could provide surveyors with the means to conduct far more complete and reliable dimensional controls in manageable times. But, its use in practice remains limited essentially because of some concerns regarding the level of measurement accuracy it provides, and the time required for manual processing of the data to extract the dimensions of interest.

This paper presents a novel approach that integrates TLS and BIM to significantly automate the processing of TLS data, and hence overall control processes. The system automatically (1) identifies the TLS data corresponding to each floor in the 3D model, and (2) applies control procedures. The approach is demonstrated here in the case of slab flatness control, with the application of the Straightedge method, one of the most common standard flatness control procedures. The approach achieves results that compare favourably with those obtained using traditional measurement techniques. Furthermore, a novel variation of the straightedge measurement technique is presented that enables more complete flatness controls for negligible additional processing time.

The rest of this paper is organized as follows. Section 2 reviews existing methods for conducting floor regularity control, and particularly the Straightedge method. It then analyses how the integration of TLS and BIM can enable a leap forward in the efficiency and completeness of dimensional control operations. The proposed approach and implemented system are then presented in Sections 3 to 5. Results of the experiments conducted to test and validate the proposed system are reported and analysed in Section 6. Conclusions are finally drawn and recommendations for future work made in Section 7.

2 Background
2.1 Surface Flatness Quality/Compliance Control
Surface flatness, or surface regularity, is “the deviation in height of the surface [...] over short distances in a local area” [10]. The control of surface regularity can be done using different methods, such as: the Straightedge method [10, 2], the F-Numbers method [2, 3], the TR34 method [15] and the Waviness Index method [4]. In the following, we focus on the Straightedge method [10, 2] that is traditionally and commonly used.

In the Straightedge method, the surveyor lays a straightedge at different locations on the surface and measures the maximum deviation under it, preferably using a stainless steel slip gauge [10]. The deviation is then compared to a tolerance to validate or reject the level of flatness of the surface. A long straightedge (2m in Europe, 3m in the USA) is used to control global flatness, while a smaller ruler (0.2m in Europe, 0.3m in the USA) can be used to control local flatness. Control of global flatness enables the discovery of larger deformations, like bending; while
local flatness is measured to identify little gaps or bumps on the slab.

In the UK, the multi-part standard BS 8204 [10] provides global tolerances specifically for the surface regularity of direct finished base slabs or leveling screeds. In the USA, tolerances for concrete slab flatness are provided in ACI 117 [3]. Similarly to BS 8204, ACI 117 provides tolerances for 100% compliance – *i.e.* 100% of the straightedge deviations measurements must be below the given tolerance. However, in contrast with BS 8204, it also requires a second set of tighter tolerances be defined for 90% compliance – *i.e.* 90% of the straightedge measurements must be within the given tolerance.

Surprisingly, no British standard specifies where the straightedge should be positioned on a given surface. A note in BS 8204 only mentions that “the number of measurements required to check levels and surface regularity should be agreed between the parties concerned bearing in mind the standard required and the likely time and costs involved.” In the USA, ACI 117 suggests that straightedges should be placed randomly on the surface. It further specifies that at least one sample must be taken for every 100 ft² of floor area and that samples must be taken parallel, perpendicular, or at a 45° angle to the longest construction joint of the test area. It is however acknowledged that “there is no nationally accepted procedure for taking measurements or for establishing compliance of a test surface with this tolerance approach” [3]. The only more detailed method was found in the CSTB (France) “Avis technique 20/10-193*V1” [11] that suggests (but it is not a standard) the use of a square grid of lines spaced by 1m.

It is widely agreed that the Straightedge method is simple to understand, inexpensive and thus still widely used. However, it presents important deficiencies including:

- The difficulty in testing large areas of floors;
- The difficulty of randomly sampling floors; and
- The inability to reproduce testing results.

### 2.2 Terrestrial Laser Scanning (TLS) for Construction Compliance Control

With TLS, a laser scanner sweeps its entire surrounding space with laser light to acquire 3D data points with good accuracy, high density, and great speed. Point clouds provided by 3D laser scanners can be used directly for measurement and visualization, but can also be post-processed to extract underlying valuable information.

The potential of TLS for quality control has long been recognized, and [1] proposed a first formalization for integrating project 3D models and sensor systems (in particular TLS) for construction quality control. A first implementation of such a system has then been reported in [7]. The method uses what the authors later called the *Scan-vs-BIM* principle [17], where the TLS data is registered (*i.e.* aligned) in the coordinate system of the project 3D BIM model. This enables the system to automatically match TLS 3D data points to each BIM model object; and infer the recognition of those objects. In [7], the authors then demonstrate an approach for automatically quantifying positional deviations (*i.e.* deviations equivalent to rigid transformations, such as out-of-plumb deviations of columns). This approach cannot however assess local shape and surface irregularities (*i.e.* deviations equivalent to non-rigid local deformations), such as floor flatness.

Regarding the assessment of surface regularity, [14] have explored three algorithms for detecting surface flatness deviation. Their main algorithm works in 3 stages:

1. Apply Gaussian noise filtering to the point cloud;
2. Fit a plane against the overall point cloud; and
3. Calculate the distance between each point and the overall plane.

Two other variations of that algorithm are also considered. However, despite a detailed analysis of their performances, the methods presented in [14] do not enable the characterization of defects in ways comparable to current standards. As a result, it is difficult to assess the performance of using TLS in general, and the performance of their approaches in particular, for surface flatness control.

#### 2.3 BIM for Construction Compliance Control

The value of BIM models with regard to specifications and compliance control is at two levels. First, the integration of specifications within BIM models would enable reliable and efficient issue and management of construction project specifications. *NBS Create* [16], released by NBS in 2013, is a software tool that enables just that: the automated identification and management of the standards and specifications relevant to all components present within a given BIM model. The user then simply needs to specify the requirements identified by the system.

Secondly, design BIM models (with integrated specifications) can support more efficient and robust construction quality/compliance control. In [8], an approach is presented that uses a project 4D BIM model with integrated specifications and automatically generates for the surveyor the list of building components to be controlled along with the related specifications based on the current construction progress. Their vision further included (a) the automated generation of detailed survey plans given those requirements and the available survey equipment; (b) the automated identification of deviations by comparison of the design BIM model and as-built data captured by the survey equipment; and (c) the automated identification of defects by comparison of the deviations with the defined specifications. However, no approach was proposed and demonstrated for those later stages.
3 Contribution and System Overview

We propose an approach that integrates TLS and BIM models and conducts automated floor flatness control. The system assumes as input a BIM model augmented with specifications and a set of TLS scans acquired on site. It then uses the Scan-vs-BIM method of [5, 17] to align the TLS scans in the coordinate system of the BIM model, and match all TLS cloud points to the different 3D objects composing the BIM model. Finally, it automatically applies the Straightedge method to control the compliance of floors. The diagram in Figure 1 summarizes this process. The advantages of this overall approach are:

- **Integration/Automation:** the process is almost entirely automated; the only step potentially requiring user input is the alignment of the TLS scans with the BIM model. Furthermore, the results can be automatically linked to the BIM model, so that they can be easily shared with and reviewed by other stakeholders.

- **Compatibility with current standards:** the system applies a standard method for floor flatness specification and control, and is thus entirely compatible with current standards. An improvement of the Straightedge method is nonetheless proposed that takes advantage of the density of data available.

Section 4 quickly reviews the Scan-vs-BIM system used at the beginning of the process. Section 5 then describes our implementation of the Straightedge method.

4 Scan-vs-BIM system

The input of the proposed dimensional quality control system includes a 3D BIM model and a 3D point cloud (composed of one or more laser scans). The first step of the Scan-vs-BIM process [5, 17] consists in aligning the point cloud with the model. For this, we use the approach in [6] based on plane matches, but other approaches can be used. Then, each point of the point cloud is matched to a BIM model object (or none) using a combination of proximity and surface normal similarity metrics. This process segments the initial TLS point cloud into a set of sub-point clouds, one for each of the 3D model objects. The user can select any object (e.g. a floor) and visualize the points associated to it, e.g. colour-coded according to their deviations from the surface of the object.

5 Automated Straightedge Method for Flatness Control

We have digitally encoded the Straightedge method for floor flatness control, so that it can be applied automatically to any floor. In our implementation, the control procedure is divided in three steps: 1. Data pre-processing; 2. Generate Straightedges; 3. Associate TLS points to straightedges and calculate deviations and compliance. These three stages are detailed in the sequel.

Note that floors must be controlled by floor section, that is defined as a continuous surface delimited by the floor boundary and/or joints. Floors should thus first be divided into conforming test sections. In our implementation, we assume that the 3D model already contains appropriately divided floors.

5.1 Data Pre-Processing

In this section, two important pre-processing steps are described:

1. Identification of the set of TLS points from the floor’s top face. For this, we build on the fact that the Scan-vs-BIM system employed in [5] associates points to each triangular face defining the surface of the object. As a result, the points on the floor’s top face are easily identified as those associated to mesh faces with normal vectors pointing upwards.

2. Organization of the floor’s top face points in a 2D square array that will be used to conduct efficient, directed point searches. The orientation and extent of the array are determined using the two main directions of the floor (we use the horizontal directions of its bounding box) and a pre-defined array cell size, $d_{array}$ (we use $d_{array}=50$mm).

5.2 Generation of straightedges

The system generates straightedges by selecting pairs of TLS points on the floor that are spaced by the necessary distance (e.g. 2m). The literature review highlighted that current standards do not prescribe the pattern in which straightedges should be positioned on the slab. But, the
literature suggests that straightedges may be positioned randomly, or possibly along the lines of a square grid. In this research, these two (Random; Grid-Square) as well as a third pattern (Grid-Star) were investigated and are described in Sections 5.2.2 to 5.2.4. Before that, Section 5.2.1 discusses the method we use to validate the length and location of straightedges generated with either of the three methods above.

5.2.1 Validation of straightedges

Each generated straightedge must be validated against two criteria: length, and location.

The distance between the two points must correspond to the specified straightedge length $L$ (e.g. $L = 2m$ for global flatness control). However, selected TLS points may not be exactly distant by $L$. We thus introduce a tolerance factor $\epsilon$ on the distance between the two points, i.e. we accept straightedges with length $(1 \pm \epsilon) L$; we use $\epsilon = 2\%$.

Then, it must be ensured that each generated straightedge is entirely contained within the floor – i.e. it does not cross any of the boundary segments – and is not closer to its boundary than a pre-defined distance $d_{boundary}$. To check whether the straightedge intersects any of the boundary segments, we work in the 2D coordinate system of the floor’s top face, on which we project the straightedge’s extremity points, $s \rightarrow s'$ and $f \rightarrow f'$. We then employ the method described in [13]. To additionally check that no part of the straightedge is closer than $d_{boundary}$ to any of the boundary segments, we simply check that $\rho(s, f')$ and $10\text{cm}$ point increments in between them are not closer than $d_{boundary}$ to the boundary. In our experiments, we use $d_{boundary} = 40\text{cm}$.

5.2.2 The Random method

The Random method to generate straightedges simply consists in randomly selecting pairs of points from the point cloud associated to the floor’s top face. Each straightedge is then validated as described in Section 5.2.1. This process is iterated until a pre-defined number of straightedges has been obtained, e.g. 100 straightedges.

The laser scanning measurement process leads to a heterogeneous spread of points on the floor, with most points located near the scanner. To ensure that straightedges are homogeneously and widely spread around the floor, we use the homogeneous floor decomposition provided by the array data structure defined in Section 5.1. To generate each straightedge, a cell is first randomly selected from the array, and a TLS point is randomly selected from those contained in that cell as the first extremity of the straightedge, $s$. Then, the second extremity of the straightedge, $f$, is searched among all TLS points contained in the cells intersecting a circle centered on $s$ and with radius $L$. Figure 2b illustrates the result obtained.

5.2.3 The Grid-Square method

The Grid-Square method creates a 2D square grid with spacing parameter $L$ and then defines straightedges between all pairs of neighboring grid intersections. The orientation and size of the grid is determined using the main directions and dimensions of the floor’s top face.

Straightedges are generated between neighboring grid intersections as long as these have valid TLS point associated to them. For each grid intersection, a valid associated point is identified as the closest TLS point within a neighborhood defined by the radius $\rho$ (we use $\rho = 25\text{mm}$). If two valid neighboring grid intersections are found, we then check the validity of the straightedge connecting them, as described in Section 5.2.1. An example of straightedges extracted using the Grid-Square method is shown in Figure 2c.

The Grid-Square method does not really make use of the density of points provided by laser scanners and consequently leads to a partial assessment of floor flatness. The random method can more easily make use of the point density by simply increasing the number of straightedges to be generated. However, this process remains random and may require the generation of an unnecessary large number of straightedges. Another straightedge generation method is thus needed that would produce straightedges that altogether cover the floor completely (including in different directions), but that would achieve this without requiring an unnecessarily large number of straightedges to be generated. We propose one that we call Grid-Star.

5.2.4 Grid-Star method

This Grid-Star method uses a similar grid as the one used by the Grid-Square method. But, to ensure that straightedges are generated in all areas of the floor, the process is altered in two ways:

- Additional grid lines and intersections are created at the end of the measurable floor section, even if these are closer than $L$ to their neighbors.
- Instead of generating straightedges using neighboring grid intersection points only, we generate a number of straightedges with their first extremity defined at each grid intersection point and the second extremities located on a circle of radius $L$ around this first extremity. To ensure a homogeneous spread of straightedges around each grid intersection point, the second extremity points are searched at regular angular intervals, $\alpha$. 
Figure 2d illustrates the result obtained for $\alpha=10^\circ$.

(a) The point cloud of the floor in its original colour.

(b) Random generation of 100 straightedges using the array structure.

(c) Grid-Square method.

(d) Grid-Star with $\alpha=10^\circ$.

Figure 2: The three different straightedge generation method considered: Random (b); Grid-Square (c); Grid-Star (d).

5.3 Find points under a straightedge and calculate deviation

Once valid straightedges have been generated (using either of the three methods above), the next stage is to identify the points that are located under each straightedge, calculate the deviation for that straightedge and compare it to the tolerance.

Given a straightedge $r$, we construct a local 3D coordinate system $\mathcal{R} = (x; y; z)$ that uses its first extremity, $s$, as the origin and its direction, $u$, as the $x$ axis. The coordinate system is then entirely defined as follows:

$$x = u; \quad y = \frac{Z \times x}{||Z \times x||}; \quad z = x \times y$$

where $\mathcal{G} = (X; Y; Z)$ is the global coordinate system, and $\times$ is the vector product operator. Finally, we define the 3D rigid transformation $M_r = (R_r|T_r)$ from the global coordinate system to the local coordinate system of $r$ ($R_r$ being the rotation matrix and $T_r$ the translation vector of the transformation).

To find which TLS points are under a straightedge $r$, we express each point $p$ in $r$’s local coordinate system:

$$p^r = [x^r_p, y^r_p, z^r_p, 1]^T = M_r p.$$ 

Then, $p$ is considered to be “under the straightedge” if: $0 \leq x^r_p \leq L$ and $|y^r_p| \leq \rho$, where $\rho$ is used to define an acceptable neighborhood around the straightedge for points to be considered “under” it ($\rho=25mm$).

The calculation of the deviation of the floor under the straightedge requires the vertical coordinate $z$ of each point $p$ under it to be well estimated. To reduce the impact of measurement noise on the $z$ coordinates, we calculate these by averaging the values of all points in their neighborhoods; we use $\rho=25mm$ as the neighborhood radius. We denote by $\bar{p}$ the resulting point.

According to [9], each point’s deviation under the straightedge, $\delta^r_p$, should be measured along the global vertical axis and not perpendicularly to the straightedge. We thus calculate $\delta^r_p$ using the following formula:

$$\delta^r_p = \frac{z^r_p}{z}$$

with $z^r_p$ the $z$ coordinate of $p^r$, i.e. $\bar{p}$ expressed in $\mathcal{R}$.

The overall straightedge deviation, $\Delta^r$, is then calculated as:

$$\Delta^r = \max \{\{\delta^r_p\}\} - \min \{\{\delta^r_p\}\}$$

The floor is then within compliance, if none of the measured straightedge deviations, $\Delta^r$, exceeds the defined tolerances.

6 Experiments

The proposed system was tested and validated using two real concrete floors. The first is the floor slab $6.40m \times 6.70m$ of the Acoustic Laboratory (AL) of the School of the Built Environment at Heriot-Watt University. The second is a section $4.80m \times 8.10m$ of the concrete floor slab of the Drainage Laboratory (DL) in the same school. These slabs are both around 25 years old, thus with potential ageing defects.
The Grid-Square approach was applied to the two floors using both the proposed TLS-based system and the traditional manual control technique. This enabled a direct comparison of their results to validate the proposed TLS-based system. The other two approaches – Random and Grid-Star – were applied using only the proposed TLS-based system, and their results compared to each other and to the Grid-Square approach.

For the manual measurement, we have carefully drawn a 2m grid on the floors with a chalk line so that the grid intersections, and consequently the straightedges, match those automatically generated by our system. Measurements were then conducted using a 2m long straightedge and a precision steel rule.

For the TLS data collection, the AL being a small fully enclosed room, two scans had to be conducted to ensure data was acquired for the entire slab. For the DL slab, one scan was sufficient. All scans were acquired using a FARO Focus 3D [12]. Following data acquisition, a 3D BIM model of each room was created using Autodesk Revit. Then, the approach of [5] was used to register the laser scans with the 3D models, and match all TLS points to the different objects composing the 3D models of the rooms. This process resulted in ~6 million TLS points matched to the AL floor, and ~1 million points matched to the DL floor.

6.1 Grid-Square results

Figure 3 and Table 1 summarize the experimental results obtained for the Grid-Square approach in comparison with those obtained through manual measurements. To support a detailed comparison, the manual measurements were conducted with straightedges at the same locations as those generated by our system. The following conclusions can be drawn from these results:

- Using 4%, 10% or 25% of the initial point clouds did not have any major impact on the final results. This means that it is not necessary to conduct extremely dense scans, which can save time on the overall process (see further discussion on time performance below).
- Some differences between the deviations obtained using the manual and TLS-based approaches can be observed. But, these are generally small, and the similarity of results is confirmed by the statistical analysis results that reveal that (1) the average difference between the manual and TLS-based measurements of the deviation under a straightedge is 1mm or less; and (2) there is clearly no statistical difference between them.
- Not only are the differences between the manual and TLS-based approach small, but the maximum overall deviation (which is used to assess the overall floor compliance) is found by both approaches to be for the same straightedge.

![Figure 3: Quality performance of the proposed TLS-based system. Comparison of the individual straightedge deviations obtained by our system (using 4%, 10% and 25% of the initial scan) and by manual measurement; the straight-edges are sorted by manual measurement deviation.](image)

![Table 1: Quality performance of the proposed TLS-based system. Statistical analysis of the difference in straight-edge deviations obtained manually and using the proposed TLS-based approach (using 10% of the initial laser scans).](table)
complete and reliable results without impacting the overall flatness control duration.

<table>
<thead>
<tr>
<th>Process Stages</th>
<th>Acoustic Lab</th>
<th>Drainage Lab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scanning</td>
<td>2x30 min</td>
<td>30 min</td>
</tr>
<tr>
<td>Scan Pre-processing (Faro Scene)</td>
<td>20 min</td>
<td>10 min</td>
</tr>
<tr>
<td>Scan-vs-BIM</td>
<td>25 min</td>
<td>15 min</td>
</tr>
<tr>
<td>Flatness Control</td>
<td>&lt;1 min</td>
<td>&lt;1 min</td>
</tr>
<tr>
<td>(Grid-Square; global)</td>
<td>1 hr 50 min</td>
<td>1 hr</td>
</tr>
</tbody>
</table>

Table 2: Approximate durations recorded with the TLS-based approach (with 10% of the original scan data).

6.2 Random and Grid-Star results

Figure 4 and Table 3 summarize the results obtained with the Grid-Star and Random approaches. Note that to enable a fair comparison of these two approaches, the number of straightedges generated by the Random method was set to the number generated by the Grid-Star method – i.e. 230 straightedges for the AL slab, and 320 straightedges for the DL slab. These results highlight a couple of things:

- The two methods provide deviations measurements that widely cover the surface of the floor, but the Grid-Star method leads to a more homogeneous and complete coverage.
- Both methods achieve similar results with regard to maximum deviation. As expected, these deviations are larger than those obtained using the Grid-Square method because these two methods generate straightedges that cover more surface and are thus more likely to identify localized surface irregularities. In fact, all our experiments used a 100% global flatness tolerance of 10mm, and it can be seen that the Random and Grid-star approaches both identified an area of the Acoustic Laboratory floor that was non-compliant, and this area was missed by the Grid-Square method (see Figure 3).

Regarding time performance, the processing times using the Random and Grid-Star methods were both minimal: less than 15 seconds for the DL slab, and less than 2 minutes for the AL slab (this longer time is due to the larger number of straightedges and larger TLS point cloud associated to the floor). Therefore, the overall durations of the flatness control operation are the same as for the Grid-Square method (as detailed in Table 2). But in contrast, if the Random and Grid-Star methods were to be applied manually with the same number of straightedges as here, then the overall durations of the control operation would have been in the order of 35 hours for the AL slab and 23 hours for the DL slab. The proposed TLS-based system thus enables more complete and reliable flatness control in potentially significantly shorter times than traditional measurement methods.

<table>
<thead>
<tr>
<th>Slab</th>
<th>Stat.</th>
<th>G.-Sq.</th>
<th>R.</th>
<th>G.-St.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>Max.</td>
<td>7.6</td>
<td>11.3</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>4.4</td>
<td>4.7</td>
<td>4.5</td>
</tr>
<tr>
<td>DL</td>
<td>Max.</td>
<td>4.3</td>
<td>6.6</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>3.3</td>
<td>3.3</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Table 3: Results for the Random and Grid-Star approaches. Mean and maximum deviations (mm) reported by these two methods along with the Grid-Square method.

7 Conclusion

TLS and BIM technologies offer great opportunities to improve the completeness, reliability and efficiency of dimensional quality control operations. This paper presented an approach that integrates TLS and BIM technologies to significantly automate floor flatness control. The Straightedge control technique has been encoded for application to any floor section with matched TLS points. Three different straightedge generation methods have been considered: Random, Grid-Square and Grid-Star; the latter being a novel method that we proposed. The experimental results lead to the following conclusions:

- In terms of quality performance, the system compares favourably with the traditional manual measurement approach with regard to both individual straightedge deviation and overall floor compliance.
In terms of time performance, the system is able to conduct a large amount of straightedge deviation measurements in negligible times. This means that more complete, hence more reliable results can be obtained in significantly shorter times than if traditional manual measurement methods are used.

The Random and Grid-Star methods both showed similar performances, generating straightedges covering floor surfaces well. However, the Grid-Star appears better as its surface coverage is slightly better, more homogeneous. Furthermore, the Grid-Star method has the clear advantage of employing a predictable straightedge generation approach, which means that it could be easily re-applied to the data by any stakeholder to confirm the results.

These initial results are very promising but still require further validation using datasets from a wide range of surfaces/floors (including newly constructed) and including comparable results obtained using traditional manual measurements.

References


Chip-based Real-time Gesture Tracking for Construction Robot’s Guidance

Ying-Hao Yu⁎*, Chun-Hsien Yeh⁎, Tsu-Tian Lee⁎, Pei-Yin Chen⁎ and Yeu-Horng Shiau†

⁎Department of Computer Science and Information Engineering, National Cheng Kung University, Taiwan
⁎Department of Electrical Engineering, Chung Yuan Christian University, Taiwan
†Department of Electrical Engineering, National Yunlin University of Science and Technology, Taiwan
E-mail: yinhaun@gmail.com, ttlee@cycu.edu.tw, pychen@csie.ncku.edu.tw, shiauyh@yuntech.edu.tw

Abstract -
Mobile robots in automation construction have been designed for applications of craning, conveying, excavating, and floor polishing. These robots nowadays are equipped with various sensors to detect environmental information and finish tasks autonomously. When robot’s navigating path needs to be rescheduled, the supervisor of robot can duly interrupt system and then redefines a new route for robot. In addition to robot remote control by radio signals, using digital camera to receive instructions from supervisor’s gestures is also effective and can avoid the drawback of networked data routing. In this paper, we propose a gesture tracking system by simulating a traffic light baton to guide a differential drive robot in construction site. Here a real-time moving object detection first tracks supervisor’s waves (gestures) with digital camera. Next the system determines guiding direction and steering angles based on fuzzy logic. All of our designs are implemented in single FPGA chip for operating under rigor environments. The experimental results demonstrate that proposed gesture tracking system is accurate and promising for chip-based gesture guidance on construction robots in the future.

Keywords -
Gesture Tracking; Moving Detection; Fuzzy Logic; Supervisory Control; FPGA

1 Introduction

Tracking pre-defined paths is one of the most important capabilities of mobile robots. The paths of robots are generally constructed in various configuration of tags or environmental features so that robots can detect and compare with onboard virtual maps. These paths and maps have to be installed and recorded beforehand in order to guarantee that robot can run on expected paths [1-4]. However, due to the changeability of construction sites, path tracking algorithms are inapplicable with predefined paths. For example, an indoor floor cleaning robot will bump into obstacles once the house fittings are changed. Although the robot designed with heuristic algorithm can learn and update maps after every new bump [3], this kind of control scheme in construction sites is unworkable in that heavy and large robot will cause serious safety issues to workers and facilities.

Besides, the algorithm called simultaneous localization and mapping (SLAM), which can recognize indoor environmental features, might overcome aforementioned problem by replacing heuristic algorithm with digital camera. Recently, the SLAM can achieved real-time processing by using algorithms of simpler training, improved classifying, steady-state operations, and high-end computers [4], [5]. The problem is that SLAM working under extremely changeable environments still remains a challenge to be resolved. Any failure of tracking or computing delay from SLAM will cause a disaster by lost robots.

In the construction sites, mobile robots have been designed for various applications of craning, conveying, excavating, road and tunnel inspecting, serving, and tiling machine [6-11]. The navigating paths of robots are usually determined by humans’ volition. For instance, the path of a road inspecting robot could be randomly rescheduled depending on the obstacles, traffic conditions, and cracks found by the supervisor. Similarly, optimal manoeuvre and parking location of a mining haul truck is also dynamically determined by supervisor for relative position to excavator, turning space, loading of the truck, and terrain of mine. Consequently, a semi-automatic control system is assumed to be more suitable than a fully automatic system in construction sites.

A typical semi-automatic system in robotics is the supervisory control. This kind of system is composed of a supervisor and robot. The supervisor controls the performance and progress of assignments and duly interrupts and reschedules operations of robot. For the role of robot, it just needs to perform routine tasks with
pre-defined control modes [10]. Such control mechanism might look inferior to a complete automatic system, but a safe working environment with heavy robot, therefore, can be secured from accident.

In this paper, we simulate a real-time hardware-based gesture tracking system for guiding a differential drive robot. The gesture tracking design on robot can update navigating path accordingly by detecting supervisor’s gestures (waves). Comparing with traditional gesture recognitions, the palm tracking is practically replaced with coloured traffic light baton in construction sites. Such design avoids interferences of palm’s tracking, which will break down under a crowded or nonideal illuminated environments. In addition, the range of a wave is generally defined as an approximative value in real-life operations. This makes fuzzy logic a suitable choice in our system to interpret the scale of a waving range. Finally, by considering the rigor environments of construction sites, all of our designs were realized in a single field programmable gate array (FPGA) chip in order to conform to criteria of lower power, cost, and installation dimension.

This paper is arranged as the follows. In the Section 2, gestures (waves) tracking algorithms based on moving object detection will be represented. Resource usages of FPGA chip design and demonstration will be shown in the Section 3. A short discussion is arranged in the Section 4, and the conclusion of paper is drawn in the Section 5.

2 Gesture Tracking Algorithms

Before starting gesture tracking off, traffic light baton’s image from digital camera is first detected by our previous works of hardware-based real-time demosaicling and moving object detection [12], [13]. Once supervisor sweeps the traffic baton, moving image will be immediately marked by our moving object detection and then tracked by hardware-based gestures tracking algorithms as the follows.

2.1 Tracking a Traffic Light Baton

We assume only a traffic light baton with specific colour that will be detected by robot’s camera. The gravity point \( P \) of baton’s moving marks can be derived by designating a rectangular tracking frame on marks. Each image picture produces one gravity point, and every five gravity points can determine the waving direction of traffic light baton by linking up first and last points.

Besides, detected gravity points sometimes need to be examined on monitor amongst waves. An additional sorting problem of gravity points arises from the different sequences between detecting and displaying gravity points. For example, as shown in Figure 1, if \( P_0 \) has to be first shown on monitor, \( P_1 \) will appear in next picture and so on. Thus five gravity points will not appear on monitor in one picture unless a sorting mechanism is designed as the follows:

1. Load the gravity points \( P_0 \) to \( P_4 \) into the shift registers \( R_0 \) to \( R_4 \) in turn, where the \( P_0 \) is the first tracked gravity point and \( P_4 \) is the last one.
2. Examine gravity points’ coordinates \((x, y)\) on image sensor array with registers pairs \((R_4, R_3)\) and \((R_2, R_1)\). The point that will first show on monitor should change to the left-hand position in each register pair.
3. With the similar operation, examine the register pairs of \((R_3, R_2)\) and \((R_1, R_0)\) continually and swap data as step 2.
4. Repeat steps 2 and 3 until the order of all gravity points in registers corresponding with the displaying sequences of monitor. Thus the most significant bit of \( R_4 \) will be the first point showing on the monitor, and the last point showing on monitor will be the least significant bit \( R_0 \).
5. Display gravity points by reading out shift registers’ data from the most significant bit in order.

Figure 1. Possible distribution of gravity points on monitor

2.2 Confirmation of a Gesture

After determining a wave’s direction, the next step is to confirm a valid gesture (wave) of supervisor. In this process, the first three waves initialize a new tracking and then the system confirms gestures by every wave. This mechanism can be achieved by recording wave directions in registers, as shown in Figure 2. In Figure 2(a), the registers \( L \) and \( R \) respectively denotes the left or rightward wave. The content of both registers will be \((1,0)\) if traffic baton sweeps from the right- to left-hand side, and the \((0,1)\) condition is for reverse direction. The operations in vertical direction are similar to the horizontal direction.
Figure 2. Registers’ conditions for: leftward wave (a) and upward wave (b)

As shown in Figure 2(b), the register $U$ of vertical waves represents upward direction and $D$ is for reverse one.

Meanwhile, additional shift registers are necessary for counting the number of waves in different directions. The Table 1 shows an operation of two shift register groups ($LX_2$, $LX_1$, $LX_0$) and ($RX_2$, $RX_1$, $RX_0$) to record the leftward waves. The number of waving leftward was recorded in ($LX_2$, $LX_1$, $LX_0$) and waving rightward was in ($RX_2$, $RX_1$, $RX_0$). For the leftward waves, the registers ($LX_2$, $LX_1$, $LX_0$) were first filled with three waves after initialization and then ($RX_2$, $RX_1$, $RX_0$) were inhibited. Consequently, the system after initialization only allowed $LX_2$ to be continually cleaned and refilled. This mechanism is also similar to the waves in other directions.

The final process of tracking a gesture is to determine the initial point of a wave. As humans’ behaviour, the same gesture will be confirmed if the initial points of waves are similar. Once the location of initial point is out of a range, the record of initial point will be reset and a wave of reverse direction might be considered. For the case of leftward waves, the traffic light baton sweeps from the right- to left-hand side, and the initial point $CX$ will be determined by coordinate $x$ on image sensor and the registers $RX_0$ and $RX_1$ as:

$$CX = \begin{cases} \frac{x_{RX0} + x_{RX1}}{2}, & \text{if } x_{RX1} > 0 \\ x_{RX0}, & \text{if } x_{RX1} = 0 \end{cases}$$  

Operations of initial point for leftward waves are illustrated in Figure 3. A starting off area is defined by $CX$ and errors $E$. Three cases in this figure can be discussed for:

Case 1: The $x_{RX_0}$ is out of starting off area on the left-hand side. Tracking system with this case will still recognize a leftward wave, but the initial point $CX$ will be updated and gradually moved to the left-hand side.

Case 2: The $x_{RX_0}$ is within the range of starting off area. The leftward wave is confirmed with last $CX$.

Case 3: Latest $x_{RX_0}$ locates on the right-hand side of starting off area. System with this case will immediately clean all waves’ records in registers and then a new rightward wave might be confirmed depending on upcoming waves.

Table 1. Data variation of shift registers for leftward wave

<table>
<thead>
<tr>
<th></th>
<th>$LX_2$</th>
<th>$LX_1$</th>
<th>$LX_0$</th>
<th>$RX_2$</th>
<th>$RX_1$</th>
<th>$RX_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial state</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1st (return)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3rd</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4th</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5th</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6th</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7th</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3. Operations of starting off area for leftward waves
2.3 Guiding Robot with Fuzzy Logic

This subsection represents the determination of steering angles for a differential drive robot. The design difficulty arises from the magnification of camera lens that causes misjudgement with the same waving range at different locations. Consequently, guiding a robot by actual spatial scale on camera is impractical.

To resolve such problem, magnification of camera lens has to cooperate with fuzzy logic in proposed system. According to magnification of lens,

$$M = \frac{h_i}{h_o} = -\frac{d_i}{d_o}$$  \hspace{1cm} (2)

$h_o$: height of object  
$h_i$: height of projected image  
$d_o$: object distance  
$d_i$: image distance

the distance between traffic light baton and camera lens ($d_o$) can be derived from baton’s height ($h_o$), projected image height ($h_i$), and baton’s image distance ($d_i$). It can be seen that waving range $w_i \times d_o$ on monitor and $d_o$ have an inverse proportional relationship, which approximates a constant of $w_i \times d_o$ for the same wave at different locations and helps us to resolve aforementioned misjudgement problem.

Figure 4 depicts the gesture control scheme for a differential drive robot. The supervisor assumed guiding a robot with expected driving directions. Proposed system on robot first tracked supervisor’s gestures then derived steering angles by fuzzy logic. Here the inputs of fuzzy set are $w_i \times d_o \text{ cm}^2$ and the outputs of fuzzy logic are expected steering angles from supervisor. Steering angles from fuzzy logic will be translated into physical values by the driving unit of robot.

The fuzzy set of $w_i \times d_o$ is defined in the Figure 5. Here the triangular membership functions were adopted to denote “range zero (RZ)”, “range small (RS)”, “range medium (RM)”, and “range big (RB)” of $w_i \times d_o$ inputs. Triggered values of membership functions can then be expressed as:

$$\mu_{RZ}(x) = \begin{cases} 
\frac{440 - x}{440} & 0 \leq x \leq 440 \\
0 & x > 440 
\end{cases}$$ \hspace{1cm} (3)

$$\mu_{RS}(x) = \begin{cases} 
\frac{x}{440} & 0 \leq x \leq 440 \\
\frac{1320 - x}{880} & 440 < x \leq 1320 \\
0 & x > 1320 
\end{cases}$$ \hspace{1cm} (4)

$$\mu_{RM}(x) = \begin{cases} 
0 & x < 440 \text{ or } x > 2200 \\
\frac{x - 440}{880} & 440 \leq x \leq 1320 \\
\frac{2200 - x}{880} & 1320 < x \leq 2200 
\end{cases}$$ \hspace{1cm} (5)

$$\mu_{RB}(x) = \begin{cases} 
0 & x < 1320 \\
\frac{x - 1320}{880} & 1320 \leq x \leq 2200 \\
1 & x > 2200 
\end{cases}$$ \hspace{1cm} (6)
Figure 5. The fuzzy set of wave ranges

For the fuzzy rules, we defined $TZ$, $TS$, $TM$, and $TL$ for the turning angles of “turn zero”, “turn small”, “turn medium”, and “turn large” of the different drive robot. Each rule was respectively assigned with weight of 0, 1, 3, or 5, as shown in Table 2. Robot’s steering scales to driving unit were assumed from 0 to 5 to control robot’s two wheels in different speeds. Finally, as aforementioned discussion, the driving directions of robot were determined by the gravity points’ tracking as supervisor’s waves.

In the end, the algorithm of defuzzification was minimum inference engine [14] as:

$$\mu_l(x) = \begin{cases} 1, & \text{if } x = x' \in X \\ 0, & \text{other} \end{cases}$$

(7)

where $\mu_{l'}$ is a fuzzy singleton of input, and triggered rule $l$ and output $\mu_{l'}$ are,

$$\mu_{l'}(y) = \max_{l=1} \left[ \min \left( \mu_{l'}(x'), \mu_{l'}(y) \right) \right]$$

(8)

The actual output of steering angles was based on centre of gravity for singletons as [15]:

$$y^* = \frac{\sum_{i} \mu_{l'}(y_i) y_i}{\sum_{i} \mu_{l'}}$$

(9)

where $y^*$ represents the outputs of steering angles by fuzzy, $\mu_{l'}(y_i)$ denotes the triggered values of membership function, and $y_i$ is the weight of fuzzy rules.

<table>
<thead>
<tr>
<th>Inputs</th>
<th>RZ</th>
<th>RS</th>
<th>RM</th>
<th>RB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outputs</td>
<td>TZ</td>
<td>TS</td>
<td>TM</td>
<td>TL</td>
</tr>
<tr>
<td>Weights</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2. Fuzzy rules

<table>
<thead>
<tr>
<th>Designs</th>
<th>LEs</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total design</td>
<td>4,678</td>
<td>6.84</td>
</tr>
<tr>
<td>Full colour demosaicking</td>
<td>1,753</td>
<td>2.56</td>
</tr>
<tr>
<td>Moving detection</td>
<td>361</td>
<td>0.53</td>
</tr>
<tr>
<td>Gesture tracking</td>
<td>1,210</td>
<td>1.77</td>
</tr>
<tr>
<td>Fuzzy logic</td>
<td>1,354</td>
<td>1.98</td>
</tr>
</tbody>
</table>

Table 3. Resources usages of proposed gesture tracking system in single FPGA

3 Experimental Results

Proposed gesture tracking system is implemented in single Cyclone II 2C70 FPGA chip from Altera. Adopted digital camera module was set for 1280×1024 pixels resolution and pictures’ frame rate were 12 fps, which is suitable for moving detection with walking speed. Total hardware resources usages of logic elements (LEs) are shown in Table 3. It can be seen that we only consumed 4,678 (6.84%) out of 68,416 LEs. This design involves the full colour demosaicking (2.56%), moving object detection (0.53%), gesture tracking (1.77%), and fuzzy logic (1.98%).

Figure 6 shows the snapshots of proposed gesture tracking. A sweeping traffic baton with blue light was first detected and marked with white stripes, and every stripe was composed of 32 pixels. Here the colour detection was simply designed by thresholds for baton’s blue colour. Cluster of moving marks was real-time tracked by a rectangle on monitor, as shown in Figure 6(a). Based on the rectangular mark, gravity points of rectangles can be plotted as the crosses in the Figure 6(b). These crosses represented the trajectory of a wave. Similarly, as shown in Figures 6(c) and 6(d), traffic light baton’s trajectories in vertical and horizontal directions were also represented by solid lines on monitor, and the marks of starting off areas are shown in the Figures 6(e) and 6(f).
Finally, the errors of various waving ranges are listed in Table 4. In this test, the traffic light baton was vertically installed and moving horizontally on a rail in order to control the quality of waves. We tested different waving ranges for 10, 20, 30, 40, and 50cm while supervisor stood front camera from 1 to 3 meters away. Each test scenario collected 10 waves including initialization stage. It can be seen that errors of waves were efficiently controlled by a maximum of 15% and the most of these errors were ≤ 10%. Meantime, average errors were all under 10%. The accuracy of proposed system is sufficient to guide a robot by successively gestures.

4 Discussion

Using a traffic light baton to guide a robot is more practical than the palm tracking. Although guiding a robot by palm is natural, palm’s detecting rate is usually deteriorated by nonideal illumination such as shadows and reflections. Moreover, additional algorithms for face recognition might also be required to track supervisor under a crowded environment. In contrast to palm recognitions, guiding robot with a traffic light baton can be realized with a simpler colour detection by using thresholds. Proposed real-time supervisory control is not only working well in the dark environments but also mitigates the interference from reflection.

Besides, the adequate and stable moving marks on traffic light baton are important to proposed system. It can be seen that accuracy of wave range depends on the number of moving marks, which represent the actual dimension of the traffic light baton. Accordingly, with 1280×1024 pixels resolution of image, there is a lower bound of traffic baton’s height or length on monitor for 2cm. This limitation can be improved by a higher image resolution without increasing of computing delay [13].

Finally, due to the moving marks which are composed of stripes, slightly move at a further location will be thought as a static object by system. It leads invalid detection at first column by 3m away from camera and errors’ fluctuations in each row, see the Table 4. Here a better detecting resolution can be achieved by decreasing the length of moving marks or with higher image resolution.

5 Conclusions

In this paper, we have proposed a practical gesture tracking for robot’s guiding system, which is based on the real-time hardware chip designs. Essential gestures such as horizontal and vertical waves have been represented in paper. Our designs involve real-time demosaicking, moving detection, and colour detection by thresholds to detect the specific target, a traffic light baton. Mobile robot’s steering angles are determined by fuzzy logic for different waving ranges. Proposed system also enables supervisor to guide robot flexibly with different distances from camera but without significant errors. Comparing with the traditional gesture guiding
Table 4. Horizontal waving errors with different distances away from camera

<table>
<thead>
<tr>
<th>Waving ranges</th>
<th>10cm</th>
<th>20cm</th>
<th>30cm</th>
<th>40cm</th>
<th>50cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum errors measured at 1m</td>
<td>10%</td>
<td>11%</td>
<td>10.67%</td>
<td>10.5%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Maximum errors measured at 2m</td>
<td>6%</td>
<td>9%</td>
<td>6.67%</td>
<td>13%</td>
<td>12.4%</td>
</tr>
<tr>
<td>Maximum errors measured at 3m</td>
<td>N/A</td>
<td>15%</td>
<td>8.67%</td>
<td>9.5%</td>
<td>10.8%</td>
</tr>
<tr>
<td>Average errors measured at 1m</td>
<td>3.9%</td>
<td>6%</td>
<td>5.68%</td>
<td>6.65%</td>
<td>7.76%</td>
</tr>
<tr>
<td>Average errors measured at 2m</td>
<td>2.6%</td>
<td>3.95%</td>
<td>3.53%</td>
<td>9.85%</td>
<td>7.96%</td>
</tr>
<tr>
<td>Average errors measured at 3m</td>
<td>N/A</td>
<td>7.4%</td>
<td>3.97%</td>
<td>5.95%</td>
<td>9.6%</td>
</tr>
</tbody>
</table>

Methodologies by using palm recognition, proposed system has the advantages of low-cost, real-time processing, miniature installation dimension, and a higher detection rate by using traffic light baton. Experimental results in this paper have demonstrated our gesture tracking system is promising for the supervisory control. The future work will focus on the implementation of gesture guiding system on mobile construction robots.

Acknowledgements
Supports from Ministry of Science and Technology, funded by the government of Taiwan under Grant NSC 99-2221-E-027-057-MY3, and Mr. Sheng-Jie Lin are gratefully acknowledged.

References


Comparative Experimental Evaluation of Dust Sensors for Environmental Monitoring on Construction Sites

A. Carbonari\textsuperscript{a}, G. Fava\textsuperscript{b} and B. Naticchia\textsuperscript{b}

\textsuperscript{a}Department of Civil, Building Engineering and Architecture, Polytechnic University of Marche, Ancona, Italy
\textsuperscript{b}Department of Materials, Environmental Sciences and Urban Planning, Polytechnic University of Marche, Ancona, Italy

E-mail: alessandro.carbonari@univpm.it, g.fava@univpm.it, b.naticchia@univpm.it

Abstract - Fine particle emissions ($\text{PM}_{10}$) from demolition and construction activities are now recognized as significant causes of pollution. So they can cause health hazards both to workers and to people living and working outside the site's boundary in the local neighbourhood. Although admissible concentration values are defined by national legislation and regulations worldwide, there is no permanent monitoring system in place, yet. Hence, in this paper one step pertaining to the development of a wireless pervasive and real-time monitoring system of dust concentration will be presented. In particular, it will focus on the types of sensors, integrated in the communication network, which can be used to the purpose. In particular, we will compare the performances of two different dust sensors. On principle, they are all suitable for wireless monitoring, but not all the sensors have the same sensitivity to changeable particles diameter sizes and types of dust. So a dedicated laboratory campaign was carried out, in order to compare their performances with the ones of a reference instrument (used as a benchmark). The comparisons show that there is a good agreement between the plots of the wireless real-time tracking system and the benchmark. However, the reliability of the sensors to detect events (e.g. sudden dust concentration variation) and particles with non-uniform particle distribution is different according to the kind of dust sensor which is integrated in the wireless system.

Keywords - Dust tracking; wireless sensor network; health and safety.

1 Introduction

Monitoring fine particle emissions ($\text{PM}_{10}$) is needed to face one of the main causes of pollution and health problems [1]. Workers in construction sites and people living in the local neighbourhood can be strongly affected by emissions from demolition and construction activities. Then, that fine dust can easily be carried in the air and penetrate further in the airways, causing well recognized health-problems.

In the construction field, $\text{PM}_{10}$ is originated by several concurrent factors, determined on one hand by site layout and organisation and, on the other, by the type of demolition or construction activities in progress. Dust and mud from roads and haulage routes on the site can become airborne through the movement of vehicles. Vehicles and some plants also generate engine exhaust emissions. In addition, the handling and storage of fine, powdery and dry materials has the potential for making the dust airborne. Additional contributions to dust generation comes from a number of fabrication processes, among which we cite cutting, grinding, sand-blasting, drilling and disk cutting.

The legislation relating to health and the environment mostly defines admissible ambient dust concentration as a running 24-hour mean, e.g. both the World Health Organisation [2] and the UK Air Quality Strategy, under Part IV of the Environmental Act [3], define a target average $\text{PM}_{10}$ value equal to 50 $\mu$g/m$^3$

In the USA, the Clean Air Amendment dated 1990 – incorporated into Title 42, Chapter 85 of the United States Code - regulates emissions from any facility into the air. The National Ambient Air Quality Standards (NAAQS) specify dust admissible in the outdoor air as a twenty-four hour average $\text{PM}_{10}$. It must not exceed 150 $\mu$g/m$^3$, whereas its annual average must not exceed 50 $\mu$g/m$^3$ [4]. Such legislation is also present in other countries: e.g. ambient air quality in Italy is ruled by Legislative Decree no. 152/2006. Annex V of this decree concerns outdoor air quality and limitation of outdoor air inclusions. It states that control measures of dust levels in outdoor air must be put in place any time dust generating activities are undergoing. The nature of such control activities shall be adequate to the hazards determined by the type and amount of dust, weather conditions and protection needed by the surroundings. So there is no pre-determined limit, whose assessment is, in fact, in charge of the person leading the dust
generating tasks. Anyway, even in Italy the reference level of outdoor air quality is usually taken as the value suggested by WHO.

On the other hand, regulations relating to occupational exposure define limit values in terms of an 8 hr Time Weighted Average (TWA) of total inhalable dust, whose thresholds differ slightly depending on the country.

Recommended control measures are those which reduce dust generated at its source [1]. Pre-project risk assessment is able to identify risks and mitigation actions, but cannot quantify their real effectiveness, which is strongly dependent on operational approaches, unless monitoring techniques are used. So far two main monitoring methods have been employed, according to relevant regulations [5]:
- pumped samplers that collect particles on a filter for later weighing or chemical analyses (generally held on site for some days or weeks), which is known as the gravimetric approach;
- continuous sampling instruments to provide fast online plots of dust concentration (e.g. light scattering), thanks to the use of optical instruments.

The work presented in our paper has the purpose of contributing to the second approach, by means of an advanced and non-invasive wireless sensor network, which is able to provide pervasive and continuous monitoring of the presence of dust on site. This new system was conceived so as to complement current monitoring techniques, by means of a platform which can be kept continuously on throughout the whole construction process.

In fact, the presence of this sensor set, capable of tracking dust concentration in real-time, is critical for obtaining a reliable picture of spatial dust distribution. Thus, once cheap, easy to deploy and pervasive sensor networks are available, they should determine several benefits, such as:
- triggering of warnings, when dust values exceed pre-determined thresholds;
- reduction in the number of environmental offences and hence in prosecutions by local authorities, which is a relevant financial burden for builders [6];
- reduction in the site engineer’s workload;
- reduction in the health damage to the workers.

In Section 2 a description of the pervasive network is presented. Section 3 will report on laboratory calibration of the sensors and Section 4 will describe on field experiments. Finally Section 5 will report our Conclusions.

2 Pervasive dust monitoring platform

This paper suggests setting up a real-time monitoring system which is capable of crosschecking the position of workers and the estimated dust concentrations over the site. Such a system would provide at least a couple of significant services in construction sites:
- capability of signaling in real-time the overcoming of any predetermined threshold values;
- gradual implementation of a database containing the cumulative value of the amount of PM concentration to which workers have been exposed, integrated over several possible meaningful time windows.

In the schematic shown in Fig. 1, every worker is supposed to be tracked using one of the available position tracking systems developed for construction sites; then the new sensors described in this paper are deployed over the site and programmed so as to give back in real time the dust concentration values at their known positions (in order to work like that, even they need to be equipped with a wireless tag device for being automatically located).

So a preliminary network prototype to monitor the concentration of particulate matter - PM$_{10}$ - was developed and tested in the machine laboratory of the DICEA Department and of the SIMAU Department at Università Politecnica delle Marche (Ancona, Italy).

2.1 The communication platform

The wireless communication network is based on the SmartNetwork Platform, an ultra-low power wireless technology (manufactured by Smart Space Solutions srl). It is capable of tracking large areas with a wireless mesh network architecture, made up of battery operated devices. The SmartNetwork includes three levels of devices: one or more coordinators, routers and sensors.

Routers forward messages across the network devices, they support complex, self-configuring and self-healing mesh networks with as many as 65,000 nodes. Routers provide network area coverage, they dynamically route around obstacles, and provide backup routes in case of network congestion or device failure.

Interaction among the three levels of devices makes the overall communication possible (Fig. 2). One or more PAN coordinators are used to initiate network
formation; routers are responsible for performing multi-hop routing of messages; end devices (also called “reduced function devices-RFD”) act as multi-purpose sensors and do not have routing capabilities.

The hallmark of SmartNetwork components is that they use a special hardware and firmware architecture capable of sensing an RF wakeup impulse with extremely low power consumption (as low as 0.05 mW), hence they can be battery powered. Once the appropriate nodes of the network are awakened, the transmission is performed through the network using the primary radio system which returns to sleep mode when no data need to be transferred. This technology allows a long time-span between two consecutive battery replacements, usually in the order of a few years [7].

2.2 Types of sensors which were compared during our trials

Our market survey gave back two main manufacturers of dust sensors: Shinyei and Sharp. Both products are based on the same operation principle: a light beam is emitted into a measurement chamber; when dust is present, the light is refracted by particles and the amount of scattered light is detected. One unique feature of the first set of sensors is that Shinyei ones use a heating resistor to create an updraft, hence it is active. On the other hand, the Sharp GP2Y1010 optical dust sensor is mostly used in air quality equipment, such as air purifiers, it has no embedded heater and it works as a passive dust sampler.

2.2.1 The active dust sampler

The active dust sampler was embedded in a package split into two chambers (Fig. 3-a). The left sided chamber including a micro-processor for processing data and sending them to the communication system, besides a lithium battery, made necessary because of the presence of a heater. The heater was placed at the bottom of the right sided chamber along with the probe, because after air was made flow upwards, the probe could collect records of dust concentration.

Technically, this sensor creates a Digital Lo Pulse output, whose pulse occupancy time is in proportion to PM concentration (Tab. 1). When dust is not present, then the Hi Pulse output is given as a results (Fig. 3-b). It means that the Lo Pulse occupancy time it detects is expressed as a percentage over the total measurement time window. Hence, every record is the result of the measurements carried out over that time span. Considering that the time for stabilization required by such sensor is about 60 s after power is turned on, we programmed it so that 60 s were used for stabilization and 30 s for measuring. Time for stabilization is due mainly to the presence of the heater, which is expected to trigger an upward air flow. As a consequence, one record every 90 s was provided by this senor.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply Voltage</td>
<td>V</td>
<td>DC 5V</td>
</tr>
<tr>
<td>Operating Temperature</td>
<td>°C</td>
<td>0 - 45</td>
</tr>
<tr>
<td>Hi Pulse Voltage (no dust)</td>
<td>V</td>
<td>&gt; 4.5</td>
</tr>
<tr>
<td>Lo Pulse Voltage</td>
<td>V</td>
<td>&lt; 0.7</td>
</tr>
<tr>
<td>Consumption current</td>
<td>mA</td>
<td>90</td>
</tr>
</tbody>
</table>

Tab. 1 Main parameters of the Shinyei PPD42NJ dust sensor.

2.2.2 The passive dust sampler

The Sharp dust sensor (Tab. 2) is equipped with a hole in the middle (i.e. measurement chamber), across which air is free to flow and dust concentration is measured by means of the light scattering principle (Fig. 4).
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply Voltage</td>
<td>V</td>
<td>-0.3 - +7</td>
</tr>
<tr>
<td>Operating Temperature</td>
<td>°C</td>
<td>-10 - +65</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>V/(0.1mg/m³)</td>
<td>0.5</td>
</tr>
<tr>
<td>Output voltage at no</td>
<td>V</td>
<td>0.9</td>
</tr>
<tr>
<td>Consumption current</td>
<td>mA</td>
<td>11</td>
</tr>
</tbody>
</table>

Technically, the Sharp dust sensor’s measurement chamber causes a pulse whose voltage is proportional to the number of particles which caused the light beam to be diffracted ($V_{s,i}$) [8]. The microprocessor works out the mean ($V_s$) based on the last 16 records. The time step at which pulses were generated was set equals to 10 s, determining about 6 records per minutes.

Fig. 4. Schematics of the passive sensor’s components.

3 Laboratory tests

Laboratory tests under controlled conditions were performed to calibrate the two types of sensors described above. Both were calibrated against the measurements of a highly accurate sensor, which was used as a benchmark.

3.1 Measurement setup

The two sensors’ characterisation and calibration was performed in the concrete laboratory of the Simau Department at the Università Politecnica delle Marche, under controlled conditions in order to avoid bias by unknown factors. A dust ventilation duct was setup in order to place the sensors and control the physical variables during testing (Fig. 5). The duct was made of a 10m long by 0.7 m wide by 0.7 m high insulated aluminium coated casing (AC), connected to a vertical substructure (VS) with a grid (at the intersection between the two) for homogenising the inlet generation.

Sea salt dust generation (DG) was provided by means of an ultrasonic fog generator placed at the inlet of the dust ventilation duct.

Fig. 5 – Deployment of the measuring setup inside and outside the ventilation duct (a); zoomed view of the instruments placed on the duct (b) and in the measurements cross section (c).

The ultrasonic fog generator uses ultrasonic technology to produce a fog composed of $10\mu m$ and less
sized water particles which gives smaller solid particle diameter as determined by eqs. (1) and (1) [9]:

\[
D_s = 0.34 \left( \frac{\sigma \delta}{\rho \nu} \right)^{\frac{1}{3}} \\
D_p = D_s \left( \frac{C_s}{\nu_s} 10^4 \right)^{\frac{1}{3}}
\]

where \(D_s\) is the fog drop diameter, \(D_p\) is the solid particle diameter, \(\delta\) is water density, \(\rho\) is surface tension (dine/cm²), \(\nu\) is the ultrasonic frequency (MHz), \(C_s\) is the solute concentration (g/L) and \(\nu_s\) is the solute density (g/L).

A filter along the duct guaranteed diffusion of the generated dust. At one end of the duct a fan (FN) was used to generate controlled and uniform air flow along the duct. The passive sensors (PS) and the active sensor (AS) were placed downstream inside the duct aside the benchmark (Grimm) laser photometer (LP). The passive and active sensors sent data to a couple of routers (RO1 and RO2) located on the top outer surface of the ventilation duct. These routers forwarded data from both sensors to one coordinator inside the laboratory, connected to a laptop at a short distance from the duct.

The benchmark instrument measured PM \(_{10}\) through an active technique. It used an optical particle counter Grimm 1.108, which is a portable laser photometer (LP) with a constant volume flow of 0.6 l/min and a digital display for real-time measurement. Finally, a removable 47 mm PTFE filter was incorporated inside the instrument in order to collect all the measured dust, so that, at any later time, an appropriate density verification/correction would be possible. This filter is in accordance with many national and ISO standards. The air flow rate inside the dust ventilation duct was generated by the rotating fan (FN) and its amount estimated by means of a multi-probe thermo-anemometer VT200 manufactured by Kimo Instruments (KA), equipped with a rotating vane anemometer with diameter 100 mm (RV), whose measurement range spans from 0.25 to 3 m/s and whose accuracy is 3% on the measurement and maximum accuracy 0.01 m/s. The passive and active dust sensors and the Grimm sampler were all placed on the same section of the duct, thus making them suction air at the same point along its path.

During the trial the dust generator (generating dust with diameter between 0.23 and 10 µm) was kept on at a constant rate for almost two hours: from 12:00 pm until 1:53 pm, while the fan’s rate was varied during the same time interval. The air speed was varied from the top speed 1.70 m/s down to the slowest value of 0.60 m/s and then increased again, as shown in Tab. 3. In this way the measurement capabilities of the two sensors were given as a function of external disturbances, e.g. wind when measurements take place either outdoors or in ventilated rooms.

\[ D_s = 0.34 \left( \frac{\sigma \delta}{\rho \nu} \right)^{\frac{1}{3}}, \quad D_p = D_s \left( \frac{C_s}{\nu_s} 10^4 \right)^{\frac{1}{3}} \]

Tab. 3 Speed of air flowing in the duct during the laboratory trial.

<table>
<thead>
<tr>
<th>Air speed (m/s)</th>
<th>Start time (hh:mm)</th>
<th>End time (hh:mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.70</td>
<td>12:00</td>
<td>12:15</td>
</tr>
<tr>
<td>1.00</td>
<td>12:15</td>
<td>12:34</td>
</tr>
<tr>
<td>0.60</td>
<td>12:34</td>
<td>1:10</td>
</tr>
<tr>
<td>1.00</td>
<td>1:10</td>
<td>1:33</td>
</tr>
<tr>
<td>1.70</td>
<td>1:33</td>
<td>1:53</td>
</tr>
</tbody>
</table>

3.2 Tests and discussion on the results

The benchmark recorded one measurement every minute, while the Sharp and Shinyei samplers worked differently. The Sharp sensors recorded one measurement every 10 s. Subsequently, post-processing worked out the average values per minute. The Shinyei sensor recorded one sample every 90 seconds, hence those data were resampled with a Matlab™ function to work out a one minute step database, when necessary. Thus, measurements from the three sensors where given the same time scale on the x-axis and the results are pictured in Fig. 6, each sensor keeping its own units.

Fig. 6 – Measurements recorded by the benchmark (a), the passive (i.e. Sharp) sampler (b) and the active (i.e. Shinyei) sampler (c).
This figure clearly shows that a qualitative agreement among the three sensors was had. As already stated in sub-section 3.1, the varying concentration sensed by the sensors was due to the variation of the air speed flowing across the ventilation duct. Although the sensitivity of the benchmark is better than the sensitivity of the other two sensors, all of them follow similar trends. But a more accurate comparison followed from the estimation of conversion factors between the two experimental dust samplers and the benchmark. The need for these factors arise from the different units that each of the sensors give as outputs. They were worked out as the ratio between the average dust concentration measured by the benchmark during the tests and the average value sensed by the passive sensors over the same time lag.

<table>
<thead>
<tr>
<th>Wind speed (m/s)</th>
<th>Passive sampler</th>
<th>Active sampler</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.60</td>
<td>314.4</td>
<td>291.1</td>
</tr>
<tr>
<td>1.00</td>
<td>320.2</td>
<td>192.0</td>
</tr>
<tr>
<td>1.70</td>
<td>276.2</td>
<td>169.2</td>
</tr>
</tbody>
</table>

Three conversion factors per each sensor were worked out, according to air speed across the ventilation duct. In other words, once wind speed are tracked and conversion factors are known, a monitoring system embedding these sensors should be able to provide a real-time rough estimation of PM$_{10}$ concentration in construction sites. Tab. 4 summarizes those factors for the combination of cases under consideration. To be noticed that the correction factors of both sensors follow different trends, in that the active sensor’s factors are lower at high air speed and higher at low air speed; on the contrary, not a similar trend is found in the case of the passive sensors. Comparing the two experimental sensors, the correction factors of the active sensor are always lower than the correction factors of the passive sensor. This suggests that the active sensor is more sensitive than the passive one, under controlled conditions.

**Fig. 7** – Comparison between the benchmark and the passive sampler (a) and active sampler (b).

**Fig. 8** – The wood cutting machine generating dust (a); view of the sensor position in the first trial (b) and in the second trial (c).

Fig. 7 shows what happens when the measurements recorded by the samplers at 0.60 m/s are multiplied by each CF, so that a direct comparison between the benchmark and the estimations provided by the
samplers can be provided. The passive sampler overestimated the intensity of the peak, but was quite accurate outside this peak’s window. The active sensor estimated better that peak but its deviations from the benchmark is bigger out of the peak’s window. Anyway, both sensors performed quite well over the whole controlled conditions tests.

4 Additional on-field experiments

A real-world scenario was offered by a company running a ship yard located in the county of Ancona (Italy).

4.1 Trials

The first trial started at 9:10 am and run until 10:20 am. The same benchmark instrument described in sub-section 3.1 was placed far from the machine but in the same room, and the two passive and active sensors were placed close by, as depicted in Fig. 8-b. The measurements collected during this time window were resampled with a 1 min wide step, same as what done with the laboratory experiments, so as to be able to compare both prototypical sensors with the benchmark, as shown in Fig. 9.

Fig. 9 – Comparison between the benchmark and the estimation provided by the passive (a) and active (b) dust samplers in the first on-field trial.

The second trial started at 12:30 pm and run until 1:30 pm. All the three sensors were placed besides the wood cutting machine, at the same height of the working desk, as shown in Fig. 8-c. Again, the measurements collected during this time window were resampled with a 1 min wide step, so as to be able to compare both prototypical sensors with the benchmark, as shown in Fig. 10.

Fig. 10 – Comparison between the benchmark and the estimation provided by the passive (a) and active (b) samplers in the second on-field trial.

4.2 Discussion

The results show that the scenarios considered in this case are considerably different from the controlled conditions created in the laboratory. In this case, just qualitative plots have been compared, because conversion factors for wood dust were not available. In fact, the conversion factors estimated from laboratory trials in section 3.2 are valid for salts. However, the qualitative comparison presented in this paper is good enough to support preliminary comments.

The first objective of this on-field experiments was to check whether the prototypical sensors could act as good environmental dust sensors. From the comparisons depicted in Fig. 9, it can be noticed that there is no correlation between the measurements of the two sensors and the one of the benchmark. There is a high likelihood that the different plots are due to the low dust concentration at such a distance from the source. That concentration was sensed by the benchmark but not by the sensors. Another possible reason is that after migration from its source, dust spread unevenly around the room, so the sensed values are markedly different. But the first circumstance looks more reasonable.

Once the sensors have been approached to the source, the passive sensor gives back quite reliable results, and the peak detected by the benchmark is present also in the plot of the passive sensors (Fig. 10). Oddly enough, the active sensor, which gave back good results under controlled conditions, did not work well in this case. This could be due to the way such sensor is driven to take measurements, which was described in sub-section 2.2.1. Firstly, it is turned on for 30s between two sets of measurements. Secondly, each measurements takes 60 s, because it is weighed over a certain air flow rate, which is moved by a heater. So this policy could have hampered the possibility to detect sudden dust concentration rise, such as the one detected.
by the passive sensors, which was instead programmed to collect one record every 10 s. As a consequence, further tests must be done, in order to check whether this phenomenon is recurrent and whether there is room for improvement, e.g. reducing measurement time lags.

5 Conclusions

Our comparison between the passive sampler and the active sampler showed that both are able to sense the particulate matter typically found in work places like construction sites. To that end, we assembled and tested an untethered monitoring setup, made of wireless devices based on the Zigbee™ communication protocol and mounting two types of dust sensors: one Sharp dust monitoring sensor and one Shinyei dust monitoring sensor.

Such a platform was set up because it would allow pervasive monitoring of mobile sources of dust. In fact, this is not possible through the use of current technologies which are mainly devoted to point measurements and dust sampling. For this reason, this new setup might be considered as a complementary system to some more accurate measurement approaches which are currently used, and it was specifically designed to solve the particular challenges posed by construction sites.

The experimental results showed that the reliability of the two sensors is strictly related to the context they are working in. In particular, the active sampler performed better when working under controlled conditions in the machine laboratory. But it was not responsive enough to detect peaks when immersed in a real work place. On the other hand, the passive sampler was more responsive, but less accurate when tested under controlled conditions. So our opinion is that more tests must be performed in real work places, in order to check further the behaviour of the two sensors. Also, a deeper analysis on the possibility to reduce the time for stabilization required by the active sensor should be performed. In fact, its scarce responsiveness might be determined by the long time required to get one record, which could flatten peaks. Additional considerations should be done in terms of cumulative readings over longer time windows.

6 Acknowledgments

The authors wish to express their gratitude to Eng. Massimo Monti and Mr. Roberto Mancini, who kindly supported the tests and analyses presented in this paper.

References

Extended Range Guidance System for the Teleoperation of Microtunnelling Machines

A. Jardón, S. Martínez, J. G. Victores, C. Balaguer
Roboticslab-University Carlos III of Madrid, Spain
E-mail: {ajardon, scasa, jcgvicto, balaguer}@ing.uc3m.es

Abstract -
Microtunneling is a trenchless construction method that is highly adequate to install pipelines beneath roads, rail roads, dams, harbours and environmentally sensitive areas. Microtunneling could be understood as a remotely-controlled, guidance operation of a MicroTunneling Boring Machine (MTBM), where references are provided by a human operator from the surface. A review of current technologies and main working principles of traditional guidance systems used to determine the position and orientation of the drill head of the MTBM during tunneling will be presented. Practical limitations and drawbacks will be discussed. A special attention will be given to those systems based on a laser-generated reference. In this paper a new video target system developed to improve the current performance of guidance systems will be presented. As it will be detailed, improving target processing of laser’ incidence allow us to extend the minimum range to achieve a distance over 400 m, without having to displace the total reference stations which are guiding the path during tunnel execution. The new target sensing principle, the implementation approach, the image processing and pose estimation algorithms will be discussed. Additionally, some preliminary results of the prototype in its current testing phase in real scenarios, gathered in parallel with commercial units, and its comparison will be presented.

Keywords -
Microtunneling machine, guidance systems, graphical interface, teleoperation.

1 Introduction
In order to successfully perform the teleoperated micro-tunneling machine drilling without excessive out-of-tolerance alignment errors and without incurring cost overruns, the performance of the guidance system is crucial [1, 2]. TBM’s shield is articulated and can be guided by the orientation cylinders placed between the first and second shield. Traditional guiding systems are based on a reference laser that marks a straight line to determine the position and orientation of the drilling head during tunnel construction. Also, this target requires complex oriented prisms systems for the estimation of position and orientation of the machine. All of this leads to delicate equipment, with complex handling and installation and which leads to higher construction costs.

The guiding system which will be described allows knowing the orientation and positions of the tunneling machine based on the capture of the incidence of a green laser on a custom-made video target. With this, the working range of the guiding system increases, avoiding having to relocate it, minimizing the impact in time and therefore, costs when drilling with it.

1.1 A brief review of existing guidance systems
The importance for the final results of the excavation pushes the ongoing improvement of target systems. When exploring Patent databases, several guiding methods for tunneling machines can be found. In [3] a method is designed based on a target with two detectors for the beam of light. A first detector is a semi-translucent mirror that deviates part of the beam to one side of the machine, while allowing another part of the beam to pass. Then a second detector is located at a known distance, following the same method. On the sides are two light detectors with photosensitive plates, to obtain the incidence position of both beams. From those deviated light points inside the target, the position of the machine is calculated against a laser reference in a general purpose computer. This way, the machine is guided by a laser beam that marks the orientation and slope of the drive.

Commercial guiding systems based on similar working principles are Gyro Navigation Systems Laser Referencing [4] or Laser based system [5]. An example of commercial gyro based is one that uses a fibre optic/mechanical gyro mounted within the TBM [6]. The Gyro System is best for small diameter drives (ID 800–1,200 mm) or for drives with muck skips and bounded views inside the TBM. A Gyro-based Guidance System uses a self-leveling, north-seeking
gyro compass for the determination of the horizontal position and azimuth of the TBM. The calculations of these systems are based on the principle of the dead reckoning, which means that it is assumed that the TBM will move exactly along the direction which is defined by its axis. The difference between the direction along which the TBM is moving and the axis of the tunnel is called the TBM drift. This drift will influence the calculations of the system. Check measurements for the calibration of the system must therefore be carried out every 20–40 metres in order to determine the TBM drift and reduce its influence to a minimum by pre-setting this value for further calculations.

For a curved drive the choice of commercially available guidance system is either a Laser based system or a Gyro Navigation System. Laser based guidance systems have been used for many decades. Traditionally, a laser station is firmly fixed inside the tunnel, projecting a laser point onto a laser target board mounted on the TBM, as shown in Figure 1. A target that is sensible to a red laser (635nm) is mounted on a TBM as an aiming unit, with high precision motorized prisms, for electronic measures of distance mounted on the same vertical axis of the target’s centre. Based on the offsets of the laser spot on the target board, the TBM operator infers the current line and grades the tunnel alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser alignment deviations. A laser emitter is located on the operator infers the current line and grades the tunnel alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser alignment deviations. A laser emitter is located on the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the fixed platforms on the back side of the tunnel. The laser beam has to be obstacle free to be able to impact in the.

Figure 1: Laser target referencing.

Following a straight line, the distance between the platforms is normally between 30 and 100 meters, which is the minimum advised distance for installing the platform from the target, because with a lower distance, the machine thrust forces generate movements in the concrete segments which would lead to lack of precision in the guiding task. Likewise, the maximum advised distance so the laser beam refraction errors do not affect the precision is 100 meters.

Examples of these commercial systems are the ELS Electronic Laser Systems Basic System or Video Target based systems. An example of the first could be found from ZED Tunnel Guidance Ltd [7]. Another ELS based commercial systems are SLS Microtunneling LT guidance systems supplied by VMT GmbH [8]. The TACS GmbH company also offers video target based systems [9]. These are guiding systems for tunneling machines based on the incidence of a laser on a target sighting, designed for tunnel alignment in which only big curvature radius can be produced. Those systems are based on sensitive panels, fixed to the TBM body, and on the usage of an inclinometer. The sensing principle of TACS’ video targets has inspired our design. It consists of two target boards and two cameras inside a robust housing (IP68) in several measures [9]. The video targets determine the coordinates of the laser hit point by means of high precision video cameras and proprietary recognition software.

1.1 Drawbacks and limitations of current guidance systems

Traditional guidance systems are based on several modules working over a laser-generated reference (red light, 620-640nm). However, the first factor that limits the accuracy and reliability of traditional laser guidance systems are the working environment. Tunnels’ condition of dust is a limiting factor for the maximum range of the guidance system. Dispersion and refraction of the laser beam over a long distance make difficult for a laser beam to impact the photo-electronic sensible panel of the front side on the target. The second limitation is related with the geometry of the projected tunnel: sudden changes in tunnel axis direction have been avoided to limit the point charges on the exterior edges of the pipe, and to prevent damage in joints. Abrupt changes in direction would also bring as a consequence great frictional resistance against the previous pressing of the complete pipe length, which could only be solved by especially elevated pressure loads. The more different the trajectory compared to the plan, the more pressure on each section of the tunnel and ground settlement in the area. In the case of concrete segment construction systems, coupled behind the microtunneling machine, the drilling heads transmit vibrations and movement on the terrain under the concrete segments and between these same concrete segments, which affect the inertial instruments and the positioning of the reference base stations when moving forward.

The third limiting factor is related to the fixation platforms of aiming lasers. Although concrete segments
are manufactured according to the original plan, they are not exempt from small relative movements, which can be sources of error if not considered in the calibration phases. As soon as the tunneling machine and the length of pipe are so far ahead that the laser beam can no longer reach the target from the access chamber, the laser should be mounted on a console, in a proper position behind the target, fixed in the length of pipe. This assumes that with the traditional systems, every 30 to 100 meters the laser or prisms should be properly relocated to achieve the guiding task, which creates a sizeable delay and a lowering in the total machine’s productivity, due to the big performance that tunneling machines offer (drilling up 40 meters per day).

The fourth limiting factor lies in potential manual errors in initializing or calibrating the laser beam’s alignment. Although human factors could be avoided by providing more automation to the guidance system, most laser systems suffer difficulties to receive laser’s projection because of excessive TBM deviations. Current ELS targets have been in reduced dimensions because the cost of photosensitive cells is very important. This leads to having a very small laser window (typically 100x100mm). Therefore, in case of machine deviations, where the laser window is missed, the position of the laser should be changed to locate it again on the laser window. For the same costs, a Video based target with optical cameras with bigger dimensions can be constructed.

2 Design and development of a novel vision based target

To improve the limitations of the current guiding systems, what is needed is a guiding system that combines the classic concepts of tunnel survey in an optimal manner, against the operation costs, achieving a guiding system that: a) Shows immediate information of the tunneling machines against the designed axis, offering an excellent precision, normally above the driving features of a tunneling machine. The tunnels are normally built with a few centimetres of error in the path, but this is caused by the impossibility of driving the tunneling machine exactly through the indicated axis given by the guiding system. b) Does not interfere on the progress or forward movement capacity of the machine, for having to reposition and recalibrating continuously the stations and reference prisms, or relocating the aiming target in relatively short time intervals, to allow monitoring the tunnel’s construction.

2.1 Functional architecture

In order to achieve the above goals, improvements in subsystems, the laser target, and the user interface have been implemented. A low cost reliable guidance system has been designed for any type of TBM (e.g. pipe-jacking or concrete segments) based on a new laser type and target combination. The target has been totally designed from scratch to allow sensing the laser beam inside, e.g. has gotten two surfaces where the beam incidence is capture by a vision system. As described in figure 2, the first surface is located in the front window, which is semi-transparent. The other surface is the interior back panel of the target. The inner walls of the target space are black painted and the cameras and on board electronics are located on the bottom. The front panel is made of a methacrylate panel with an embedded 1 mm clearance net (grill).

![Functional layout of the novel vision target device.](image)

**Figure 2: Functional layout of the novel vision target device.**

2.2 Implementation of beam incidence capture system: Vision target

The new video target developed is sensitive to green laser (532nm) instead of red laser as in commercial targets. It performs high resolution image processing and simple but reliable pose location algorithms. To complete the design of a vision based guidance system, this video based target has to be integrated in a modular guidance system that could be mounted in several types of machines with few adaptations.

The general architecture has three parts. The first is the smart target, which is the remote equipment (right part), where we have the target formed by the cameras, embedded computer, and the encoder (drawn in the left part of the image) for odometry. The encoder block is optional. In this module the image capture and processing modules will run obtaining the information from the inclinometer. The second part is formed by the local equipment in the command cabin, which executes the graphic interface, implemented in a computer with a touch screen. To connect them a wired or wireless data link allows the exchange of data to the guiding interface and the smart target. This block runs the representation system and the information representation relevant to the guiding process, in a touch screen. Here the data provided by the calculation module are represented and an interaction with it can be made. Also the communications module will be executed where the
protocol and the transmission method for the processed information will be established for several tunneling machine environments. The third part, which is essential, is a green colour laser emitter (532nm), where the axis has been properly positioned by using references to the base station, by traditional topographic methods. The complete system can be seen on figure 3. The visual system has been developed under the premise to use a green laser emitter, because it has a longer range and less dispersion problems than a red laser (635nm).

The on-board electronics of this remote system is based on low-cost robotic hardware and software technologies. The core of the system a low-cost embedded “ROBOARD”, a 32-bit credit-card sized x86 CPU, widely used to develop low cost robotic applications. This CPU [10] manages two USB cameras and the inclinometer unit. It is fully compatible with x86 libraries and its cost is about 200€. Image acquisition is performed by two low-cost (about 40€) and simple USB Logitech Quickcam Pro 9000 (V-U009), working at 2MP resolution. OpenCV computer vision libraries [11] and YARP middleware [12] manage the image vision and wired/wireless data link to the tunneling machine control station (normally on surface and not underground, except in open shielded TBM), where a local PC running a GUI application allows to get machine location data on execution time.

![Figure 3: Hardware architecture of the new guidance system.](image)

The new video target, which has external dimensions of 700x225x160 mm and beam reception windows of 170x170 mm, is installed with a clamping support to the tunneling machine, mounted on its higher panel. The guiding computer fulfils the protection standards IP65/NEMA4, indicating it is protected against dust and low pressure water, which is expected to suffer the most severe environment conditions and should be properly protected for its usage. The system’s power is 220/230 V in a guiding cabin for the GUI computer and 9 V DC (10W) for the target module.

They are joined by a UTP C5 network communication cable with military RJ-45 connections. About 15 m of cable were used as a necessary length. In order to make the image processing from the cameras more robust, a notch High Performance Laser-Line Filter, from Edmund Optics of the 12.5mm diameter model, was installed in the back camera pointing towards the translucent window, getting rid of all the frequencies except the laser one with a wavelength of 532nm and bandwidth of 2nm. This eliminates the problem of obtaining an image with excessive exterior tunneling machine light.

The use of open source libraries and low-cost, but rugged hardware approach, make an increased reliability with respect of expensive commercial systems possible, allow a more simple operation and setup phase, and incur in lower operational costs. The most expensive piece of equipment is the tactile panel due to the high IP code requirement required for an underground environment. However, this unit could be removed thanks to the data link implementation, which allows operating the system from virtually anywhere.

2.3 Vision algorithms for beam incidence capture

Once the image is captured, the next step is to process it to determine the central point or centroid, which represents the laser beam hitting the panel. The problem in these conditions is to obtain coordinates from the centre of the object that would be detected as the laser beam. It should be simple, because it basically shows a black image with a single point caused by the laser. The vision target’s interior walls have been painted in matte black. This makes easy the image recognition. However, the translucent panel causes more problems. This panel is subject to external environment conditions, either because of the luminosity provoked by the luminaries in the interior of the machine and that would be impossible to hide to the target, or the filth and water in the environment. For these problems, the window has a net in the translucent panel and a notch.

![Figure 4: System Setup for UC3M’s guidance system.](image)
filter to recognize only the wavelengths close to the ones from the laser of 532nm for green light.

The fundamental problem to precisely recognize the incidence point is to discriminate between the different points found on the image, which can be either the desired laser point, as reflections produced by the laser or the external illumination, which is why the image has to be treated to separate this point from the others, and even if the laser beam is lost, provoked by situations like a material or a worker is on the laser's trajectory or it falls out of the range, the program detects it and communicates the corresponding message.

The image processing, once loaded, binarizes the image based on a threshold, searches on the image all the possible objects with a circular shape that are within a characteristic diameter (because the laser a beam of 12mm nominal diameter) and calculates the centroid of the given object. In case no object with such characteristics is detected, this could be because the threshold value is too high, thus eliminating even the point generated by the laser, reducing this threshold a fixed value and reproducing the previous algorithm over and over until it looks similar to the one needed. In case it detects several objects that do fulfill these characteristics, having reached the binarization threshold (where it doesn't detect the laser incidence), the biggest areas are chosen as acceptable points, communicating to the guiding employee there could be some miscalculations caused by excessive light on the target. From the point obtained as the laser beam, the program communicates X and Y coordinates as well as longitudinal deviations produced in this image plane, diameter of the detected object and number of possible laser beam points found in the last program's iteration, in case a possible error in the beam detection needs to be notified [13].

As shown on the figure 5, for the proper measurement of the incidence of the laser in the frontal window, a camera with a notch filter attached to it has been used so it allows only the green laser frequency (532nm) to pass. Despite the fact that it lowers the amount of perceived light, it improves obtaining the centroid of the image, because reflections and light sources that could incur in false estimations of the incidence position of the beam in the frontal panel are removed. To convert the information of the images of the laser incidence on the target’s back window (frontal incidence image) and the back panel (back incidence image) from the orientation of the machine, the installation position of the target against this and the inclination of the machine against its forward axis (roll) should be known. A low cost but precise inclinometer can be used for this purpose. Knowing the fixed location of the target with respect the tunneling machine, the next section explains how to calculate the machine’s orientation with respect to the laser line, and the machine location with respect to the laser base location.

2.4 Position and orientation estimation

Here a brief description of the referenced setup while drilling and associate references will be presented. Initially the system is calibrated by using the laser beam reference against the coordinate system of the topographic system. Likewise, the target is located at the beginning of the tunnel with the Theodolite topographic laser station.

When initializing the localization algorithm, the following reference systems (coordinates systems or SDC) between the different parts of the guiding task are considered:

a) SDC\\_w: World coordinates system, fixed to the entrance of the attack well as a geo-localized global reference system.
b) SDC\text{Laser}: laser position coordinates system, which will be fixed as a reference for the aiming system that will calculate the position of this laser beam, but that will be relocated when its range surpasses the longitudinal reaching limits of the beam (normally under 500m for a green color laser) or when the tunnel’s radius of curvature geometrically prevents it from reaching the aiming system being used.

c) SDC\text{EPB}: reference system to express the tunneling machine’s coordinates against the SDC\text{w} coordinates system, as the main output of the system.

d) SDC\text{target}: in current commercial systems, located at some point of the traditional target, typically in the centre of the laser’s detection panels. Commercial systems only use an incidence plane and the SDC is fixed. Therefore, it is fixed from installation and should be referenced when the machines starts, by either the topographer or the machine’s manufacturer.

Once the coordinates in the image of the laser beam have been obtained, they have to be referenced to the same coordinates in millimetres above the central point of the image as the origin of the coordinates. This is a typical pose estimation problem well modelled by the proper SDC. It has been identified that it is then possible to solve the pose transformation from a set of equations which relate the SDC orientations of the rear and front panel, with their 2D image coordinates, taking into account that both are parallel planes. To do this, a C++ class has been developed that creates simultaneous equations relating the corners of the images as points of known coordinates, generating a fixed calculation matrix that, multiplied by the coordinates obtained by the vision system, obtains the coordinate system that transformed finally to the deviations from the origin of the tunneling machine as a solution.

The details of the equations of calculation and composition of rotations and displacements are detailed in [13]. They have been programmed using MATLAB scripts and afterwards converted to C++, so they can be integrated in the system software architecture. The modules of calibration and image capture, the inclinometer readings and other sensors (i.e. gases), and alarms (i.e. by incidence lost) are executed on the target. All the modules communicate using YARP that also runs on the pilot’s touch panel. Here the GUI, the calculation module, Data Logger module and the windows for warnings and errors are also executed. The implemented architecture has the guiding system in the machine itself, performing the connection directly by Ethernet protocol. A standard twisted network cable has been employed, properly protected. In case the guiding screen would be located outside, the same structure would be kept but a physical media to interconnect the different systems at long distances would be needed.

2.5 Pilot interface

In order to successfully carry out a teleoperated microtunneling machine drilling layout plan, not only is important how to achieve the referencing but also equally important is how this information is presented to the pilot.

The size of the tunneling machine must be known, as well as the position of the target inside the given machine. Figure 7 shows the necessary setup data that enables a correct calculation of the software guiding trajectory. Using the size in (mm):

A: Total tunneling machine length (11000).
B: Diameter of tunneling machine (2000).
C: Distance from back of tunneling machine up to target's translucent window (8250).
D: Distance from target's translucent panel to front of the tunneling machine (2750).
E: Height from the central symmetry axis of the tunneling machine to target's longitudinal symmetry axis (21).
F: Horizontal distance from the central axis of the machine to target's centre (676).

The table of guiding records is important so the user that controls the tunneling process monitoring can do a full parametric tracing of each forward movement produced in the tunneling machine. This data is shown in the table in the lower part of figure 10.
The main parameters to show in a guiding system are the horizontal and vertical deviations, and the called drift and tendencies. The deviations are the distances between the axis of the project and the axis of the tunneling machine, measured in the perpendicular plane to the first. The tunneling machine is usually presented as an arrow and the deviations are presented in two points of this axis, the front and the back. These deviations are given every time a position is taken, because when giving coordinates of these points, they are compared against the path stored in memory. One of the most important parameters for the pilot is the tendencies. The tendencies indicate the movement that the machine would have after a forward movement, giving an idea of the scope of the correction being applied. The pilot guides the machine exert certain pressures in the guiding cylinders and in the using ones, and these pressure values are taken by the program to calculate tendencies.

Another improvement of the systems relies in the user interface, in which that information is displayed and updated dynamically with the movement of machine. In the following section, a new way to display the position of the machine using a 3D representation is presented. This way, the pilot responsible for the operation can understand better where the machine stands and what the following actions to perform on the control of the machine to get take on the planned route are.

2 Experimental results

The preliminary results of the prototype in its current testing phase in real scenarios were obtained at a workplace where the guidance system was mounted into an open-shield boring machine. From this moment, the TBM started boring in the field workplace. Due to existing limitations and the impossibility to alter the machine's production rates, the tests were performed in parallel with a commercial system. Our developed system was mounted in parallel with the EUROHINCA proprietary system based on a customised VMT’s SLS Microtunneling LT system, according to the employed machine [8]. In the initial starting phase (Phase 1), the laser station, which comprises VMT’s standard Leica TCA 1203plus motorized total station and which includes an integrated diode laser mounted parallel to the visual axis and a sensor system that allows automatic targeting of prisms, was mounted on a custom built measurement pillar between the jacking-pit thrust rams. The green laser diode emitter was mounted in parallel and also referenced. For reference measurement a further pillar with a survey prism was installed with a reference prism outside the shaft. The green laser used was a Topcon model TP-L4G, mounted and calibrated for its usage, powered with a battery. On the first stage, distance tests were executed at 50 m, 100 m, and 200 m in the environment of the Collector Tunnel of Jalon River, a tunnel project with internal diameter of 2000 m and 479 m long [14]. After initial setup and verification, the first distance test was done at 50 m.

Figure 8 shows the incidence of the beam in the front (up) and rear (down) window taken at 50 m. The measurement taken at 200 m turned out to be complex, because it required relocating the commercial system in parallel and then calibrating it. Afterwards, without moving the green laser from its original position, the incidence of the green laser beam in the target was obtained, resulting in the images shown on Figure 8.
The Figure 9 shows the dispersion of the beam due to the distance and the presence of suspended particles because of the roadheader working. Even though during the test the resulting beam incidence is highly faded, thanks to the robustness of the vision algorithms, the data from the position of the target was obtained. Then, the data is shown by the guiding interface. In this case, and even though much dispersion exists, the algorithm could estimate the centroid of the zone of the greatest light intensity, meaning, where there is more incidence. Therefore, it indicates the beam’s direction, which highlights this data as an improvement factor of current targets. This allows, for example, using conventional mirrors, to redirect the beam in case of incidence loss.

As shown on Figure 10, the main screen of the interface developed for its validation during work has been implemented on a touch PC with IP67 protection. The information is presented in such a way that the pilot comprehends the effects of his/her actions on the trajectory in a simple manner, because of the 3D visualization, and can understand the behaviour of the machine because of the actions and correct them in an easy and intuitive way. During excavation, the laser beam was maintained continuously on the ELS (electronic laser system) target during the TBM advance.

The calculated values of position are displayed on the system’s computer screen and stored in the database. The three-dimensional coordinates of the axis point in the ELS target unit plane are compared and then saved as the projected course of the TBM measured by conventional procedures.

![Figure 9: Front window view and rear window view (200m)](image)

![Figure 10: Guidance 3D path as showed in the GUI](image)

For each instant, a calculation of estimated location of the TBM measured with our system was performed. In table 1, a comparison of deviations and estimated errors for both systems are shown. All the measurements obtain a 0.2 mm linear and a 0.1º angular repeatability. As trueness in the measurements is close to precision values, a good accuracy is obtained. It is proven that the developed prototype reaches sharper precision values for the laser beam capture. This implies better capacity when obtaining the machine’s position.

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Accuracy</th>
<th>Angular</th>
<th>Angular</th>
<th>Angular</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD mm</td>
<td>VD mm</td>
<td>Alfa °</td>
<td>Beta °</td>
<td>Gamma °</td>
</tr>
<tr>
<td>ELS1</td>
<td>1</td>
<td>0,05</td>
<td>0,05</td>
<td>0,05</td>
</tr>
<tr>
<td>ELS2</td>
<td>1</td>
<td>0,05</td>
<td>0,05</td>
<td>0,05</td>
</tr>
<tr>
<td>UC3M</td>
<td>0,267</td>
<td>0,025</td>
<td>0,022</td>
<td>0,016</td>
</tr>
</tbody>
</table>

The second table shows a comparison between the angles and deviations on current targets and the UC3M prototype.

<table>
<thead>
<tr>
<th>allowed drifts</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD mm</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>ELS1</td>
</tr>
<tr>
<td>ELS2</td>
</tr>
<tr>
<td>UC3M</td>
</tr>
</tbody>
</table>
Table 2 has the maximum horizontal and vertical deviations as well as the maximum angular pitch and inclination measurable with the designed target's prototype. Note that values are similar to actual models but should be considered the lower cost in the components.

![Figure 11: Guidance system working in workplace and laser incidence at GUI](image)

It was estimated that a prototype can be implemented with bigger vision panels (about 250x250 mm) and shorten the length of 700 to 400 mm, to maintain the angular precision, having proved in situ that there is enough space allowing this application.

Note that the simple design of the target allows rebuilding it in custom sizes without great efforts, especially if the size of the vision panels has to be increased. This makes shortening the length of the target possible, to maintain the current angular resolution.

3 Conclusions

A modular a low-cost reliable guidance system for TBM's has been presented. The preliminary results of the prototype in its current testing phase in real scenarios (working in parallel with previous commercial units for comparison) were obtained in a real workplace where the guidance system was mounted on an open-shield boring machine. Increased reliability has been achieved with respect to commercial systems, using low-cost embedded hardware with a Linux based O.S. and OpenCV libraries, allowing a wired or wireless data link to TBM control station (normally on surface and not underground). This low-cost approach is not only for development of the target itself, but also allows lower setup phase and operational costs of the guidance system. Note that the most expensive equipment is the tactile panel, which could be avoided if communication protocols of machine manufactures would be open.

The vision based incidence capture is more flexible and reliable in worst conditions (extended range & lower curvature radius allow less repositioning of reference laser beam station), allowing saving time in comparison to conventional procedures, by increasing the distances travelled without repositioning the reference laser. With this, avoiding much influencing in the work activity is accomplished, preventing losing time, allowing a complete control of all the components of the system from a computer. An easy to handle software HMI over a tactile hardware suitable for tunnels was also implemented.

As expected, Video Target guidance systems are severely prone to problems of occlusions. With the presented system, the target’s capabilities and limitations are exploited, working without needing slow and expensive relocations and recalibrations even if the machine has great deviations. A proper target installation in the setup phase of the tunneling machine causes the beam's incidence favoured and does not come out to quick from the target's range, avoiding future relocations. This minimizes the maintenance effort of a surveyor during the whole working process. Additionally, with the combination of this extended range guiding system and a proper concrete segment design, even lower curvature radius can be achieved.

4 Acknowledgements

This research was substantially supported by RoboCity2030-II-CM project (S2009/DPI-1559), funded by Comunidad de Madrid and co-funded by Structural Funds of the EU. Especial thanks to EUROHINCA’s staff for their help in conducting field trials.

References


Abstract -

One of the most prohibiting factors when attempting to reuse structural steel members or systems of members is the time and labour required to accurately determine dimensions. Current practices dictate that all required measurements are recorded by hand using tape measures and calipers which adds a significant cost to reused steel. To mitigate this cost, a semi-automated method for identifying structural steel components and systems is proposed that uses data acquired in the form of a 3D point cloud. Current research in the field of automated object recognition currently has two major limitations: (1) a priori knowledge, such as a building information model (BIM) is required, or (2) only simple, flat surfaces can be identified. The purpose of this study is to preliminarily investigate the possibility of automating the process of (1) cross section identification, (2) end connection geometry of bolted connections, and (3) relative component position of multi-component, planar structural systems such as trusses. Cross section identification is performed by creating filters that match standard structural sections and then convolving them over images of the cross section data. The end connection geometry is identified using Hough algorithms to detect lines and circles representing the limits of the component and the bolt holes, respectively. Planar structural systems are identified using Hough algorithms to detect lines which represent the components of the system. The results from the proposed methods show a strong potential for fully automated processes to be able to identify structural steel components and systems without a priori knowledge.

Keywords -
Steel Reuse, Object Identification, 3D Imaging, Automation

1 Introduction

Steel reuse has long been an underutilized method for reducing the life cycle cost of structural steel construction. It is believed that steel reuse could be increased by up to 150% of current reuse rates if there was a greater economic incentive to reuse steel [1]. This was found by performing an extensive survey of members of the steel construction industry. These members included steel service centers, demolition companies, scrap steel dealers, steel fabricators, and engineers and designers. A common concern among these groups was the economic benefit, or cost, of reusing structural steel because of the low cost of new steel components and the high value of scrap steel. A common method of reducing the cost of specific activities is to automate that activity, or part of that activity. In the field of civil engineer, a number of research endeavours have been undertaken with the goal of automating the identification process of physical objects. The automation process begins with data acquisition. A number of 3D sensing methods have been developed to reduce the cost associated with data acquisition. These methods can be divided into two categories: image based systems, and time-of-flight based systems. An image based system records many digital images of a scene and then registers these images together using the principles of collinearity and coplanarity [2]. This process is often referred to as photogrammetry or videogrammetry, depending on whether digital photographs or frames from digital video recordings are used. Alternatively, time-of-flight based systems, which are typically referred to as laser scanners, measure the time required for a signal to travel from the base unit to the object and back. The main advantage of using image based systems is the speed of data collection [3]. Image based systems can perform 3D mapping in real time [4]. The benefit of using a time-of-flight based system is the higher accuracy that is achieved. Regardless of the collection method, the result of this type of data collection is a 3D point cloud of the structure’s surface geometry. This point cloud will normally contain millions of individual data points.

After the physical geometry of the structure has been recorded, automated object recognition can be performed. Automated recognition of civil engineering components is not a new field of research. Bosché [5] presented a means of detecting construction objects for the purpose of automated progress tracking as well as determining dimensions for dimensional control of steel structures. This built on the original method developed by Bosché and Haas [6] which compared 3D point cloud data to 3D
building information models (BIMs) to determine the presence of a component. Dimensional control of concrete structures, specifically the flatness of concrete slabs, was investigated by Tang, Huber, and Akinci [7]. Marble façade panels have also been analyzed for flatness using 3D point cloud data as well [8]. A common limitation of this line of research is that a priori knowledge is required, whether it be through a BIM or the knowledge that the object in question should be perfectly flat.

The field of machine vision attempts to automate detection, often without a priori knowledge. One of the cornerstones of machine vision is known as the Hough Transformation. The Hough Transform is used to identify shapes in 2D or 3D images. While the theory and underlying objectives of the Hough Transform remain the same, the implementation, and effectiveness of the implementation, can vary greatly [9]. This detection method is commonly employed for the automated detection of planar elements in binary images [10]. For 3D applications, automated edge and surface reconstruction has been performed from point cloud data [11] [12]. Providing edges and surfaces with semantic information has only been performed for very simple surfaces, such as ceilings, floors, and walls [13]. Throughout all of this research, the ability to automatically identify complex objects, such as those that would be found in a structural steel building, has not been established.

Automating the process of identifying existing structural steel components would significantly reduce the cost of reusing such materials, thereby making steel reuse a more economically viable alternative to new steel. It is forecast that this sort of economic improvement to the steel reuse industry would facilitate a significant increase in the reuse rate [1].

2 Methodology

2.1 Overview

The research method was separated into three preliminary studies: (1) cross section identification, (2) end connection geometry identification, and (3) joint location identification. The goal of these studies was to demonstrate the possibility of automating the identification process of structural steel components without a priori knowledge. As this was a preliminary study, the employed research method was semi-automated and limited in its applicability.

The cross section identification was performed on standard wide flange beam cross sections, but the developed algorithms were applicable to any standard cross section. The end connection geometry identification was performed on a plate containing holes. The intent of this study was to simulate the end of a beam with bolt holes located in the beam’s web. The joint location identification algorithm was performed on a planar truss. The developed algorithms were based on the assumption that all members of interest would be aligned along a common plane.

2.2 Data Collection

The data collection for each research area was conducted in the same manner but on different structures. Data collection was carried out using a terrestrial 3D laser scanner. This laser scanner records geometric data as a list of x,y,z coordinates from objects within line-of-sight of the base unit. This meant that the laser scanner recorded the geometry of the desired structure as well as any nearby objects such as trees or other structures. The accuracy of the particular laser scanner used in this study is ±3 mm at 25 m [14].

The geometric data from one structure was recorded for each of the research methods presented. The cross section identification used data from a structural steel teaching aid (Figure 1a). This structure consists of a number of different standard steel cross sections, including wide flange beams, hollow structural sections, and channel sections. The end connection geometry identification was performed on data collected from a steel plate containing eight bolt holes (Figure 1b). The joint location identification was performed on data collected from an exposed truss supporting the roof of an arena (Figure 1c). Each of these data sets was comprised of multiple instances of laser scanner data that have been merged into a common coordinate system using commercial software.

Figure 1 - Point cloud data used for cross section identification (a), end connection geometry identification (b), and joint location geometry (c)
2.3 Cross Section Identification

The process of cross section identification began by pre-processing the point cloud data. In this step, members that were to be analyzed were isolated from the rest of the point cloud and aligned such that the principle axis of the member was aligned with the z-axis and the strong axis of the cross section was aligned with the x-axis. This step was performed manually in 3D computer-aided design (CAD) software.

This separated and aligned point cloud is then sliced orthogonally to its principle axis with the purpose of creating binary images that represent the cross section geometry of the scanned member (Figure 2a). These images were created by first projecting the points contained in a single slice onto a plane that was parallel to the slice. This plane was then pixelated and pixels containing a point were marked as being filled. A varying number of the binary images were created based on the slice thickness and length of the member’s point cloud. The slice thickness and pixel size of the resulting binary image were user defined parameters. These binary images then needed to be compared to equivalent binary images that represent the geometry of known structural steel sections. For this purpose, a database was created that contained a filter for each standard structural steel cross section. These filters were designed based on simplified geometry, where components of the cross section were always treated as rectangular (Figure 2b).

A convolution algorithm was implemented to compare the binary image of the scanned data with the binary image of the filter. Each filter was compared to each slice’s binary image, and a match was established based on the number of corresponding pixels from both images. The matches for each filter and each slice were recorded and the highest match value from all slices was selected as the match for that particular member.

2.4 End Connection Geometry Identification

End connection geometry identification began with pre-processing the point cloud by trimming any points not associated with the plate and by aligning the plate with the yz-plane. The aligned point cloud was then converted into a binary image (Figure 3a) by projecting the data points onto a parallel plane, pixelating the plane, and then marking pixels that contain at least one data point as filled. With a smaller pixel size, and therefore more detail, it was likely that this binary image would contain many empty pixels that should be filled based on the assumption that data points are taken across the entire plate’s surface. To fill these pixels, an expansion and contraction algorithm was implemented. This algorithm began by filling each pixel that neighbors a filled pixel. Alone, this method would alter the size of the plate and holes by expanding the region that was classified as the plate’s surface. To remedy this, the filled pixels are contracted whereby each filled pixel that neighbors an empty pixel was marked as empty. The resulting image is one that represented the plate’s surface as filled pixels (Figure 3b). Finally, all non-edge points were marked as empty (Figure 3c) because the detection algorithms used for the geometry identification required that only edge points be present in the binary image.

Figure 2 - The binary image of sliced point cloud (a), and a filter (b)

A convolution algorithm was implemented to compare the binary image of the scanned data with the binary image of the filter. Each filter was compared to each slice’s binary image, and a match was established based on the number of corresponding pixels from both images. The matches for each filter and each slice were recorded and the highest match value from all slices was selected as the match for that particular member.

Figure 3 - The binary image of a steel plate with bolt holes after creation (a), filling (b), and edge detection (c)
The detection of important features, namely the edges of the plate and the size and location of bolt holes, was completed using Hough algorithms. The algorithms used for this particular study were pre-built, commercially available implementations of the Hough transform for line detection and the Hough transform for circle detection. The particular Hough algorithms that were implemented returned the start and end locations of detected lines and the center and radius of detected circles.

Post-processing of the detected lines and circles began by extending and/or trimming detected lines such that their ends met at a common point. This was performed so that the detected lines would properly represent the edges of the plate. Further, and optional, post-processing was performed based on two assumptions: (1) the plate is expected to be rectangular, and (2) the holes are expected to be arranged in a grid pattern. The detected lines and circles were adjusted to meet these assumptions. First, each line was rotated about its center so that it ended in a horizontal or vertical alignment. The lines were then extended and/or trimmed, as before. Next, the center locations of holes were aligned such that holes in a common row would have a vertical position equal to their average vertical position before adjustment and holes in a common column would have a horizontal position equal to their average horizontal position before adjustment.

2.5 Joint Location Identification

The pre-processing for the joint location identification algorithm began in a similar manner to the end connection geometry identification. The point cloud was first trimmed to only include the points of interest; specifically, the truss to be analyzed. The truss was then aligned such that the plane of the truss was parallel to the xy-plane. As before, a binary image was then created by projecting the data points onto a parallel and pixelated plane, where the pixels containing a point were marked as filled (Figure 4a).

The analysis phase began with an implementation of the Hough transform to detect lines in the image of the truss. These lines were detected iteratively by detecting a single line and then clearing pixels that are in close proximity to that line. Figure 4b and Figure 4c show this progression for the first two detected lines. Figure 4d shows a later stage in the iterative process where the first line representing a web member has been detected. This process results in a number of detected lines that is much greater than the number of truss elements that exist (Figure 5a). The elimination of redundant lines was performed during post-processing.

![Figure 4 - Binary image of a truss before line detection (a), after the first line has been detected (b), after the second line has been detected (c), and after the first line representing a web member has been detected](image1)

![Figure 5 - Detected lines before post-processing (a), after merging (b), after trimming and extending (c), and after end point adjustment (d)](image2)
Post-processing of the detected lines was carried out in a three step method. The first step was to eliminate redundant lines to produce results where the number of detected lines matched the number of members in the planar structure (Figure 5b). This step was performed by determining groups of lines that had similar slopes and close proximities. Each group of lines was then merged into a single line in an average position based on the group. The limits of the line were determined based on the extreme limits from the lines in the group. The next post-processing step was to trim and/or extend each line so that lines met at common points (Figure 5c). For the case where more than two lines were approximately intersecting, an adjustment was made to force this intersection to a single point (Figure 5d).

3 Results

3.1 Cross Section Identification

The cross section identification process resulted in a match between the scanned member and a standard structural steel section. In total, nine members were analyzed that varied in size and proportion. These members were also hand measured to determine their true cross section.

When determining the capacity of structural steel members, three properties are important: (1) section modulus, (2) web area, and (3) cross sectional area. The section modulus, web area, and cross sectional area can be used as a very simplified and relative measure of bending capacity, shear capacity, and axial capacity, respectively. The measured and predicted, based on the identification process, value for each of these strength metrics were compared (Figure 6).

It was found that the cross section identification algorithm presented in this work had errors of between 15% overestimate to 41% underestimation. For the web area, the error was between 16% overestimation and 23% underestimation. Similarly for cross sectional area, the error was between 21% overestimation and 36% underestimation. This particular sample set had two members that were correctly predicted.

3.2 End Connection Geometry Identification

The predicted values for the plate dimensions were quite accurate when compared to measured results. The height of the plate had an average error of 1.5 mm and the width had average error of 4.5 mm. The overall plate dimensions were approximately 400 mm x 300 mm. These errors are within the expected amount of error from the scan data alone. The holes radii experienced significantly more error. The average error for hole radius was 6.3 mm but the error was very uniform for all the holes. The minimum error was 6.0 mm and the maximum error was 6.6 mm.

Figure 6 – Comparison of predicted and measured section modulus (a), web area (b), and cross sectional area (c)

The location of bolt holes is an important measurement to have accurate knowledge about when incorporating a used component into new design. The measured, predicted, and adjusted bolt hole locations were calculated (Figure 7). The hole locations shown in Figure 7 are relative to the lower left corner of the plate. In this relative position, the error in hole center before post-processing alignment was between 1.0 mm and 3.6 mm. After post-processing alignment, this error was reduced to between 0.0 mm and 3.1 mm. The error can be reduced if the relative position of the plates is translated to minimize the error in bolt hole center locations. After translation and adjustment, the error was further reduced to between 0.4 mm and 1.5 mm. These errors are within the expected accuracy of the laser scan data and the fabrication tolerances.
3.3 Joint Location Identification

Physical access was not available for the truss used in this preliminary study due to safety concerns regarding its location in the building. As such, a measured length for each of member was determined by manually tracing the scan data of the truss in 3D CAD software and measuring these lines. Table 1 shows and compares the calculated and measured lengths, where the web members have been labeled sequentially from right to left (i.e. the rightmost web member is Web 1 and the leftmost web member is Web 8).

<table>
<thead>
<tr>
<th></th>
<th>Calculated</th>
<th>Measured</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Chord</td>
<td>8.14 m</td>
<td>8.00 m</td>
<td>0.14 m</td>
</tr>
<tr>
<td>Bottom Chord</td>
<td>6.09 m</td>
<td>6.60 m</td>
<td>0.51 m</td>
</tr>
<tr>
<td>Web 1</td>
<td>1.89 m</td>
<td>1.85 m</td>
<td>0.04 m</td>
</tr>
<tr>
<td>Web 2</td>
<td>1.93 m</td>
<td>1.85 m</td>
<td>0.08 m</td>
</tr>
<tr>
<td>Web 3</td>
<td>1.89 m</td>
<td>1.85 m</td>
<td>0.04 m</td>
</tr>
<tr>
<td>Web 4</td>
<td>1.89 m</td>
<td>1.85 m</td>
<td>0.04 m</td>
</tr>
<tr>
<td>Web 5</td>
<td>1.89 m</td>
<td>1.85 m</td>
<td>0.04 m</td>
</tr>
<tr>
<td>Web 6</td>
<td>1.83 m</td>
<td>1.85 m</td>
<td>0.02 m</td>
</tr>
<tr>
<td>Web 7</td>
<td>1.95 m</td>
<td>1.85 m</td>
<td>0.10 m</td>
</tr>
<tr>
<td>Web 8</td>
<td>1.83 m</td>
<td>1.57 m</td>
<td>0.26 m</td>
</tr>
</tbody>
</table>

Excluding the large errors associated with the top chord, the bottom chord, and Web 8, the error for each member was not greater than 0.1m. This translates to less than 6% error in the length measurement for these members. Larger errors for the three aforementioned members can be explained by the trimming and/or extending step of the identification procedure. This step would result in the top chord and Web 8 being artificially lengthened to meet at a common location. Also, the bottom chord would be trimmed to end at the intersection of the outermost web members.

4 Conclusions

The three methods for identification of structural steel components demonstrate the possibility of accurately automating this process. The cross section identification procedure was able to predict the capacity of the cross section within ±50%. Given the simplified and preliminary nature of this study, this shows promise for future developments. The joint geometry identification method predicted the plate dimensions and bolt hole locations with acceptable accuracy, but the radius of each hole was predicted with significant but predictable error. The joint location identification technique obtained accurate results for most members of the analyzed truss. Invalid assumptions resulted in significant error for some individual members. Regardless, these three preliminary studies demonstrate the possibility for future developments resulting in the automated identification of structural steel components without a priori knowledge.

References


and Informatics. 26:90-102, 2009.


Near-Miss Accident Detection for Ironworkers Using Inertial Measurement Unit Sensors

Sepideh S. Aria, Kanghyeok Yang, Changbum R. Ahn*, and Mehmet C. Vuran
a Master’s Student, Computer Science and Engineering Department, University of Nebraska–Lincoln;
b PhD student, Construction Engineering and Management, University of Nebraska–Lincoln;
c Assistant Professor, Construction Engineering and Management, University of Nebraska–Lincoln;
d Associate Professor, Computer Science and Engineering Department, University of Nebraska–Lincoln;
*Corresponding Author (cahn2@unl.edu, United States)

Abstract -
In the construction industry, fall accidents are the leading cause of construction-related fatalities; in particular, ironworkers have the highest risk of fatal accidents. Detecting near-miss accidents for ironworkers provides crucial information for interrupting and preventing the precursors of fall accidents while simultaneously addressing the problem of sparse accident data for ironworkers' fall-risk assessments. However, current methods for detecting near-miss accidents are based upon workers' self-reporting, which introduces variability to the collected data. This paper aims to present a method that uses Inertial Measurement Unit (IMU) sensor data to automatically detect near-miss accidents during ironworkers' walking motion. Then, using a Primal Laplacian Support Vector Machine, a developed semi-supervised algorithm trains a system to predict near-miss incidents using this data. The accuracy of this semi-supervised algorithm was measured with different metrics to assess the impact of the automated near-miss incident detection in construction worksites. The experimental validation of the algorithm indicates that near-miss incidents may be estimated and classified with considerable accuracy—above 98 percent. Then the computational burden of the proposed algorithm was compared with a One-Class Support Vector Machine (OC-SVM). Based upon the proposed detection approach, high-risk actions in the construction site can be detected efficiently, and steps towards reducing or eliminating them may be taken.

Keywords: Sensing and Communication, worker safety, near-miss, Inertial Measurement Unit sensor.

1 Introduction
The construction industry is still dangerous, accounting for about 21% of fatal injuries in the United States[1]. Among these fatal injuries, falls to a lower level have been ranked as the foremost fatal accident type in the construction industry, representing 33% of all fatal accidents [2]. In order to reduce the number of fatal fall accidents, the Occupational Safety and Health Administration began regulating the use of accident-prevention measures (e.g. personal fall arrest systems); however, the administration cannot address certain types of accidents that occur due to dangerous circumstances [3]. Among construction trades, ironworkers have been exposed to the highest lifetime fatal accident risks [4]. However, estimating the fall risk of an ironworker is still very challenging due to the sparse amount of detailed information on actual fall accidents. Thus, there is insufficient knowledge to give forewarning to potential subjects of fall accidents.

For this reason, the identification of near-miss accidents has been brought forward as a consideration that could help prevent future accidents in the construction industry [5], [6]. According to Phimister [7], a near-miss accident is defined as an event that did not cause any harm but that had the potential to become an accident under slightly different conditions. The logic runs that behind one major accident there are numerous near-miss accidents and a few minor accidents [8]. Thus, data about the number of near-miss accidents could be used as a harbinger of an upcoming major accident in general cases.

In this context, this research proposes a method for detecting the near-miss accidents of ironworkers that uses wearable inertial measurement unit (IMU) sensors.
containing accelerometers, gyroscopes and magnetometers. As discussed in previous publications, laboratory experiments were conducted on two unskilled ironwork subjects. In order to detect near-miss accidents, this research classified the condition of the workers’ postures during a movement and proceeded to detect the near-miss incidents observed while the subjects walked. Machine learning algorithms were then applied (i.e., Support Vector Machine (SVM) and OC-SVM) to increase accuracy of predicting of near-miss accidents while decreasing computational burdens.

2 Background

In order to increase the safety of construction workers, diverse technologies (e.g., vision, sensors) have been applied to the construction industry. Vision-based methods utilize single or multiple video- or vision-cameras to track a worker or piece of equipment to prevent exposure to dangerous situations [9]–[12]. These methods investigated appropriate algorithms to observe workers and equipment on construction sites using individual or multiple vision apparatuses. However, in construction sites there are many obstacles that can interrupt devices’ vision. In addition, near-miss accidents are normally subtle incidents that are challenging to detect using vision-based methodologies.

With regards to monitoring a construction worker, different types of sensors have been applied to acquire informative data to classify workers’ activities and behaviours. Joshua [13] applied accelerometers to classify workers’ masonry activities in order to investigate workers’ productivity. Taneja [14] investigated inertial measurement unit (IMU) sensors for location-tracking in a building site as compared to other sensors that used established local area networks (WLAN) and radio frequency identification (RFID). Various types of sensor have been used for workers’ activity classification and behaviour monitoring in research; however, this research is one of the initial attempts to use sensors for near-miss accident detection in the construction area.

Our previous study [15] applied IMU sensors to classify workers’ postures and motions and to detect near-miss accidents for ironworkers. Data was obtained from one test subject. Due to scarcity of near-miss incidents, we faced limitations in training the best classifying function. We preferred an OC-SVM in the previous research because it was mostly successful in training a classifying function. In order to increase the feasibility of this near-miss detection approach, this research utilized a different algorithm and compared its accuracy with previous research results. This work is based upon the detection of ironworker’s fall incidents using IMU sensors, which have high sensitivity to capture the subtle difference between normal and near-miss accidents. This research conducted a preliminary laboratory experiment and used worker’s motion data to implement the near-miss accident detection approach.

3 Research Objectives & Methodology

Extensive growth in the construction field has resulted in an increasing demand for improving the safety of construction workers. This call for safety improvement in turn gives rise to the application of state-of-the-art machine learning algorithms on data. A class of promising algorithms for classification purposes is semi-supervised learning. Semi-supervised learning algorithms estimate a target classification from a few labelled examples alongside a large collection of unlabelled data. In this research, we aim to use a semi-supervised learning algorithm and IMU sensor data to develop a system for detecting ironworkers’ near-miss incidents. This objective is achieved through five phases: (1) setting up the test bed, (2) collecting data from IMU sensors and videotaping the data collection period for generating labels, (3) analysing data to get the desired frequency and extracting features, (4) training the classifier and, finally, (5) evaluating the performance of the classifier function. Figure 1 illustrates the overall research steps.

Figure 1: Near-miss-incidents detection research phases

During the data collection phase, an IMU sensor was attached to the subjects’ waists and collected data on the workers’ motions. Motions included walking and unintended near-miss activities on a steel frame. Subjects were walking on the frame for a few minutes with a steady speed. This rectangular frame was comprised of two four-inch flange I-beams and another two two-inch
width steel beams. The dimensions of this frame were 12 feet 1 inch by 6 feet 6 inches (see Figure 2-b). Videotaping the data collection period aided with the assignation of appropriate labels for deciding whether the ironworker was in a safe condition or was experiencing a near-miss incident.

The type of IMU sensor used to collect the ironworkers’ motion data was a SHIMMER 9DoF with three axes each for the accelerometer, gyroscope and magnetometer (See Figure 2-a). IMU sensors recorded motion data at a frequency of 51.2 Hz and transferred the data to a laptop computer via Bluetooth. Concurrently, a video recorder filmed the experiment to create reference data labels to assist in training the classifier. To keep track of the beginning and end of the data collection process and to synchronize collected motion data with video recording, an impact was given to the sensor at both the start and end of data collection. Figure 2 illustrates the experiment equipment and layout.

![Figure 2: Laboratory experiment layout and equipment](image)

From raw data received from IMU sensor, we observed a frequency of 51.2 Hz. However recorded video of the experiment didn’t have this frequency since objective labelling cannot reach the same sampling speed of IMU sensor. In order to match the frequency of the data received from the IMU sensors and the labels produced with the monitoring video recordings, we reconstructed both raw data and video labels by a sampling rate of 32 Hz with 50% data overlap. Within each sample window, data was sampled using the features for both the accelerometer- and the gyroscope-measured data points present in that window. The accelerometer and gyroscope each have x, y and z axis, hence there were six values for each sampling technique per each data point. In each sampling window, we measured mean, standard deviation and peak values per x, y and z axis, for both the accelerometer and gyroscope. This makes up 18 features. Another 20 features are extracted by using correlation, spectral entropy, and spectral centroid functions. Considering all of these features, 38 features were extracted.

Considering that workers’ near-miss incidents will create irregular patterns in IMU sensor data, the detection of near-miss incidents can be formulated by training a classifier function, which differentiates signal patterns that do not conform to expected signal patterns—in this case, those signals that are departures from workers’ stable postures during a movement. A near-miss is a motion in which the subject loses balance slightly. Finding near-miss incidents therefore depends on subjective labels assigned by reviewing recorded video from the data collection period. However, subjective labelling is not feasible in all circumstances and is computationally expensive. Additionally, due to the rare and dynamic nature of near-miss incidents, it is difficult to record labelled data that includes all sorts of conditions leading to a near-miss incident. Therefore, we favoured a semi-supervised classifier function that can benefit from both labelled and unlabelled data. In order to evaluate the performance of the classifier function, data was divided into 60% training data and 40% test data. Accuracy in predicting near-miss incidents against the total number of data points and the computational time elapsed to achieve this accuracy were measures to assess the performance of the classifier.

4 Identification of near-miss incidents

Recently, many algorithms have been proposed to enhance the quality of a semi-supervised classifier function [16], [17]. The premise of semi-supervised learning is that a marginal distribution of a decision boundary can be estimated based upon labelled data, and that each point of a cluster can be distinguished from data points belonging to other clusters by a curve that separates a dense area of a cluster from a non-dense boundary area. The boundary between different classes is not as dense as the area in each class. This characteristic of the semi-supervised learning algorithms is called the Cluster Assumption. Another assumption in this algorithm is the Manifold Assumption, which states that the boundary lies on or near low-dimensional manifold and that the classifier function moves smoothly along this boundary. The Manifold Assumption produces low-dimensional space using key class features, and therefore is effective in the applications where data has noise.

In this paper we focus on a Primal Laplacian Support Vector Machine (LapSVM) approach. The original Laplacian Support Vector Machine (LapSVM) was proposed by Belkin et al [18]. The original algorithm of LapSVM had a dual formation, which was defined by a number of dual variables equal to l, the number of labelled points. In LapSVM, if the number of labelled data points is l and the number of unlabelled data points is n (where usually n ≫ l), then the relationship between data points is found by a linear system of n equations and variables. Belkin et al. [15] also proposed the Manifold Regularization method, which is based on the geometry
of marginal distribution: Assuming that the probability of the distribution of data has a Riemannian manifold (say, \(M\)), labels of two points that are close to one another in the intrinsic geometry of \(P_x\) will be similar or the same sense the conditional probability distribution \(P(y|x)\) should change little between two such points. LapSVM follows the principals behind manifold regularization with hinge loss function [19]. Hinge loss function forgives noise introduced to the training data.

In recent years there has been a major focus on employing a primal approach to solve nonlinear LapSVM problems rather than the dual approach that was being used in the original LapSVM. The original—or dual—LapSVM requires two steps in training, but using the primal form allows us to collapse training to a single step. This is done by setting a maximum number of iteration and checking the stability of the classifier after every few iterations of training. The classifier function is finalized when a maximum number of iterations or stability is reached. The advantages of the primal approach over the dual approach are the use of the greedy technique in building the classifier, an efficient solution to the original problem with no need for variable switching, and the faster computation of the approximate solution with an unseen pattern in it in combination with a priori-specified probabilities. For better evaluation of classifying boundaries shaped by a trained classifier, the test set has to include data points from both normal and near-miss incidents. Therefore, test data was picked from a data set that includes both types of data points.

5 Results

In our experiment, training and test data was selected from both labelled and unlabelled data. Each data set included 973 data points after sampling raw data points. Labelled validation data provides more information to compare than stable condition data. If a classifier function is trained using only stable condition data, it results in a non-informative decision system. However, validation data introduces more information to the decision system. The solution for the Primal LapSVM efficiency problem was generated by employing the Preconditioned Conjugate Gradient (PCG) [18]. PCG is an iterative algorithm that finds the numerical solution for a linear system containing many variables. A LapSVM is an example of such a system, and the decision boundary is an efficiency problem that we solved using PCG. We set the maximum number of iterations to 1000 iterations and set the classifier function to keep checking the stability of the classifier every three steps to determine whether the classifier was precise enough to quit the training process. Whether the classifier is precise enough or not depends upon measuring the error rate and stopping the training once the error rate is less than the user-defined value. In this experiment, the boundary value was set to 0.3.

Results show that for both subjects, the classifier could successfully detect all stable conditions. For a
sampler size of data, a Radial Basis Function (RBF) kernel showed a better decision boundary when compared to Linear or Polynomial kernels (See Figure 3). RBF in nature usually acts as a low-band pass filter, it acts by smoothing decision boundaries formed by hyper plane defined by support vector machine. This smoothing is at the cost of some loss estimated by loss function, meaning we would allow some outliers in the decision boundary for sake of keeping hyper planes as smooth as they can be. Whether we can use it for different fall-related studies depends on many factors. We can try different kernels and based on “out of sample” results from cross validation we can choose the best kernel. Another factor would be computational complexity. Linear kernels usually compute much faster than radial or poly kernels. Data received from each test subject was processed using an RBF Kernel to train the classifier function while the distance between data points were measured by Euclidian distance. In our previous research [13] we used an OC-SVM in which the classifier was trained only on the positive class; however, in this research, the training and test sets were used without pre-processing in the sense of normalization. In both studies, the data set was divided to 60% training set and 40% test set. We used the PCG method to solve the LapSVM optimization problem rather than Newton’s method to prevent unnecessary iterations [18]. We can see that for both test subjects, Primal LapSVM using PCG optimization produced a classifier in a shorter amount of time compared to the dual approach used in the OC-SVM. Also, the experiments showed that Primal LapSVM using the PCG method achieved the same accuracy as before, if not better. While the error rate for both test subjects was satisfactory, the algorithm also resulted in a faster training process (See Table 1 and 2). A more in-depth look at the accuracy of the Primal LapSVM is presented in table 3. Comparison between Primal LapSVM and OC-SVM is based on previous research [17] and was performed only on subject 1. For this comparison we measured Precision and Recall.

Since we aimed to detect near-miss incidents—which are rare—we need a metric that would compare the number of detected incidents against the total number of incidents in the data set. The Negative Predictive Value compared the total number of true near-miss incidents in the data set to the total number of data points classified as near-misses by the classifying function. (As this value gets closer to 100%, fewer near-miss incidents are wrongly classified as stable by the classifier function.) The error rate was counted as the count of incidents where the prediction of the classifier function didn’t agree with real-world data.

<table>
<thead>
<tr>
<th>Table 3 Performance of Primal LapSVM on Two Test Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
</tr>
<tr>
<td>Precision</td>
</tr>
<tr>
<td>Recall</td>
</tr>
<tr>
<td>Negative Predictive Value</td>
</tr>
<tr>
<td>Error Rate</td>
</tr>
<tr>
<td>Training Duration (Sec)</td>
</tr>
</tbody>
</table>

6 Conclusion and Future Work

In this research, we improved previous methods of automatic detection of near-miss incidents [17] by both enhancing the level of accuracy of the trained classifier and decreasing the computational complexity involved in the training. Using a Primal LapSVM, the solving optimization problem for the classifier function was reduced from cost of $O(n^3)$ to $O(kn^2)$. Significant improvements in memory consumption and the time spent on generating an approximate classifier raises hope for applying greedy techniques for incremental classifier building in future. This study focused on near-miss accidents during walking motion as an advocate for
measuring the success of the proposed algorithm. Study showed very promising near-miss accidents detection. When used with different movements, a two level classification is required. The first level of classification aims at detecting each action. Second level of classification will detect near-miss accidents specifically trained for that motion, which is left for the future work. However, this new model proposed in this research does provide the construction industry with an opportunity to improve safety and identify fall accidents before they actually happen.

7 Acknowledgement

The authors would like to acknowledge the excellent support and technical assistance from Cory Lyons, Project manager in Davis Erection Co., Inc. of Lincoln, NE, who helped design and conduct the experiment. We also thank Dr. Jiayu Chen, a Post-Doctoral researcher at the University of Nebraska-Lincoln, and Tim Adams, an undergraduate student at the University of Nebraska-Lincoln, for their considerable help in conducting experiments and collecting data. This work was financially supported by the Nebraska Research Initiative. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Nebraska Research Initiative.

8 References

Localization inside Tunnels Using Machine Vision

Y.M. Hsieh* and Y.C. Liao*

*Department of Construction Engineering, National Taiwan University of Science and Technology, Taiwan
E-mail: ymhsieh@mail.ntust.edu.tw

Abstract -

Localization (position tracking) inside tunnels is difficult. GPS signal is not available inside tunnels, and installation of wireless access points is expensive. In this work, we propose the use machine vision for localization in tunnels. The advantage for such approach is that it requires no installation of any infrastructure, needed with radio triangulation approach. Prior researches have shown that machine vision can successfully identify positions in both outdoor and indoor environments. However, such application has never been tried inside tunnels. In this paper, we present our work on developing appropriate algorithms for localization in tunnels. One potential application for localization in tunnels is to help tunnel maintainers acquire accurate location information in tunnels when they do tunnel inspections and maintenances.

Keywords -
Localization; machine vision; evaluation; tunnel

1 Introduction

Tunnels need to be inspected, monitored, and maintained during their lifetime. Without proper inspection and maintenance, accidents may happen and endanger its users. In 2006, there was a fatal accident in the Central Artery Tunnel in Boston, USA. The ceiling fell off and caused death. The accident could have been avoided with proper inspection [1]. Similar tragic accident happened in Tokyo-bound Sasago Tunnel, Japan. Emergency inspections conducted after the incident identified 16 similar defects out of 59 tunnels of similar type [2]. Both accidents might have been avoided with proper inspection. Therefore, it is important to conduct proper inspection, monitoring, maintenance to ensure tunnel safety.

In this work, we try to develop a positioning or localization system inside tunnels to assist tunnel inspectors. The tunnel inspectors drive or walk through the tunnel to be inspected. During this course, inspectors pay attentions to spots that have visual anomalies or past issues. They then record the spot by taking notes, pictures, or voice recording. These records are automatically tagged with its whereabouts inside tunnels with the system to be developed. Thus, the tunnel inspector can focus on his inspection work without needing to identify the position for the record. Thus, the system can help inspectors be productive, reduce human errors in recording positions, and hopefully can achieve better inspection quality.

In the past, automated localization in closed space such as mines and tunnels can be done through wireless sensor network [3]. Such technique, however, requires setup of wireless sensor nodes. This leads to an initial cost for such system. The cost includes sensor nodes and their batteries. The system also needs to be maintained, including replacing faulty nodes and charging batteries. Alternatively, one can use Wi-Fi to achieve localization with the same technique. However, using Wi-Fi needs to setup wireless hotspots and their wiring for electrical power. In addition, this hotspot infrastructure also needs to be maintained.

To overcome above-mentioned shortcomings, we propose using machine vision to construct automated positioning system. The proposed method requires no prior installation of infrastructure. Nor does it need prior knowledge of the environment (i.e. no need for “figure-prints”). In this paper, we first present the theoretical background for applying machine visions in localization. Then, we introduce a research platform for localization in tunnels. Finally, we share issues for localization in tunnel and give remarks for out platform.

2 Localization using Machine Vision

The basic idea behind localization using machine vision is illustrated in Figure 1. When the user is at position A, the stereo camera (formed by two cameras) picks up some feature points. Let us assume one of the feature point is point P. With proper setup of the stereo camera, the three-dimensional (3D) coordinates with physical dimensions can be calculated using the approach later described. The calculated coordinates’ origin is
located at the centre between two cameras that constitute the stereo camera. Now, suppose the user moves to position B. The stereo camera picks up another set of feature points. If the feature point P is still picked up by the stereo camera, then point P will have another coordinate because the origin has moved with the stereo camera. Using the original and updated coordinates of point P, user’s movement can be then tracked. The method will not work if there are no overlaps of feature points. Therefore, the system needs to continuously “see” and track features in the tunnel. Such approach has been proven successful on Mars [4], outdoor environments [5] and indoor settings [6]. However, our work seems to be the first attempt in tunnels.

Figure 2 illustrates the overall procedure for localization using machine vision. These steps are explained in subsequent sections.

2.1 Camera Calibration

This work uses stereo vision formed by two cameras that are aligned in frontal-parallel configuration. Each camera must be calibrated first to obtain parameters needed for later steps. This calibration serves three purposes. First, to under how physical world points are projected into pixels of obtained images in each camera. Second, to correct distortions caused by each unideal cameras. Third, to correct deviations from the frontal-parallel assumption. This means images from both cameras need to be transformed so that they are co-planar and row-aligned.

The camera calibration is conducted using procedures recommended in OpenCV [7-8]. The procedure involves the following steps: 1) prepare a chessboard image, 2) obtain at least five different sets of images from the camera setup, and 3) invoke the calibration algorithm. At the end of calibration, four important sets of parameters are obtained. The first set is camera intrinsic matrices [9] (one matrix for each camera) for projecting 3D physical world points onto image plane on camera sensors. The second set is distortion vectors for correcting lens distortions [10-12]. The third set of parameters is called the essential matrix. This transformation matrix transforms image plane of the right-hand-side camera to that of the left-hand-side camera. The final set of parameters is called the fundamental matrix that transforms pixels of the right-hand-side camera to corresponding pixels of the left-hand-side camera. The essential matrix & fundamental matrix is calibrated using Bouguet’s algorithm [13].

It should be noted the calibration image must on a rigid and flat plane to be ideal. Otherwise, assumptions about the reference chessboard image are violated and resultant calibration parameters will not be accurate. Inaccurate calibration parameters would lead to errors in 3D coordinates, affecting accuracy of localization. It turns out using tablets computers such as iPad to present these reference images are rather ideal since it is perfectly flat and it is rigid.
2.2 Image Acquisition

In our research, we take two different approaches for image acquisition. The first approach obtains images from physical cameras, so that the implemented system can be tested in the physical world. The second approach obtains images from virtual cameras, calculated using computer graphics technique. The second approach allows us to obtain perfect images free from camera-lens distortion and improper camera alignment. Further, we have power controlling many factors that cannot be controlled easily in the real world. These factors include lens distortion, camera alignment, lighting condition, etc. As a result, our platform can be used to study how each factor contributes to the errors in localization.

In our work, the physical image acquisition is done using commodity webcams. Their acquisition costs are low, and they are supported on both Windows and Linux platform. Commodity webcams are connected to computers through USB. USB interfaces have limited bandwidth, and therefore limit the frame rate of acquired images. Fortunately, prior researches suggest higher frame rates do not necessarily give better results, 1 – 2 FPS (frames per seconds) should be sufficient.

There are several issues associated with images obtained from physical cameras. First, they have image distortion problems due to less-than-ideal lens. This is especially true for commodity webcams. Second, the acquired images are often “noisy” under low light conditions. The noise may interfere stereo-pair matching is later steps. In addition, some webcam softens and compresses acquired images in order to reduce noise and/or to reduce needed bandwidth between webcams and host computers. This internal processing introduces artefacts in acquired images. As a result, they may affect the accuracy of formed stereo images. These are the reasons why we also use virtual cameras to generate ideal images to evaluate issues in the localization system in development.

2.3 Image Processing

After the images are acquired from cameras, each image goes through a three-step process: un-distortion, rectification, and enhancement.

Un-distortion uses distortion vector obtained in earlier calibration to correct for lens distortion. The accuracy of resultant images depends strongly on the accuracy of calibration.

Rectification corrects images from both cameras to be in the frontal-parallel configuration. This step uses the fundamental matrix obtained in calibration to achieve the desired configuration. Similarly, how well this rectification is depends on the quality of calibration.

Enhancement processes undistorted and rectified images with image-processing techniques. These techniques are often found in photo-editing software such as Photoshop. We use image-processing techniques to help enhance features so that identifying feature points can be easier. This step is crucial in tunnels. This is because in subway tunnels and railway tunnels, they tend to have low light condition. The acquired images, without image processing, do not possess enough feature points. The system or the method of tracking movements using stereo vision simply breaks without sufficient number of feature points. Therefore, this enhancement is key to successful movement tracking inside tunnels.

2.4 Stereo-Pair Forming

Stereo-pair forming is performed after images are obtained and processed from stereo camera, which consists of left-hand-side (LHS) and right-hand-side (RHS) cameras. Stereo pair forming consists of the following steps. 1) Identify feature points. 2) Match feature points. 3) Filter matched pairs. 4) Compute 3D coordinates. These steps are explained subsequently.

Identify feature points. This step tries to identify some characteristic points in LHS and RHS images. These features are called feature points or key points. These points will be used to track personnel movements in tunnels. This is also the reason why we may want to conduct image enhancement to help identify these characteristics points. They can also be regarded as dynamic markers in the context of augmented reality (AR). AR typically use markers to help AR system identify a particular location to present some information at that specific location. In our system, we dynamically pick these markers from the scene, so that no prior installation of any marker is needed for movement tracking.

There are several algorithms for identify feature points, such as SIFT [14] and SURF [15], etc. Each algorithm uses different “descriptor” to describe a feature point. In our work, we mainly use SURF and SIFT for identifying feature points. They offer scale invariance, which is important for tracking feature points that involves camera movement.

Match feature points. After feature points are identified in both LHS and RHS images, they are matched basing on their descriptors. Each matched pair of feature points are supposedly the same point in the physical world that is projected to image planes of LHS and RHS cameras. At later steps, the 3D coordinates of these matched point pairs will be calculated. In our current implementation, we use brute-force matcher to try to find as many matches as possible. As a result, one feature point in the LHS image may find many corresponding points in the RHS image because their descriptors have little difference. Thus, one feature point
in LHS image may result in many point pairs. Each point pair has one feature point in the LHS image and one feature point in RHS image. These two points in a point pair have the same or similar point descriptor.

Filter matched pairs. The point pairs obtained from matching descriptors of feature points may be wrong. As one point pair will give a marker for tracking movements, they need to be correct. It is thus necessary to filter out wrongly matched feature points. Otherwise, the tracking of movement will be invalid. Currently, we use rather simple rules to conduct this filtering. 1) The pairing should have one-to-one correspondence. If one feature point in the LHS image matches multiple points in the RHS, and vice versa, it is filtered out. 2) The matched feature points should have similar vertical pixel coordinate (y-coordinate in the image). It should be noted there might still be invalid point pairs after this filtering. The platform that we build allows us to test effectiveness of filters with virtual tunnels. This platform enables us testing different algorithms or procedures without field trips.

Compute 3D coordinates: After point pairs are formed and filtered, disparity in each point pair is computed. Each point pair uses triangulation in Figure 3 to compute a 3D coordinate corresponding to a point in the physical 3D world. For feature points in each pair, the pixel coordinates and the computed disparity can be used to “re-project” the feature point into 3D coordinates in physical world with the use of intrinsic parameters for each camera. Supposedly, the re-projected 3D coordinate from LHS feature point and RHS feature point in a point pair are the same. However, due to possibly mismatched point pairs, these re-projected coordinates may be different. This is known as re-projection error. Then, feature-point pairs with large re-projection errors are dropped. Finally, we have a set of feature point pairs with 3D coordinates.

2.5 Movement Tracking

After stereo-pair forming, a set of feature point pairs is obtained with 3D coordinates. When we have two such sets at different time, we can track movements in this period. However, this tracking can only be done when these two sets of feature point pairs have intersections. In other words, some feature point pairs exists in both sets.

The same technique for forming stereo-pair is applied to match feature point sets in two different time. In other words, match feature point descriptors. This is an important factor contributing to the popularity of SIFT and SURF descriptors. They claim to be robust against transformations (scales, rotations), enabling tracking feature points obtained at different time.

Once matching stereo-pairs at different time is completely, the affine transformation between these two sets of stereo-pairs (obtained at different time) can be calculated. The affine transformation matrix between them can be calculated. Then the movement can be inferred from the affine transformation matrix. There are six independent degrees-of-freedom (three translations and three rotations) that need to be determined in the affine transformation. Therefore, at least six correctly matched stereo-pairs in two different time are needed. If more stereo-pairs are matched, least square method or QR decomposition can provide better solutions or accuracies. However, because feature matching can be incorrect. It is better use methods such as RANSAC (RANdom Sample Consensus) to help detect and eliminate effects from outliers.

Once the affine transformation matrix is determined, the translation component is used to track movement occurred to the camera, which is equivalent to the movement of the system user.

3 Implementation

We use C++ language to construct our research platform for localization in tunnels. The platform uses two well-known class libraries: OpenCV [7] and VTK [16]. Using these libraries greatly accelerates the development of our platform.

OpenCV is an open-source computer-vision library. It runs on Windows, Linux, MacOS, iOS, and Android. Three main sub-libraries in OpenCV are used: calib3d, features2d, and highgui. We use calib3d for camera calibration and 3D re-projection, features2d for finding feature points in images, and highgui for graphical user interface. Functions used in OpenCV are summarized in Table 1.

![Figure 3. Triangulation using a feature point pair [7]](image)
Table 1. OpenCV function used

<table>
<thead>
<tr>
<th>Image acquisition</th>
<th>Calibration</th>
<th>Image Processing</th>
<th>Feature detection</th>
<th>Position tracking</th>
<th>GUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>cvCaptureFromCAM</td>
<td>findChessboardCorners</td>
<td>remap</td>
<td>SurfFeatureDetect or::detect</td>
<td>solve</td>
<td>Imshow</td>
</tr>
<tr>
<td>Capture images from webcams</td>
<td>Finds the positions of internal corners of chessboard images.</td>
<td>Correct images to be free from lens distortion and be in frontal parallel configuration</td>
<td>Feature point detection using SURF</td>
<td>Use QR factorization to solve for affine transformation matrix</td>
<td>Show images</td>
</tr>
<tr>
<td>Calibration</td>
<td>cornerSubPix</td>
<td>GaussianBlur</td>
<td>SurfDescriptorExtractor::compute</td>
<td>estimateAffine3D</td>
<td>waitKey</td>
</tr>
<tr>
<td>Refines the corner locations to get sub-pixel accuracy.</td>
<td>Blur images to have less noises</td>
<td>Compute SURF feature descriptors</td>
<td>Feature point detection using SIFT</td>
<td>Use RANSAC to choose the best affine transformation matrix</td>
<td>Wait for key strokes</td>
</tr>
<tr>
<td>stereoCalibrate</td>
<td>stereoRectify</td>
<td>addWeighted</td>
<td>BFMatcher::match</td>
<td>Position tracking</td>
<td>drawMatches</td>
</tr>
<tr>
<td>Calibrates stereo cameras. Obtains intrinsic matrices, distortion vectors, essential and fundamental matrices.</td>
<td>Computes rectification transforms for each head of a calibrated stereo camera.</td>
<td>Enhance edges in blurred images</td>
<td>Brute-force feature-point matching</td>
<td>Draws matched feature points between images</td>
<td></td>
</tr>
</tbody>
</table>

VTK is an open-source class library for visualization. It supports Windows, Linux, and MacOS. It is used to create 3D virtualized tunnels with texture mapping. The texture was acquired using high quality single lens reflex (SLR) digital cameras under controlled lighting. Once the virtual tunnel is created, two virtual cameras are placed at desired configuration and locations to simulate LHS and RHS images acquired by cameras. These simulated images are shown in Figure 4. They represent images obtained under idealized conditions. Furthermore, we can evaluate the quality of calibration algorithms by looking at camera parameters, since they are controlled and known for these virtual cameras.

In addition to the above class libraries, we also used OpenMP [17] for multithreading. Two threads are used in the developed system. Image acquisition and processing for LHS and RHS cameras are processed by their own threads. Doing so effectively reduces processing time.

4 Demonstration

Figure 5 shows the computer generated reference chessboard for calibrating virtual cameras. Doing so allows us to inspect the quality of calibration of all camera parameters, including intrinsic matrices, distortion vectors, essential matrix, and fundamental matrix. The platform also enables us to determine the most convenient yet effective poses of the chessboard to calibrate these parameters.

Figure 6 shows the rectified images obtained from cameras (either virtual or physical ones). Showing this image can confirm that the calibrated essential matrix is correct, and after rectification the images from LHS and RHS cameras are indeed in front-parallel configuration with row alignment. If the images are acquired from virtual cameras, we can further inspect the intrinsic matrix and check their values to see if they match the settings of virtual cameras.

Figure 7 shows feature points identified and matched before filtering by a chosen algorithm in the system. Each line in the figure connects matching feature point in LHS and RHS images. The original LHS and RHS images are shown in Figure 4. These matches are obtained after rectification, thus all lines should be horizontal if matches are correct. It is obvious that
several lines that are not horizontal, suggesting the feature matching do need filtering.

Figure 5. Ideal reference images for calibration

Figure 6. System showing rectified images from cameras

Figure 7. Showing matching feature points between LHS and RHS images.

Figure 8 shows feature point matching in LHS camera at different time in the same manner as in Figure 8. Since the camera is moved, the matching lines do not necessarily be horizontal. However, the movement in the demonstrated figure is small, and therefore they should be close to horizontal. One can easily spot these wrong matches easily in Figure 8. Figure 9 shows filtered results from Figure 8 using some filter algorithm. It can be seen that the results are significantly improved. However, there are still some mismatches left undetected. Therefore, it is necessary to use techniques such as RANSAC to perform position tracking while filter out outliers at the same time.

Figure 10 shows the validation of the chose position-tracking algorithm. The red curve with cross symbols represents true path of camera movement. The blue curve with circular symbols is the inferred movement from the procedure summarized in Figure 2. The point in the centre of the grid is the starting point. It is seen the current implement for position tracking does not offer good results since the two curves shown in the figure have some deviations. We have identified several issues associated with the position tracking, and are discussed in the next section.

Figure 8. Showing matching feature points from the same camera at different time.
5 Issues in Localization in Tunnels

So far, we have very limited success using the developed system for localization in tunnels. The positions tracked have un-negligible errors. This is evaluated in the virtual tunnel generated using our system. Thus, it is expected using the system in the real world would produce even worse results. Hence, we compiled the issues identified so far. Five issues have been identified affecting the accuracy of localization:

Calibration of camera setup can be inaccurate. During calibration, cameras need to shoot from at least five different positions. These five positions, however, cannot be chosen randomly. They need to “obtain a rich set of views” [7]. Without careful choices of these views, errors of few centimetre can be expected.

Identify feature points in tunnels can be challenging. Comparing to other applications of localization using machine visions, tunnels have relatively little features. Some of these features need to have good lighting conditions in order to reveal. Therefore, if the technique is to be applied in tunnels, additional light sources may be necessary. Image processing techniques have little help.

Feature point matching between LHS and RHS images can be wrong. One key step for tracking positions is having obtain good reference points in physical space. These reference points are obtained by finding good feature points in LHS and RHS cameras and forming stereo pairs. Unfortunately, the stereo pairs can be formed wrongly. It is necessary to use more than just feature descriptors to form stereo pairs. Good filtering strategies are needed in order to form stereo pairs correctly.

Stereo-pair matching can be wrong in two different time. Not only matching feature points in LHS and RHS cameras can go wrong, but also matching feature points in different time can go wrong as well. This feature point tracking is necessary because we need to track movements of cameras. Again, RANSAC technique proved helpful but not enough. Better filtering techniques are needed in order to track movements correctly.

Solved affine transformation matrix from two sets of stereo pairs can be go wrong or inaccurate. After stereo pairs are formed at two different time, affine transformation matrix can be solved. However, if one fails to find enough matching stereo pairs at different time, then the affine transformation matrix cannot be solved. In addition, if some stereo pairs are wrong, then the solved affine transformation matrix results in wrong displacement. Good enough filter algorithm is needed to find outliers in stereo pairs at different time.

6 Conclusive Summary

In this paper, we describe our work on creating a platform to study localization in tunnels. This study seems to be the first attempt in such environment. The research platform enables us to study good algorithms to use in such application before conducting field trials. Class libraries used in our study are briefly introduced and issues that we have encountered are discussed. Hopefully we will get better results in near future.
Acknowledgement

This work is funded by National Science Council of Taiwan under grant number 101-2221-E-011-111-MY2. Authors would like to thank Prof. T. D. Wang of National Taipei University of Science and Technology for providing tunnel-lining images for our study.

References


Abstract -

The tenet of TQM is the use of measurements to monitor an operation in order to avoid poor quality outputs. The construction faces unique barriers to utilize effective technologies to capture critical process data in real time. This paper presents the results of extensive field testing of wireless systems to study the effect of the “noisy” environment on sensing and signal propagation. The findings confirmed the hypothesis that construction activities and equipment impact the performance of such networks. In an experimental application on site, ZigBee network to collect data about temperature at pre-designed locations was integrated with an off-site station. Working with the system led to the conclusion that ZigBee is a viable technology to be implemented on the construction site, reliable and easy to use. The concrete crew receiving real time data on temperature from the embedded sensors were delighted as they have never been provided with such new and important information before.

Keywords - Sensing and Communication; ZigBee Thermocouple; Signal Propagation Pattern; Information Hub

1 Introduction

The 1980’s brought the first comprehensive concept to manage quality within a large organization that was referred to as Total Quality Management (TQM). The central tenet of TQM was the drive for continuous improvement of the operation while delivering high-quality products and services to customers. Lean manufacturing was another concept that received a lot of attention in the 1980’s and depended on the availability of data. Maguad [1] summarized that: “Implementing lean production is facilitated by a focus on measurement.” Of course, the preventive approach to reduce waste (=lean) requires in-process measurements instead of final product inspection. For example, the traditional concrete quality control procedures, like many other construction processes, are completed well after they have ended. For example, while curing temperature and moisture content affects the mechanical properties of concrete [2], limited data in project databases particularly in BIM records is associated to construction of concrete elements. The advances in sensing as well as communication technologies are the backbone of the transition toward a real-time process oriented quality approach. Boyer and Swink [3] articulated that “… the advent of low cost data collection technologies opens opportunities for improved observations of phenomena.” In particular embedded sensors have been introduced to monitor long-term serviceability of concrete structures [4]. Akinci et al. [5] developed a “formalism” for active quality control on construction sites using embedded sensing systems. Gordon and Akinci [6] put emphasis on mobility of implemented technology in order to collect data from the point of interest with limited access during construction. However, the quality of wireless link between nodes is affected by many factors such as: a) communication protocol, b) network configuration, c) site activities and d) weather conditions. The experiments presented in this paper aim to identify the practical challenges of implementing wireless data capture technologies such as ZigBee sensor network to collect selected process data.

2 Process Quality Measurement: The Path to Eliminating Supply-Chain Waste

Bellah et al. [7], between others, reminded us that:” TQM is a philosophy and set of practices that aim to eliminate all forms of waste from all product manufacturing and service delivery processes.” How can waste elimination lead to quality while improve performance? The connection becomes clear when one equates waste to lack of quality not only of the final product but also during production because waste requires resources and adds to the cost of the product thus reducing its value. Canel et al. [8] defined waste as: “… anything other than the minimum amount of equipment, materials, parts, space, and workers’ time, which are absolutely essential to add value to the product or service …” Thus the systematic elimination...
of waste is also a systematic assault on the factors underlying poor quality and fundamental managerial problems. Robinson and Malhotra [9] highlighted this fact: “... quality practices must advance from traditional and product-based mindsets to an inter-organizational supply chain orientation involving customers, suppliers, and other partners.” Perhaps the most essential difference in this transition of traditional activities is a shift from a product to a process orientation.

3 Wireless Sensing Technologies for Construction

Present implementations of ambient environment sensors on the construction process level are limited to on-board instrumentation (OBI). The agile, rugged and ever-changing spatial locations of operation require hardened and flexible sets of electronic sensors to read temperature, pressure, wind, accelerations, vibration, humidity, barcodes etc. Furthermore, the extensive number of supply chains, depending heavily on truck-based delivery methods, created the need for embedding monitoring technologies into construction equipment like trucks, drilling rigs or tower cranes, or even the construction material enabled to communicate data about workflow, quality and safety. Of course, the mobility of embedded sensors not only requires mobile power-supplies but also a wireless communication network. Bae et al. [10] provided guidelines for optimal design of wireless sensor network (WSN) topology in order to maximize the performance of a ZigBee based bridge health monitoring system. They asserted that the WSN performance depends on the material type and object thickness through which the electronic signals have to travel. However, many barriers have to be overcome in order to wirelessly connect many “islands of information” at construction site are: a) Technical hardware and power, b) interoperability problems, c) security issues, d) ownership of data, e) user-friendliness, f) cost and g) ruggedized housing against dust, heat and rain [11]. Attempts to wirelessly collect data at field level started by the introduction of IEEE wireless protocols like 802.11b [12]. Although wireless conduits were backbone of information flow, the communicated data was collected manually using mobile devices [13]. Research has been done to measure capability of available wireless technologies to transfer construction specific information on-site [14] and off-site [15]. While limited in number and scope, results of feasibility tests showed the lack of knowledge regarding propagation pattern of radio signals on construction work areas [16].

Kim et al. [17] divided applications of WSN into two main categories: a) Environmental monitoring and b) identification of a mechanical system through measured system responses. We add the third category as “process waste control” through monitoring key physical variables in order to diagnose any deficiency from specifications in real time.

4 Performance Features of a Site-Based Wireless Sensor Network

As mentioned, establishing an effective wireless data collection system covering a construction site requires knowledge about the behavior of spatially distributed sensors and electronic signal propagation. Site activities affect WSN performance in two principal ways: a) Interference from electromagnetic fields generated by operating engines and b) blockage of line of sight between sender and receiver. In order to mitigate such negative effects, many wireless systems, such as ZigBee, are employing Received Signal Strength Indicator (RSSI) technology. However, fluctuations of RSSI due to signal fading significantly affect various applications’ accuracy [18]. Experiments presented in this section aim to quantify ZigBee signal’s decay in relation to its “root causes” providing critical knowledge to design any ZigBee based on-site tracking and control application.

The root causes have been studied by monitoring quality of a wireless link between a fixed central data logger and a portable node. National Instruments WSN Link Quality Logger is a virtual instrument (VI), programmed in LabVIEW to determine link quality of WSN nodes in an application environment (http://www.ni.com/example/31346/en/). A snapshot of the program interface is provided in Figure 1. The interface can be used to set the sampling frequency, monitoring battery life, and more importantly recording the Link Quality Index (LQI) – a number out of 100.

4.1 Electromagnetic Emission on Site

An accelerating electrically charged object emits packet of electromagnetic (EM) energy in a fixed frequency that creates the so-called EM field. Theoretically, frequency and intensity of the EM fields at site are the two key factors affecting the performance of WSN. However, because of ZigBee’s position in the radio frequency spectrum, the frequency based interference of EM sources with wireless signals is less contingent. But, the undesirable electromagnetic radiation (EMR) still can interrupt network performance through other mechanisms such as: a) Causing network components especially the sensitive electronic chipsets to malfunction, b) increasing signal to noise ratio in the environment especially in antenna proximity, and c)

869
signal attenuation in a weakly ionized medium. Noise emission regulations such as FCC docket 20780 in United States and VDE 0871 and 75 in Germany have provided guidelines to protect digital and electronic equipment against a broad range of 10 KHz to 1 GHz. In particular, focus of VDE 0875 on low-frequency devices up to 150 KHz shows that network gears are potential victim of EMR of the power lines feeding tower crane as well as the heavy equipment engines.

4.1.1 Concrete Pump

First step to field deployment of a wireless monitoring system is to ensure that operation of the equipment required for the process does not threaten network’s reliability. Concrete pump is one of the equipment with potential destructive EM emission. Figure 2 demonstrates test configuration where a mobile node is fixed beside the pump hopper and a truck mixer discharges the concert. The first design of experiment included recording of LQI in two different conditions a) Idle pump and b) pumping in progress. The design was later modified to examine the third treatment when pumping and feeding the hopper happened simultaneously. It was observed that mixer’s engine that runs the hydraulic drives of the drum created higher noise in reverse rotation. That was enough to question whether the change in the direction of the drum rotation to “discharge” the concrete introduces different situation than “charging” rotation of the mixer. Three explained conditions comprise the treatments of the ANOVA. Link quality data for each treatment obtained from the field test is documented in Table 1. The test statistic \( f = 3.82 \) is bigger than the critical threshold \( f_{0.05, 2, 15} = 3.68 \) thus supports the rejection of the null hypothesis of equal means. Although the ANOVA says that there is significant difference among means but it doesn’t say which means differ. The procedure of finding the specific differences is built upon the fact that while similar treatment means are close enough to be covered by the bell curve of normal distribution with \( \sigma = \sqrt{\frac{\text{Mean Square Error}}{\text{Number of Replications}}} = \sqrt{\frac{52.8}{6}} = 3.0 \), the dissimilar treatment means will fall out. In another word if the difference between two means is larger than the width of the normal curve, those two treatments are “different”. That is common to estimate the width of normal curve with 6\( \sigma \) for practical purposes. However,

<table>
<thead>
<tr>
<th>Treatments</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idle Pump</td>
<td>54</td>
<td>60</td>
<td>63</td>
<td>50</td>
<td>55</td>
<td>53</td>
<td>55.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Pumping in Press</td>
<td>54</td>
<td>47</td>
<td>36</td>
<td>41</td>
<td>44</td>
<td>56</td>
<td>46.3</td>
<td>7.7</td>
</tr>
<tr>
<td>Pumping and Feeding</td>
<td>49</td>
<td>39</td>
<td>36</td>
<td>56</td>
<td>38</td>
<td>54</td>
<td>45.3</td>
<td>8.8</td>
</tr>
</tbody>
</table>
because the P-value of the ANOVA (0.046) is very close to test significance level of 0.05, it would be reasonable to consider a smaller fraction of the normal curve for example 3σ. Note, that selected width still contains fairly high percentage (87%) of the area under the normal curve. Now, the difference between treatment means can be compared to identify similar treatments. Following the described procedure, signal performance is similar around an operating pump either fed with concrete or not but is different when pump is idle. The mean signal quality index is 9.5 units larger than working pump’s mean, a difference slightly larger than 3σ (=9) that confirms the effect of the pump operation on the link quality.

5 Deploying Attained Knowledge to Collect Process Data

Development of advanced information management systems in construction promises improved process coordination [19]. However, current practices of manually acquiring field data needs to be automated in order to fully deploy work-integrated information management models [20]. A wireless climatic information system for construction sites developed by [21] is a proper combination of automated collection and wireless dissemination of critical data. Still, WSN is not widely implemented to gather “process” data. Embedded measurement of concrete maturity using thermocouples is a potential application that needs extra investigation.

Heat evolution of concrete mixtures not only affects material properties such as strength [22] and yield stresses [23] but also may cause thermal cracking leading to disqualification of hardened concrete [24]. Furthermore, if available during the construction process, temperature development is an indicator of formwork striking time for concrete structures [25] as well as traffic opening criterion for concrete pavements [26]. Calorimetry test has widely been used to model hydration properties of early-age concrete [27]. However, due to equipment limitations in-situ concrete strength development cannot fully be modeled at the laboratory. For example, while it’s believed that isothermal test has advantages over adiabatic test; Xu et al. [28] showed that hydration parameters calculated from semi-adiabatic test better match actual measurements. Moreover, other construction specific factors such as a) framework, b) workmanship, c) thickness of concrete element, d) heat exchange of concrete with the surrounding environment, e) weather conditions and etc. affecting hardening rate are not included in practical models. Therefore, if possible, field measurement is the most appropriate method to determine temperature development of cast in place concrete.

5.1 Instrumentation and Methodology

As pictured in Figure 3 a programmable NI3212 wireless node at work-front locally interprets T-thermocouple voltages into temperature and sent the measurements to the central NI9792 wireless sensor network (WSN) gateway. However, because concrete pump was the desired central point of the network a careful design was required to wirelessly link the end note to the central data logger. ZigBee was selected as the communication protocol because of its low power consumption and secure networking. Based on the knowledge acquired during the field surveys, a relay node was placed beside the crane to ruggedize the “information conduit” off the work-front. Figure 3 also shows the mobile multifunctional user interface (eCKiosk) developed to provide live temperature data to the job foreman. After pouring, data logging frequency was lowered to one measurement per 10 minutes to save battery life of wireless nodes until the formwork was removed two days later. Real time weather condition was also recorded using data provided on Australian Bureau of Metrology’s website.

A distributed node can be set either as an “end node” or a “mesh router” node. Deployment of IEEE 802.15.4 radio in selected devices facilitates low power communication among as many as 64000 nodes in a network. Router configuration let the end nodes find new path back to the central gateway when existing rout is blocked or signal is lost. However, a router configuration keeps the node continuously awake synonym to higher power consumption. NI Measurement & Automation Explorer (MAX) was utilized to arrange nodes to extend network’s distance as well as tune up reliability versus power consumption. The typical network configuration used during field test is presented in Figure 4. An NI9792 coordinates the network by managing node authentication, buffering...
transmitted packets of data on top of bridging WSN to the internet for information sharing purposes. During preliminary tests the link reliability between NI3212 end node at pouring spot and network gateway beside the pump truck significantly dropped when rebar mesh was installed in the formwork and the node was tied to a wooden timber supporting the formwork in order to be protected against any unpredicted stroke during the process (See Figure 5a). Furthermore, selected concrete pour was inside a three meter deep foundation pit surrounded by in-situ concrete sheet piles supporting the vertical excavation. Prediction of rain for the period of data collection in addition to the confined space with significant height difference which was proven to decay signal strength [29], forced the research team to add a router node in order to enhance reliability of the system. As shown in Figure 4, install of the additional node enhanced end node’s link quality from mostly “Poor” and “no signal” to “Good”. Because of the access to the power outlet, the intermediate node was supposed to work on external power in the initial plan. But, according to safety rules a powered device was not allowed unattended over the night. Therefore, router node was also powered by battery during the experiment. In particular, switching to battery did not affect performance of the network at all. A thermocouple is made by joining two different metals while the contact point may be simply twisted or welded. Any change in joint temperature produces a small thermoelectric voltage in millivolt range that is converted to the temperature using Seebeck coefficient. That coefficient is a function of materials composing the thermocouple and determines the operating range of a specific thermocouple. For example, the T thermocouple embedded in the formwork is made from cooper and as one conductor and constantan (copper-nickel alloy) as another. Similarly, a T type thermocouple with operating range of -50°C to +250°C from another manufacturer was used to measure surface temperature of freshly poured concrete. With typical accuracy of 1°C the thermocouple was a proper choice for fresh concrete quality monitoring purposes. Compatible with selected hardware, the LabVIEW was proper system design software for our measurement and control application. The graphical programming environment of LabVIEW made creating custom software solutions dramatically simple. The code developed to acquire, analyze, display and store the concrete temperature is presented in Figure 6. Adiabatic and surface embedded thermocouples were connected to TC0 (+/-) and TC03 (+/-) terminals of NI3212 correspondingly. The NI3212 thermocouple input node is mentioned as Node3 in the network tree rooted to NI9792 as central data logger. Collected data was written on a hard disk as well as shown simultaneously to the concrete crew through eCKiosk. While regular waveform chart displayed the measurements in real time, two temperature meters were also added to facilitate fast reading of the data. Furthermore, two indicators were considered to notify concrete foreman when each monitored variable exceeded pre-defined thresholds. As mentioned in the methodology, the sampling frequency was set to 0.1 in pouring and then decreased to one measurement per every ten minutes in curing period. A higher sampling rate during pouring was to catch the sudden changes of temperature because of frequent drizzling. Figure 7 depicts the data presented to the crew during pouring process.

6 Summary and Conclusion

Total Quality Management (TQM) depends on proactive measurements to ensure a flawless production with high quality outputs. Poor quality supplies and processes not only lead to break downs and low productivity but can also result in costly repairs when the product does not pass final inspection. However, real time monitoring of construction processes requires a secure stream of data from the work-front collected by
a sensor system. The goal of this paper was to assess the effectiveness of such wireless systems through the experimental deployment of a set of wireless thermocouples to record temperature variation inside as well as on the surface of recently poured concrete. Because the behavior of wireless signals across an active construction site was unknown, a series of pre-
tests were conducted to understand the extend of possible interferences with the communication signals. For example, the effects of a pump under different operating conditions were statistically analyzed. This was followed by field testing an entire Wireless Sensor Network (WSN) coupled to a mobile and multi-purpose hub interfaced with the internet. Designed to monitor concrete construction, it enabled a more flexible condition-based control of the process especially when pouring was frequently stopped and resumed because of rain. Continuous data capturing for three days showed the robustness of the predicted signal propagation patterns used in designing the wireless network. Observations confirmed that ZigBee’s data rate of 250 kbit/s was sufficient for the single transmission from sensor to gateway. On the ever changing construction site, its capability to transmit data through intermediate nodes was effective in overcoming signal attenuation. The experimental field test also demonstrated the potentials of such measurement technologies still novel in construction. The availability of new information about ambient and peak adiabatic temperatures or other weather conditions provides a proactive and nondestructive quality control of the concrete curing process. However, the value of collecting process data is not limited to quality management. While the test focused on temperature data other valuable information can be gained “for no extra cost.” Examples are idleness of resources, or surface drying can indirectly be interpreted from the same data. Of course, there are a myriad of other electronic sensors that would drastically expand the possibilities to ensure high quality of processes and final products.
References


[23] Wu D., Fall M. and Cai S.J. Coupling temperature, cement hydration and rheological behaviour of


Implementation of As-Built Information Modelling Using Mobile and Terrestrial Lidar Systems

S. M.E. Sepasgozar\textsuperscript{a}, S. Lim\textsuperscript{b}, S. Shirowzhan\textsuperscript{c}, and Y.M. Kim\textsuperscript{d}

\textsuperscript{a,d} Engineering Construction and Management, School of Civil and Environmental Engineering, University of New South Wales Australia.
\textsuperscript{b,c} GIS and Lidar, School of Civil and Environmental Engineering, University of New South Wales Australia.
E-mail: samad.sepasgozar@gmail.com, s.lim@unsw.edu.au, shirozhan@gmail.com, peterkim.ym@gmail.com

Abstract - Recently it is observed that there is an increasing trend of using lidar point clouds in construction, since lidar is able to provide highly accurate 3-dimensional representations of objects. The flip side of a lidar point cloud is its massive size and therefore prolonged time for data processing. Our research aims to develop an efficient framework for as-built modelling in terms of time, cost and performance. For this, we collected lidar point clouds using two different data collection methods (i.e. mobile and terrestrial lidar systems with a few centimetre level accuracy) and constructed solid 3-dimensional models of a building. The procedures of creating as-built models from both data acquisition methods are compared to understand the capability of each for automatic building information modelling. In order to create as-built models, a framework consisting of eight key stages from data capturing to constructing the building information model is developed. It was found that the framework using mobile lidar enables contractors to create as-built models for complex objects in a timely manner, whereas the framework using terrestrial lidar provides us with more accurate as-built information models. The implementation results of the two frameworks using mobile and terrestrial lidar systems vary between 5-30 mm and 1-45 mm, respectively. It is anticipated that the proposed study becomes a step forward to full automation of lidar-based building information modelling.

Keywords - As-built; Scanner; Mobile lidar; Terrestrial; BIM, Construction

1 Introduction

An emerging technology such as lidar provides an opportunity to move forward to a cost-efficient way of as-built creation. Construction engineers are willing to adopt such a technology in order to increase the accuracy and speed of the digital as-built generation process. The United States General Service Administration’s Office of the Chief Architect (OCA) has commanded that every federal facility projects should be documented in 3-dimensional (3D) coordinates using the laser scanning technology for acquiring building spatial data [1]. However, the current practices’ accuracy still needs to be improved for as-built documentation purposes [2].

After conversion to a suitable format, the data captured using lidar scanners can be stored in building information modelling (BIM) as 3D objects with the attached attribute information. However, collecting and incorporating reliable as-built data for BIM is still challenging, when considered accuracy, time and cost constraints [3]. According to Giel and Issa [2], there is a great disconnection between the data derived from remote sensing and the creation of as-built using BIM.

Details and specifications of many construction objects that are left behind walls, covered by another item or buried for another scheduling, could not be detected in each stage. The main reason for this problem is that surveying-quality scanners are relatively expensive and difficult to adjust or relocate the equipment through the construction site even by highly skilled technicians. Therefore a tool capable of detecting objects and collecting accurate coordinates of the objects in a timely manner would have a significant value for contractors and contribute to the information flow and modelling in construction.

The importance of as-built for the owners and project managers has been well documented [4]. Several technologies are available for creating as-built models. Recently, lidar is introduced to the construction industry for different purposes [5-7]. The capability of lidar as a tool to collect a large amount of accurate 3D data is investigated and applied using the existing technologies. While many studies were attempted to provide new solutions to collect and record the completed construction objects for as-built purposes, the effort to automate the process of an as-built creation on time is
still in the early stage and very challenging [3, 8]. For example, Dore and Murphy [9] presented a semi-automatic approach for generating BIM façade models for existing buildings, and concluded that future work requires procedural development for generating entire building models, rather than parametric library objects such as column and pediments. In addition, they suggested that future work should be more comprehensive to serve end-users [9]. Therefore there is much work needed to improve the accuracy of the current practices considering cost and time in construction [2]. Recently, improving the level of details in as-built is recognised as an open question [3, 10]. On the other hand, integration of the detected fine objects and the data into BIM as a new paradigm of a knowledge sharing system is still challenging, and the use of appropriate hardware, structure, and ad-hoc algorithm still needs to be enhanced [3, 11].

This on-going study aims to provide a procedural framework to rapidly incorporate 3D lidar point cloud data captured by a mobile lidar into BIM. The main objective of this paper is to develop a rapid as-built model (rABM) using the state-of-the-art mobile and terrestrial lidar equipment. The originality of this paper lies in implementing a framework to use a novel hand-held mobile lidar for real-time indoor data acquisition for creating as-built BIM. The presented framework enables contractors to create as-built BIM in a cost-efficient and timely manner. In addition, comparing two new technologies reveals a possibility for current construction needs for creating as-built for complex buildings in limited and small areas. On the other hand, the comparison assists vendors and innovators to enhance the feature and capability of next generation of scanners to cover limitation of as-built creation modelling barriers.

The paper firstly reviews existing techniques for creating as-built which are mostly manual processes; and identifies technology gaps and barriers to the automation process. Secondly, the framework is implemented and the results are verified using extra independent data sets of a sample educational building at the University of New South Wales, Sydney, Australia. Thereafter, the results of a comparison matrix of two filed operations and results will be discussed. Finally, investigating other sample buildings will be suggested as future work.

2 Overview of As-built Creating Methods

The process of as-built information modelling using new technologies can be divided into two main phases: data acquisition and building information modelling.

2.1 Data Acquisition Methods

Several techniques can be used for data acquisition in order to create as-built in construction. However, there are limited techniques to acquire data with possibility for creating digital modelling.

2.1.1 Traditional Practice

Traditional as-built practices are mainly based on graphical standards for 2D drawings [12]. To develop such illustrations in two or three dimensions, traditional measuring equipment is used. The accuracy of these traditional as-built methods is within the required tolerance [13]. However, this traditional method of data acquisition produces a mass of drawings so that their management and usage are time consuming [14]. For example, Wang and Love [15] explained that traditional site layout method is labour intensive and required many times re-measuring.

2.1.2 Photogrammetry

Historically, sketch and photos are used to supplement traditional method to assist CAD operators to enhance the information and accuracy of traditional as-built planning without any ground control [16]. Recently, researchers try to use photogrammetry techniques to produce digital and parametric data for as-built information modelling. Photogrammetry refers to geometric information derived from photographs [17, 18]. However, this method has limitations [17, 19]. For example, extracting object points from a wide angle shots nearby an object is difficult [17]. This approach is not able to produce the required information about the topography of irregular shapes in detail, and cannot provide the details of curves and irregular shapes, whereas lidar scanners can capture such details easily. Photogrammetry usually cannot be used independently in creating as-built, and it is not an ideal solution for as-built [20]. Recent studies attempt to integrate digital photogrammetry with lidar scanners [16, 21].

2.1.3 Lidar Solutions

Lidar is a laser imaging technology that is increasingly employed for capturing scenes with millimetre to centimetre accuracy. It provides fast, accurate, comprehensive and detailed 3D data about the scanned scenes at the rate of hundreds of thousands of point measurements per second.

Lidar scanners collect data in the form of point clouds which are shapes and dimensions of objects in real space converted and represented as a collection of points in a 3D digital space.
Recently, lidar is widely used for construction purposes [5, 22]. Particularly, several studies attempt to use lidar for as-built creation [6, 23]. However, there are two main problems in this stream. First, geometric information such as lines and surfaces cannot be easily extracted from millions of points data of objects [24], and are recommended as future research [24, 25]. Second, a limited number of scanners such as terrestrial scanners are suitable for BIM [6], and the state-of-the-art of technologies are not investigated fully.

2.1.4 Comparing Possible Solutions

A summary of comparison of as-built methods is provided in Table 1. Corresponding to the challenging problem of reliable data collection of constructed objects in a timely manner, scanners can be proper devices of data collection for as-built creation in construction. The comparison suggests that laser scanners are an accurate and fast solution that will be employed in this paper. It should be noted that the previous scanners used for as-built creation purpose were bulky as it is stated in Table 1. In order to solve these problems, we will use a novel mobile equipment to acquire data.

Table 1. Comparing three as-built methods

<table>
<thead>
<tr>
<th>Item</th>
<th>Traditional</th>
<th>Photogrammetry</th>
<th>Lidar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware</td>
<td>Tools</td>
<td>Portable tool</td>
<td>Scanners</td>
</tr>
<tr>
<td>Portability</td>
<td>Hand held</td>
<td>Hand held [18]</td>
<td>Bulky</td>
</tr>
<tr>
<td>Skill</td>
<td>Low</td>
<td>Low</td>
<td>Medium-high</td>
</tr>
<tr>
<td>Equip. cost</td>
<td>Hundreds</td>
<td>Hundreds [18]</td>
<td>Thousands</td>
</tr>
<tr>
<td>Resolution</td>
<td>Low [17]</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Centimetre</td>
<td>Centimetre [17]</td>
<td>Millimetre</td>
</tr>
<tr>
<td>3D Modelling</td>
<td>Manual</td>
<td>Post-process</td>
<td>Automatic</td>
</tr>
<tr>
<td>Retrieval</td>
<td>Manual</td>
<td>Manual</td>
<td>Automatic</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Accurate</td>
<td>Accurate [18]</td>
<td>Most accurate</td>
</tr>
<tr>
<td>Spatial speed</td>
<td>Not real time</td>
<td>Not real time retrieval [19]</td>
<td>Real time retrieval</td>
</tr>
<tr>
<td>Range</td>
<td>-</td>
<td>Medium [19]</td>
<td>Long [19]</td>
</tr>
</tbody>
</table>

2.2 Building Information Modelling

BIM is a collection of data which digitally represent the relevant characteristics of a building [26]. It is created through a process of modelling which includes practices such as distribution and storage of these datasets. BIM is a rich data platform, but it is not practical if it only represents the design, and not the building as it is built.

As-built BIM includes all the changes made on the building during the construction phase; it can be an up-to-date representation model of the building. The widespread use of as-built BIM is prevented by the lack of a time-efficient and accurate method of easily creating BIMs [8]. The capabilities of BIM as representation models and its requirements for as-built BIMs are well investigated [2, 9, 27, 28]. However, Huber et al. [8] claimed that the topic of representing as-built BIM is in its early stage. Furthermore, current procedures rely on manual processes which are labour-intensive, time consuming and susceptible to errors [3]. According to Tang et al [3] existing work focuses on modelling the simplest objects of a building rather than modelling complex objects such as doors and windows. In addition, they showed that there are significant disconnections between surface-based modelling and volumetric modelling representations [3].

In summary, the review reveals three significant challenges for creating as-built BIM. First, previous work mainly used terrestrial scanners and stationary lidar equipment [2, 3, 8, 9]. Second, modelling fine and complex objects of buildings have received little attention [3, 8]. Third, a procedural approach toward automatic as-built creation of the entire building is still unachievable [9].

3 Methodology

3.1 Study Framework

Using new technologies, innovative algorithms assist us to create semi-automated and rapid as-built modelling (rABM) for creating and updating as-built building information models in each stage of the project. With this intention, this paper attempts to take the first step forward to develop the model to create more accurate as-built with low skill labours in a short time. Based on the literature and previous work, the overall process for as-built creation is proposed consisting of scanning, processing, and creation. In addition, the algorithm and feature of the new hand-held rapid laser which is used in this paper dictated us to follow a new procedure to create rABM. The result will be verified by a data set acquired from a terrestrial lidar.

The rABM consists of eight stages as shown in Figure 1. In the data collection step, one building at the Kensington Campus of the University of New South Wales has been scanned using the mobile mounted range-sensing system, and a terrestrial scanner. Next, the data is registered and obvious noises were removed. Main elements including openings, walls, floors and ceilings were segmented in the next step. Then, the extracted elements were combined as the identified as-built elements. Field work was conducted to assess and verify the level of accuracy obtained using the dense lidar points.
3.2 Device and Algorithm of Data Acquisition

A novel handheld mobile lidar is used to implement the framework. The results are validated using the state-of-the-art terrestrial scanner.

3.2.1 Mobile Lidar

The first set of data was captured utilising a handheld mobile mapping device including a lightweight lidar scanner. The utilised device is a 3D sensor system that consists of a rotating and trawling 2D lidar and a MicroStrain miniature internal measurement unit (IMU) mounted on a spring mechanism. This equipment, namely Zeb1 designed on the basis of a 3D simultaneous location and mapping (SLAM) solution [29], while other mobile mapping methods are based on GNSS that cannot be used for indoor mapping purposes. The laser of the scanner is a Hokuyo UTM with 270-degree field of view, 30 meter maximum detectable range. The total weight of the system is about 510 g. The handheld device has a small battery that is sufficient for the whole operation and a data acquisition laptop which was carried in a backpack by the operator.

3.2.2 Terrestrial Lidar

The second data set was collected using a state-of-the-art multi-station called Leica Nova MS50 at two locations from less than 3 meters distance from the objects. The maximum distance measurement and the maximum range are 50 and 1000 meters, respectively, for this multi-station.

3.3 Details of Study Extent

The study extent and the sample partition on the 4th floor are shown in Figure 2 (a). A part of building with windows, doors and stairs were selected as it is complex enough to explore the accuracy of the work for different architectural elements, i.e. modelling such fine objects and details are still challenging. The general attributes of the case study are represented in Table 2. For consistency in understanding the analysis, pseudonyms are used for the parts such as W4 to refer to the west wing and C4 to refer to the middle corridor of the 4th floor which is being studied.

<table>
<thead>
<tr>
<th>Table 2. The profile of the object/case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Use</td>
</tr>
<tr>
<td>Type</td>
</tr>
<tr>
<td>Scope</td>
</tr>
<tr>
<td>Materials</td>
</tr>
<tr>
<td>Point</td>
</tr>
</tbody>
</table>

4 Results and Discussion

The results of implementation of the framework are separately discussed in this section.
4.1 Mobile lidar

Scanning the indoor environment of two levels of the building using the handheld rapid-laser mapping system took about 20 minutes. The point cloud segment of 4th floor is overlaid on the layout plan of the same place (see Figure 2a) and the open-loop trajectories of the scanner are shown in Figure 2 (b). The sample is shown in A–A view (Figure 2 a) from the north-south corridor of level 4 called C4. Using a segmentation technique, the 3D point cloud of this area was analysed to obtain the required dimensions for parametric modelling. In the current practice, registration is a semi-automated process. Data processing is also a semi-automated process that includes manual and automated filtering to remove noises and unwanted data, such as points from moving objects and reflections.

4.2 Terrestrial lidar

The multi-station terrestrial scanner (Figure 3) was set up at two locations as shown in the layout plan Figures 2a and 3. The process of data collection took about 30 minutes from two locations including stationing and scanning.

In order to compare the result of two different data sets, the same segment of A–A is selected. This data collected by terrestrial lidar contains less noise points compared to the mobile lidar.

4.3 Modelling Process

Pre-processing of the data sets requires firstly, conversion of the text files to .ply and .las files for visualisation purposes and secondly, detection and reduction of noise (i.e. unwanted objects such as furniture). The data sets then were segmented to wall, floor and ceiling. Thereafter, both processed mobile and terrestrial lidar data sets were imported and processed into Autodesk Revit and converted within the program into a compatible format. Two segmented components: wall (including windows and an opening) and a section of stairs of the sample C4 are modelled as shown in Figure 4d. Levels were allocated in components and a model was created.

4.4 Comparison of the Results

Table 3 shows the measurement results from three different methods: traditional, mobile and terrestrial lidar. While the performance of terrestrial lidar is better than mobile lidar in terms of windows measurements, mobile lidar shows better results in measuring the door dimensions.

<table>
<thead>
<tr>
<th>Object</th>
<th>Traditional</th>
<th>Mobile</th>
<th>Terrestrial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stair</td>
<td>30</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td>Door</td>
<td>1840</td>
<td>1810</td>
<td>1780</td>
</tr>
<tr>
<td>Window 1</td>
<td>1680</td>
<td>1700</td>
<td>1690</td>
</tr>
<tr>
<td>Window 2</td>
<td>684</td>
<td>690</td>
<td>683</td>
</tr>
</tbody>
</table>

L refers to length of the object. W refers to the width of the object.

Mobile lidar scanning took about 20 minutes for the entire collection at 3rd and 4th floors. However, operation of terrestrial scanning consists of stationing and scanning for only the wall from one side took 20 minutes. Similarly, stationing and scanning for stairs took about 13 minutes. In total, using multi-station terrestrial lidar took about 33 minutes only for the sample partition including the wall and five of stairs behind the wall. This shows that the mobile hand-held...
scanner is very fast, and does not need a skilled technician for adjusting and using the equipment.

The accuracy of the result of implementations of the frameworks using mobile and terrestrial lidar equipment comparing traditional measures varies from 5 mm to 30 mm and from 1 mm to 12 mm. The results from terrestrial lidar are more accurate and closer to the field work than mobile lidar for small openings.

Another difference between the mobile and terrestrial lidar is cost. The cost of the mobile lidar is three times cheaper than the multi-station scanner. At the same time, using mobile lidar equipment does not need a skilled operator. However, processing the data collected by the mobile lidar needs a skilled expert compared to the terrestrial lidar which is fully commercialised. A summary of the comparison of two experiments is listed in Table 4.

All in all, the results show that the terrestrial lidar is more accurate than the mobile lidar in most of the cases, and from visual inspection less noise points can be seen in the terrestrial lidar data than the mobile lidar data. At the same time, the data collection process of the terrestrial lidar is not as fast and easy as the mobile lidar. Both terrestrial and mobile lidars are accurate enough for construction purposes.

The terrestrial lidar equipment is expensive ($25,000) compared to the mobile lidar ($60,000). In sum, it is preferred to implement the framework using mobile lidar considering the speed of collection from a large area of complex buildings by a low skilled operator. Implementing rABM using mobile lidar, the final as-built modelling in BIM created as shown in Figure 5. However, terrestrial lidar is preferred for fine and specific objects in construction such as windows.

In addition, the prior robotics applications are typically concerned with navigations and detection algorithm; whereas the feature of mobile lidar equipment
enables us to use it in small and limited areas of complex buildings where adjusting a terrestrial lidar equipment is not possible.

Table 4. Comparison of the results of Mobile and Terrestrial lidar

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mobile lidar</th>
<th>Terrestrial lidar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating</td>
<td>Without Adjustment</td>
<td>With Adjustment</td>
</tr>
<tr>
<td>Portability</td>
<td>510 g</td>
<td>7.6 kg</td>
</tr>
<tr>
<td>Required skill</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Estimated cost</td>
<td>About $25,000</td>
<td>About $60,000</td>
</tr>
<tr>
<td>Resolution</td>
<td>Medium</td>
<td>Very high</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Millimetre</td>
<td>Millimetre</td>
</tr>
<tr>
<td>Spatial data Retrieval</td>
<td>Semi-automatic</td>
<td>Semi-automatic</td>
</tr>
<tr>
<td>Spatial data accuracy</td>
<td>±5 to 30 millimetre</td>
<td>±3 to 12 millimetre</td>
</tr>
<tr>
<td>Spatial data speed</td>
<td>Real time retrieval</td>
<td>Real time retrieval</td>
</tr>
<tr>
<td>Range distance</td>
<td>20</td>
<td>1000</td>
</tr>
<tr>
<td>Operation time (for the sample)</td>
<td>10 minutes</td>
<td>30 minutes</td>
</tr>
</tbody>
</table>

5 Concluding Remarks

This study aimed to develop a framework for rapid as-built modelling using 3D point cloud data captured by the state-of-the-art technologies. The 3D lidar scanners are used in order to speed up the acquisition flow of the required information from a raw lidar point cloud. In this study, we implemented the process to obtain dimensions of one complex building. The dimensions are verified using mobile and terrestrial lidar, and the results are more accurate than the current as-built which were created by a construction team, and the accuracy is higher than the previous work at some areas such as openings and stairs. Therefore, mobile scanners are suitable for collecting indoor data for buildings and can be used in conjunction with terrestrial scanners which are able to collect both outdoor and indoor data.

This study shows some significant potential benefits that can be improved by implementing the framework in terms of time and accuracy of creating as-built BIM. However, there is much work to be done to fully automate the whole process of as-built creation in construction.

As for future work, we suggest to select other types of buildings and construction sites such as tunnels, railways and mining objects. In addition, examining the framework for creating staging as-built during construction is needed as future work.

Acknowledgement

The authors appreciate C.R. Kennedy Pty Ltd for their support in providing the second set of data.

References


Three Dimensional Spatial Metrics for Compactness Assessment of Urban Forms

Sara Shirowzhan\textsuperscript{a} and Samsung Lim\textsuperscript{b}

\textsuperscript{a,b} School of Civil and Environmental Engineering, University of New South Wales, Australia
E-mail: S.Shirowzhan@student.unsw.edu.au, S.lim@unsw.edu.au

Abstract – Developing measurement tools to assess urban sustainability is one of the new streams in built environment; however, automated methods of urban form assessment are technically difficult. While planar characteristics of urban forms have been studied traditionally using the spatial metrics from remote sensing data such as Landsat images, height information of urban environments is now playing an important role in urban energy exchange (e.g. solar energy gains), urban air circulation (which is affected by interactions between terrain and buildings) and urban microclimate. As incorporating the planar metrics and the height information has been rarely attempted in the sustainable urban form studies, we propose an automatic technique for the quantification of 3-dimensional urban compactness which is one of the main factors influencing sustainability. In this paper, we aim to utilise one of autocorrelation statistics known as Moran’s $I$ for the assessment of urban form compactness, considering both layout and elevation attributes. Additionally, Getis-Ord $G$ statistic is also used for further investigation on concentration of low or high urban features.

Keywords - Lidar, Sustainable urban form, Urban neighbourhood pattern, GIS

1 Introduction

Accurate and reliable measurement techniques are necessary for urban sustainability assessment as they provide a basis for characterisation of various urban neighbourhoods. Recent metrics such as Leadership in Energy and Environmental Design for Neighbourhood Development (LEED-ND) rating system credits can be used as sustainable urban form metrics. Urban Neighbourhood Pattern and Design (UNPD) is one of the major categories of LEED-ND rating system with a prerequisite of compact development. The main goal of this category is to build a community \cite{1}. A problem of this category is that the spatial dimension of compactness is often ignored. While social dimension is important in urban neighbourhood research, the spatial and structural assessment of urban neighbourhoods is also necessary. Apart from the shortcomings of the existing spatial assessment, height information of buildings has not been utilised appropriately even though building height plays a crucial role in both climatic design and human scale design. Climatic design aims to create convenient and comfort built-up areas by designing 3-dimensional (3D) building shapes and appropriate arrangement of the buildings in an urban environment that requires adaption to the local climate. Therefore, it is important to investigate 3D urban metrics that can quantify spatial dimensions of urban sustainability.

Urban metrics provide measurement methods of comparing the characteristics of different urban districts. There are several metrics for measuring characteristics of urban forms at metropolitan and building block scales. Spatial metrics have been known as effective measurement tools for characterizing urban forms in large study extents such as metropolises \cite{2}. For example, Huang et al. \cite{2} suggested seven metrics to reveal five elements of urban forms, namely, compactness, centrality, complexity, porosity and density to compare cities in developing and developed countries. The metrics characterising building pattern, complexity and compactness also exist; for example, Floor Area Ratio (FAR) and Building Coverage Ratio (BCR). However, metrics quantifying urban neighbourhood characteristics such as compactness remain scarce. Sustainable urban form measurement focuses on decision for either compact or sprawl development at the metropolitan scale. Compactness measurement at the neighbourhood scale is important for sustainable development assessment as the urban morphology and geometry affect wind speed, air quality at ground level \cite{3} and energy exchange \cite{4}. Indeed, traditional spatial metrics cannot characterise 3D urban forms because characterization of 3D urban forms requires both 3D data and adaptive spatial metrics. One of the fundamental barriers for 3D characterisation of urban forms is the large extent of the spatial analysis because 3D data (such as lidar point clouds) over an urban environment may not cover the entire urban area or could not provide a sufficiently high resolution. While remote sensing is reported as a reliable source of information over urban areas \cite{5}, the proven advantage of 3D urban remote sensing to analyse 3D
urban patterns and processes has not been fully explored. Urban form patterns have been traditionally derived using spatial urban metrics applied to remote sensing data. This bottom-up approach, deriving information from structure to process [6], can be improved using lidar. The advantage of lidar for urban pattern assessment studies is to provide accurate elevation information of both terrain and objects e.g. buildings and trees. The potential of lidar data for characterising urban forms can be used when we adapt the spatial and urban metrics to 3D space.

This research aims to develop a spatial 3D metric of compactness. We employ Moran’s I (MI) as a spatial metric of characterizing 3D compactness for the sustainable urban form assessment purpose. A novelty of this research lies in applying an autocorrelation statistic of MI to elevation information of lidar point clouds to investigate the level of 3D compactness of urban features, including man-made and natural objects.

In this paper, we review the current conceptual approaches in urban pattern modelling and the theoretical background of autocorrelation statistics of MI and Getis-Ord G (GOG). To fill the gap of 3D compactness measurements at the urban neighbourhood scale, we propose novel 3D metrics that will be tested in the simulation and implementation step. Before applying the statistics, the lidar data will be classified to ground and non-ground points. The statistics then will be applied to Digital Surface Model (DSM) and Normalised DSM (NDSM) of 6 urban districts grouped in two urban land uses including diverse object shapes, elevation and urban fabric.

2 Urban Pattern Conceptual Approaches

At the metropolitan scale of urban pattern recognition, two main conceptual approaches can be recognized; namely, traditional and modern perspectives. As Herold et al. [6] discussed, in the traditional top-down view (i.e. from process to structure), processes produce structures, while in the modern bottom-up view (i.e. from structure to process), structures can be a representative of processes. Although these conceptual approaches are reported for spatio-temporal analysis of urban forms, the main idea works for a snapshot of urban forms. That is, the modern view employs remote sensing data and spatial metrics for analysis of urban patterns whereas the traditional perspective believes that the processes are the major drivers of urban form and structure.

While the modern approach is acceptable in urban growth modelling at the metropolitan scale, it can be enhanced using remotely sensed 3D data at the neighbourhood scale. However, there are two main problems for such improvement. One problem that could be emerged when using 3D data is that the height information will be ignored spontaneously if the study extent is an entire metropolis or a big city. To address this problem, the scale of analysis has to be modified to an urban neighbourhood. Therefore, 3D data can be used for characterising the 3D pattern of urban neighbourhoods.

Another problem is lack of appropriate spatial metrics adapted to 3D space. The conventional spatial metrics are capable of characterizing planar urban forms in a large extent. However, we need metrics that are capable of characterizing urban forms at the neighborhood scale.

In general, urban fabric is categorised to ‘fine’, ‘coarse’ and ‘mixed’ where fine fabric includes mostly small size objects and coarse fabric mostly contains large objects. In many cases change of fabric from coarse to fine, or reverse, can be a sign of land use change. For example, industry sites are mostly recognized as coarse fabric because these areas include large size buildings. On the other hand, fine fabrics are mostly residential areas. In 3D space, these categories can be expanded based on the amount of high, medium or low buildings. Table 1 shows possible 3D urban fabrics. In 3D space, a central business district can be either mixed-high or coarse-high fabric. Downtowns can be examples of coarse high fabric.

Table 1. Urban fabric categories in 3D space

<table>
<thead>
<tr>
<th>Planar fabric</th>
<th>Height attribute of urban fabric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine</td>
<td>Fine-High</td>
</tr>
<tr>
<td>Coarse</td>
<td>Coarse-High</td>
</tr>
<tr>
<td>Mixed</td>
<td>Mixed-High</td>
</tr>
</tbody>
</table>

Figure 1 contains our proposed bottom-up approach for neighbourhood pattern analysis in 3D space. As can be seen, the overall goal is to start from Step 1: Structure, to achieve Step 4: Pattern. In this proposed bottom-up approach, remotely sensed 3D data would be employed using developed 3D urban metrics to derive, measure and compare urban neighbourhood patterns.

3 Moran’s I for Measuring 3D Compactness

The relationship between pair variables is known as correlation and the degree of correlation is measured through statistical coefficients [7]. Sign and the numerical value of correlation coefficient are important; for example a positive or direct relationship
is determined by a plus sign and a value close to 1 (|a| \rightarrow 1) shows higher strength of a relationship than a value close to 0 (|a| \rightarrow 0) [7]. Autocorrelation refers to the correlation between pairs of observations in a single variable [7]. This means that, in autocorrelation, the relationship between the values of a variable is investigated. Autocorrelation can be calculated for a variable changes over time, for linear spatial series and for two-dimensional spatial series. The phenomenon of spatial autocorrelation can be defined as the relationship among the values of an attribute in distributed areal units on a planar surface [7].

The null hypothesis for spatial autocorrelation is Complete Spatial Randomness (CSR) between observation values and _p_-value is calculated to test the hypothesis: if _p_ is small (less than 0.05), the null hypothesis is rejected and it indicates that clusters in observations exist.

MI and GOG are known as autocorrelation statistics measuring compactness and clusters of low or high values, respectively. In a defined study area and for an observation _i_ and neighbour observations of _j_ in a distance of _d_, MI for attribute of _x_ is defined as [8]:

\[
I(d) = \frac{n \sum w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W \sum (x_i - \bar{x})^2}
\]  

(1)

and GOG is defined as [8]:

\[
G(d) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}x_i x_j}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j} \quad j \neq i
\]  

(2)

where _w_{ij} is spatial weight and defined as the inverse distance between _i_ and _j_, and _W_ is the sum of all weights.

Getis and Ord [8] compared MI and GOG. They stated that "G statistic measures overall concentration or lack of concentration of all pairs of (_x_i_, _x_j_) such that _i_ and _j_ are within _d_ of each other. MI on the other hand, is often used to measure the correlation of each _x_i_ with all _x_j_’s within _d_ of _i_ and, therefore, is based on the degree of covariance within _d_ of all _x_i_” [8, p.196].

If we assume _K_1 and _K_2 as constants of the MI and GOG equations, as below, MI can be calculated based on GOG and then it helps for better comparison.

\[
K_1 = \frac{1}{\sum_{i=1}^{n} x_i^2}
\]  

(3)

\[
K_2 = \frac{n}{W \Sigma (x_i - \bar{x})^2}
\]  

(4)

\[
G(d) = K_1 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j
\]  

(5)

\[
I(d) = K_2 \sum w_{ij}(x_i - \bar{x})(x_j - \bar{x})
\]  

(6)

Therefore,

\[
I(d) = \left(\frac{K_2}{\bar{x}}\right) G(d) - K_2 \bar{x} \sum (w_{ij} - w_{ji}) x_i + K_2 \bar{x}^2 W
\]  

(7)

MI in Equation (1) and _G_ in Equation (2) are unequal when the weighted sums \(\sum w_{ij}\) and \(\sum w_{ji}\) are not equal to \(W\bar{x}\) and therefore, the weight patterns are different [8].

In characterizing the correlation among the same set of weighted observations, the difference between MI and GOG is that MI quantifies the spatial autocorrelation of a feature attribute considering both feature locations and feature values, concurrently but GOG measures the compression of high or low values of an attribute [9].

These two statistics can be compared using their corresponding z-scores. A z-score for MI statistic is defined as [9]:

\[
z_I = \frac{I(d) - E[I(d)]}{\sqrt{Var I(d)}}
\]  

(8)

where _E_ and _V_ represent the expected mean and variation, respectively.

In addition, a z-score for _G_ statistic is defined as [8]:

\[
z_G = \frac{G(d) - E[G(d)]}{\sqrt{Var G(d)}}
\]  

(9)

Measuring compactness through MI only considers the object’s location on a planar surface, hence, they are characterising only planar compactness but if we apply the autocorrelation statistics of MI and GOG on elevation attribute of urban objects we would achieve 3D compactness considering both layout and elevation of objects. _G_ statistic can be used as complementary statistic to derive the information for determining concentration of high or low values.

Two spatial statistics of MI and GOG have been used so far for measuring centrality/compactness [10-12] and concentration of high or low values [9]. MI has been recognized as an effective measurement tool of compactness of socio-economic data at the metropolitan scale [10]. However, in literature, its potential to characterize 3D compactness of urban fabric has been rarely explored.

4 Methodology

This research involves in two steps including simulation and application to a case study. The patterns in the simulation step are obtained from the well-known morphological urban patterns.

4.1 Simulation

We obtained 30 and more schematic buildings in a constant simulation extent since at least 30 square input features are required for calculation of MI and GOG [9]. Three different patterns are constructed for constant features by changing the elevation attribute of buildings which make three types of neighborhood pattern in 3D space. These are mono-centric, polycentric and decentralized (see Figures 2a to 2c).
The study extent is a coarse-high fabric surrounded by fine fabric. The aerial image over all the study extent is shown in Figure 3a. While the aerial photo can demonstrate whether a fabric is fine or coarse, 3D data is required for 3D fabric determination. We used lidar point clouds with 20 cm horizontal and 12 cm vertical accuracy. Figure 3b shows the triangulated lidar point clouds and the districts which are considered for applying autocorrelation statistics. Lidar data in Figure 3b shows the DSM. To exclude the objects’ height, NDSM is constructed from extracted non-ground points (see Figure 4a). A DSM contains elevation of both terrain and attached objects whereas a NDSM includes only absolute height information of objects such as buildings and trees. The difference between DSM and NDSM over same district can be distinguished through comparison between Figures 4b and 4c.

We explored the potential of both MI and GOG to test how their results are different for different 3D urban fabrics of fine-low (b), fine including some large tall buildings (a and c), coarse-low (f), coarse-medium (e) and coarse-high (d). The districts are selected based on their urban objects’ configuration. They can be categorized based on residential and educational land uses. These land uses have clearly different urban fabric configuration. Residential land use is a fine fabric and educational land use is a coarse fabric. To test the potential of MI for measuring 3D compactness we will apply MI and GOG statistics on the elevation attribute of lidar point clouds in both DSM and NDSM, derived from the point clouds. Indeed, we quantified the level of urban features compactness, including natural and man-made objects, in 3D space.

5 Results

To distinguish how various 3D patterns of urban neighborhoods can result in different numerical values of MI measurement, MI was applied on both 3D patterns in simulation study and 3D patterns of urban districts shown in Figures 2 and 3.

5.1. Simulation Results

Three different 3D patterns of a neighborhood are analysed where their difference comes from the change of elevation attribute of buildings and their layouts are exactly the same. Simulation results of MI over mono-centric (Figure 2a), polycentric (Figure 2b) and decentralized (Figure 2c) patterns are 0.13, 0.09 and -0.23, respectively. The results for more compact patterns of high rise buildings are positive whereas the result for decentralized pattern is negative.

These results confirm the results obtained by Tsai [10] in simulation study. However, in that study the data type, layout pattern, number of cells and the attribute value of each pixel were different.

5.2. Implementation Results

As described before, two products of lidar point clouds are DSM and NDSM. The DSM object’s height contains terrain and the object height. For an object in NDSM the objects height is absolute elevation. The autocorrelation statistics of MI and GOG are applied to both DSM and NDSM.

5.2.1. MI Statistic Results

Table 2 indicates the results of applying MI to DSM. As it shows, the results of residential land use fabric range between 0.84 and 0.92. The MI values for the districts with fine fabric including large and tall buildings are higher than the fine fabric. The MI values for the districts in educational land use with dominant coarse fabric are obtained between 0.80 and...
While the difference between the results of residential and educational land uses are not clear, the difference between the fine-low (b) and coarse-high (d) fabrics is considerable. MI for a fine-low fabric is achieved minimum among all the results which is 0.84 but the result for the coarse-high fabric is the maximum value among the results which is 0.97. The MI value for fine fabric of c is higher than the ones including high and large buildings (a and b). The results for a more compact form (d) in coarse fabric is higher than case (f) which is coarse-high but contains less buildings and the high buildings are dispersed.

### Table 2. Applied MI to DSM

<table>
<thead>
<tr>
<th>Land use</th>
<th>Urban fabric</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>a. Fine</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>b. Fine</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>c. Fine</td>
<td>0.92</td>
</tr>
<tr>
<td>Educational</td>
<td>d. Coarse</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>e. Coarse</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>f. Coarse</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Approximately similar pattern can be seen when we apply MI to NDSM. As can be seen in Table 3, fine-low fabric (b) has the minimum value of MI, (e) and (f) have lower levels of MI than case (d) which is more compact. The difference of the results of applying MI to DSM and NDSM is that in case of DSM, maximum value of MI was obtained for case (d) which is a fine-low fabric but in case of NDSM, maximum value is obtained for case (c) which is fine fabric including some large and tall buildings.

### Table 3. Applied MI on NDSM

<table>
<thead>
<tr>
<th>Land use</th>
<th>Urban fabric</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>a. Fine</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>b. Fine -low</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>c. Fine</td>
<td>1.08</td>
</tr>
<tr>
<td>Educational</td>
<td>d. Coarse</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>e. Coarse</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>f. Coarse</td>
<td>0.65</td>
</tr>
</tbody>
</table>

All the z-scores and p-values are in significant part of a normal distribution curve. The z-scores higher than 1.65 are in significant area with 90 percent confidence level. All our achieved z-scores are higher than 2.58, which means that the level of confidence achieved for all cases is 99 percent. As described before, the null hypothesis here is Complete Spatial Randomness (CSR) which is rejected by 99% confidence in all of our results.

### 5.2.2. G Statistic Results

As discussed in section 3, G statistic can be used in conjunction with MI for complementary information; MI measures compactness of both location and numerical value of an attribute distributed overall the data set but GOG is a concentration measurement tool for high or low values [9]. Table 4 demonstrates the results of applying GOG to DSMs of the urban districts. As it indicates, maximum value is obtained for district (b) and minimum value is obtained for districts (a) and (f) followed by (c) with a partial difference.

### Table 4. Applied G on DSM

<table>
<thead>
<tr>
<th>Land use</th>
<th>Urban fabric</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>a. Fine</td>
<td>43 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>b. Fine</td>
<td>98 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>c. Fine</td>
<td>47 × 10^{-6}</td>
</tr>
<tr>
<td>Educational</td>
<td>d. Coarse</td>
<td>52 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>e. Coarse</td>
<td>64 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>f. Coarse</td>
<td>43 × 10^{-6}</td>
</tr>
</tbody>
</table>

Table 5 contains the results of applying GOG to NDSM. The maximum G value is achieved for district (b) and minimum G value is obtained for district (d) followed by (a). Maximum G value is obtained for same district (b) and district (a) is in lowest level of G values in both Tables 4 and 5.

### Table 5. Applied G on NDSM

<table>
<thead>
<tr>
<th>Land use</th>
<th>Urban fabric</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>a. Fine</td>
<td>828 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>b. Fine</td>
<td>1328 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>c. Fine</td>
<td>1188 × 10^{-6}</td>
</tr>
<tr>
<td>Educational</td>
<td>d. Coarse</td>
<td>667 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>e. Coarse</td>
<td>1078 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>f. Coarse</td>
<td>1088 × 10^{-6}</td>
</tr>
</tbody>
</table>

In all results for G statistic, the z-scores remain in significant part of a normal distribution curve and the null hypothesis (CSR) can be rejected by 99% level of confidence.

The districts where their MI value is maximum have minimum G value and the districts with lowest MI value are obtained maximum G value. This happens because MI measures compactness of features and values location whereas G measures the compression of high or low values. Districts (a) and (c) where we obtained maximum value of MI and minimum value of G are fine fabric in residential land use and the urban objects are compact in layout so their MI are higher than other districts. In this residential fine fabric, low objects are concentrated compared to cases (d) and (e) where higher buildings are clustered.

### 6. Discussion

We promoted the bottom-up approach (from structure to process) of studying urban patterns using 3D remote sensing data and by proposing 3D metrics. The questions remain on how to improve the top-down approach to explain the 3D pattern of urban areas. It is expected that the drivers and factors influencing on 3D pattern and growth of urban areas be explored by urban planners and economists. In detail, these questions are: 1) How the top-down approach can characterise 3D urban growth? 2) Which factors, drivers and processes influence on making different
3D urban patterns? 3) How urban modelling and spatial urban theories can be developed to consider both urban layouts growth and various vertical patterns?

While the proposed 3D metric of MI is capable of distinguishing among different 3D patterns of urban neighbourhood caused by variation of only height attribute in simulation study, it needs to be integrated with the information derived from $G$ statistic in the implementation step. Indeed, using MI in conjunction with GOG in each district details our derived information about each district.

As the sample size and study areas are different in the implementation step, the MI results have to be compared within their context characteristics such as the number and size of urban objects, the area of study extent and the distribution of objects’ height pattern.

Among 3D urban fabric categories (Table 1), the categories of fine-low, fine including some large and tall buildings, coarse-medium, coarse-medium with decentralised objects and coarse-high are analysed in this paper and still the compactness values from MI for other patterns need to be explored.

The 3D compactness analysis in the implementation step could measure the compactness of all urban objects on a DSM including natural and man-made objects. For further assessment of compactness of man-made objects (buildings), we need to apply the statistics to the classified building points. For studying the compactness of natural objects such as trees and vegetation class we have to apply MI to only the classified vegetation points.

7. Concluding Remarks

In this paper, we aimed to improve 3D urban metrics for 3D pattern recognition of urban neighbourhoods. MI was firstly applied to simulation study over 3 urban neighbourhood patterns where their difference come from the difference in height attribute only. In the simulation study, a compact form achieved a positive value of MI whereas a decentralized pattern obtained a negative value of MI. The results from the implementation of MI over urban districts indicate that the general idea of clustering works and can differentiate the 3D compactness level among the districts. However, as the urban areas are more complicated than the simulation patterns, we need further information to enhance our understanding of different patterns. Therefore, in the implementation step we added $G$ statistic in our data processing and found that it could enhance our understanding through assigning a lower value to the patterns with shorter urban features than the patterns including taller objects.

It is recommended to apply the proposed 3D spatial metrics to other urban fabric types for a comprehensive study of 3D compactness of urban neighborhoods. For future work, applying MI to the classified buildings and vegetation is suggested to find the level of compactness of the built-up areas and vegetation compactness, respectively.

References

Abstract -

Human detection is an important topic that can be used for many applications, it is mainly found in areas that required surveillance such as airports, casinos, factories, construction and mining sites. In this paper, a novel human detection method is introduced to extract the human figure from an input image without prior information or training. This method firstly uses a head and shoulder detection scheme based on curve detection with scaled gradient magnitude and orientation maps. It is then followed by a human body estimation scheme based on gap detection and golden ratio. Finally, the human figure is extracted through thresholding local gradient magnitude regions and horizontal filling. Tests on various images have shown that this method is capable of detecting and extracting human body figures robustly from different images.

Keywords -

Human Detection; Head and Shoulder Detection; Human Body Estimation; Human Figure Extraction

1 Introduction

Human detection can be defined as a process of detecting the presence of human features from images or videos. Generally, in order to detect and track humans robustly, an algorithm needs to be able to extract common features among different people, so they can be separated from the background. As different people tend to have different features which are usually caused by their variable appearances and postures, this task can be quite challenging.

Human detection is generally used in areas that required surveillance [1, 2], these areas include airports, casinos, factories, construction and mining sites. The basic functions of human detection are used to detect and track the human, while the advanced functions which contain some form of post-processing are used to achieve additional goals. These goals may include people counting [3], face recognition [4] and behaviour recognition [5]. The advanced functions are generally used for high security environments such as airports and casinos, these systems tend to be expensive which equipped with high speed communication lines and powerful processors [6].

Human detection methods can be divided into video based [7] or static image based [8]. The major difference between these two types of methods is that video based approaches could also utilise the motion features such as background subtraction [9] and optical flow [10], this is impracticable with a single image.

The main contribution of this paper is the introduction of a novel human detection method which includes a new method of detecting the head and shoulder, a new method of estimating the size and position of the human body and a new method of extracting an accurate figure of the human body. Also, this human detection method does not require any prior training.

This paper is organised into 5 sections. Introduction and related work are given in Sections 1 and 2, the proposed human detection method is introduced in Section 3, experimental results are provided in Section 4 and conclusions are given in Section 5.

2 Related Work

The most influential human detection method is the Histogram of Oriented Gradients method (HOG) [11]. HOG is a feature descriptor that heavily relies on the extraction of gradient orientations and the use of linear Support Vector Machine classifier [12]. HOG is renowned for its human detection ability by utilising large amount of training images. However, this means the algorithm itself is also heavily limited by the quality and quantity of these training images. This is because, in order to learn a classifier successfully, HOG requires continuously recurring shape events in the given blocks of the training images [13]. HOG generally suffers from speed issues and the training itself is a time consuming task. Many existing methods tend to build their algorithm based on HOG or utilise it as an additional feature for their algorithms [14–20].

One method argues that humans in standing positions tend to have distinguishing colour characteristics [14],
therefore addition colour information is employed with HOG descriptors. This includes colour frequency and co-occurrence features in each channel of Hue-Saturation-Value (HSV) images. It also employed the partial least square regression analysis [15] onto the descriptors. These additional elements improved the accuracy of the HOG, however at the same time the computational cost is greatly increased.

Another method utilises the omega-shape features of humans [16]. This method argues that HOG feature based classifier are generally accurate but very slow to work with, while Haar feature based classifiers [17] are generally fast but suffer from poor accuracies. Thus, by combining these two classifiers, robustness could be achieved. This is done by first employ the Haar feature classifier to exclude obvious negative image patches and then employ HOG for the remaining image patches. This method greatly improved the computation speed with a small drop in accuracy. However, the algorithm only performs detection and tracking of the head, rather than the whole body.

Scale-Invariant Feature Transform (SIFT) can also be used to detect human [18–20]. SIFT [21] is an algorithm to detect and describe local features by extracting distinctive invariant features. SIFT descriptors are similar to HOG in the sense that both descriptors utilize gradient features. The difference is that HOG is computed in dense grids while SIFT is computed in sparse grids with orientation alignment. By combining HOG and SIFT descriptors, higher detection rate can be achieved. However, the efficiency of these methods tends to suffer with online processing, since both descriptors have reasonably high computational costs.

Therefore, this research is aimed to develop a human detection method that capable of extracting the human figures without the need of pre-training.

3 Proposed Method

The proposed human detection method consists of four stages: scaled gradient mapping, head and shoulder detection, human body estimation and human figure extraction. A flowchart of the method is shown in Figure 1.

3.1 Scaled Gradient Mapping

The first step in the scaled gradient mapping stage is to convert the input image to greyscale, then the gradient is computed using two simple kernels, as given in Equations (1) and (2).

\[
G_x = [-1 \quad 0 \quad 1] \quad (1)
\]

\[
G_y = [\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}] \quad (2)
\]

where \(G_x\) and \(G_y\) are the gradient components along \(x\) (horizontal direction) and \(y\) (vertical direction) axes of the image respectively.

The exact gradient magnitude and orientation values are then computed for every pixel, as given in Equations (3) and (4), then saved into two gradient maps, known as the magnitude map and the orientation map. The origin of the coordinate system is at the upper left corner.

\[
G = \sqrt{G_x^2 + G_y^2} \quad (3)
\]

\[
\theta = \tan^{-1}(G_y, G_x) \quad (4)
\]

Both gradient maps are scaled into 8 smaller sizes in which the height and the width of the maps are divided by a scale factor \(S\) (where \(S = \{2, 3, 4, 6, 8, 12, 16, 24\}\)), this is done to accommodate various sizes of human in the image. During the scaling procedure, both maps are divided evenly into square blocks with length equal to \(S\). For the magnitude maps, the local maximum magnitude values inside each block are saved. While for the orientation maps, both positive and negative directions of the gradient along \(x\) and \(y\) axes are used to obtain the orientation for the scaled map, as given in Equation (5).
\[ \theta_s = \arctan(2(G_{y,\text{max}}^{-} - G_{y,\text{max}}^{+}, G_{x,\text{max}}^{+} - G_{x,\text{max}}^{-})) \]  

where \( \theta_s \) is the new orientation value for the scaled orientation map with the scaling factor \( S \), \( G_{x,\text{max}}^{+} \) and \( G_{x,\text{max}}^{-} \) are the maximum positive and negative \( x \) components of the gradient respectively, \( G_{y,\text{max}}^{+} \) and \( G_{y,\text{max}}^{-} \) are the maximum positive and negative \( y \) components of gradient respectively.

### 3.2 Head and Shoulder Detection

As both shoulders and the top half of the human head have distinctive curvatures, curve detection is used to detect possible head and shoulder in an image using curve templates on the scaled gradient maps. The templates consist of a left and a right curve blocks which are oriented outward with origins lie in the bottom right and bottom left corner respectively. The two curve templates with their orientation values are given in Figure 2.

All orientation maps are then searched for patches that contain similar orientation values using a constant 3×3 search window. The average difference between the patch and the template relate to its magnitude are calculated and compared to a degree of difference parameter \( p \), as given in Equation (6).

\[ G_{\text{max}} \frac{\sum_{i=1}^{4} (\theta_i - \mu_i)}{\sum_{i=1}^{4} G_i} \leq p \]  

where \( \theta_i \) and \( \mu_i \) are the \( i \)th orientations in the block of the patch and template respectively, \( G_i \) is the \( i \)th magnitude in the block of the patch, \( G_{\text{max}} \) is the maximum magnitude of the image, \( p \) is the maximum average difference parameter, default value used is 30° (\( \pi/6 \)).

When a region satisfies the maximum average differences allowed, it can be said that there is a high chance of seeing parts of the head or shoulder. The curve detection results of these two templates are then joined together by horizontal filling, that is filling the space between them. This is done only if the pixel distance between the left and right curve is smaller than a width threshold \( d \), this width threshold is determined based on the maximum detectable size of the human figure in the current image, as given in Equation (7).

\[ h \frac{\varphi}{2\varphi + 1} \leq d \]  

where \( \varphi \) is the golden ratio \( (\approx 1.618) \), \( h \) is the maximum height (in pixels) of the image, \( d \) is the maximum allowed width (pixel distance) between the left and right curve for horizontal filling.

Golden ratio can be an excellent tool in estimating the proportions of the human body \([22, 23]\), the model of the human body is created based on this to estimate the size and proportion of the human figure, as given in Figure 3. The size of the body is estimated between the top of the head to the knee. This model is also used to determine the width ratio \( d \) in Equation (7).

The Gaussian probability distribution function used for the model is given in Equation (8).

\[ f(j) = e^{-(j-r)^2/\alpha^2} \]
where $j$ and $r$ are the current and ideal results of a comparison between two blobs, $\sigma$ is the standard deviation.

The comparisons and their related parameters are listed in Table 1, where $w$ is the width of the head. The overall probability is then calculated, as given in Equation (9).

Table 1. Parameters used for the comparisons of two blobs

<table>
<thead>
<tr>
<th>Type</th>
<th>Comparisons Between Two Blobs</th>
<th>$r$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>centroid difference in $x$</td>
<td>0</td>
<td>$w$</td>
</tr>
<tr>
<td>B</td>
<td>width ratio</td>
<td>$1/\phi^2$</td>
<td>$1/(2\phi^2)$</td>
</tr>
<tr>
<td>C</td>
<td>centroid difference in $y$ to width ratio</td>
<td>$1/\phi$</td>
<td>$1/(2\phi)$</td>
</tr>
<tr>
<td>D</td>
<td>area ratio</td>
<td>$1/\phi$</td>
<td>$1/(2\phi)$</td>
</tr>
</tbody>
</table>

$$f(j) = f_A(j)\frac{f_A(j) + f_C(j) + f_D(j)}{3} \quad (9)$$

where $A$, $B$, $C$ and $D$ are the four feature comparisons between two blobs, $A$ is given more weight since a small centroid difference in $x$ direction tends to indicate a much higher chance of seeing the head and shoulder blobs.

3.3 Human Body Estimation

Once the head and shoulder blobs are identified, gap detection is then performed to locate possible gaps below the armpit and between the legs. These gaps can be assumed as regions with multiple edges and significant orientation changes. The detection is done by searching all orientation maps for patches that contain similar orientation values to the gap template, as given in Figure 4. The gap template is oriented inward with interception point lies on the centre of the bottom edge.

In the figure, the filled areas are the head and shoulder blobs, the 3 red blocks are the estimated gap regions, while the 8 blue blocks are the estimated body regions divided from human body size estimation (Figure 3). All the blocks overlaps at least one other regions (blocks 4 and 5 overlaps the head and shoulder blobs), this is done to ensure the maximum coverage and the accuracy in detecting the human body.

The size and position of the blocks are determined based on golden ratio $\phi$ and are listed in Table 2, where $B_i$ represents the block number, $x_0$ and $y_0$ indicate the coordinates of the top left corner of the blocks, $S_x$ and $S_y$ are the centroid coordinates of the shoulder blob while $H_x$ and $H_y$ are the centroid coordinates of the head blob, $w$ is the width of the head, positive directions for $x$ and $y$ are right and bottom respectively, $H_{min}$ and $S_{min}$ are the minimum $x$ position for the head and minimum $y$ position for the shoulder respectively, $G_9$ and $G_{10}$ are the outer most gap location in $x$ for region 9 and 10 while $G_{11}$ is the centroid coordinate of gap in $x$ for region 11.
The head and gap regions (blocks 1–5) are determined first and are then used to locate the body regions (blocks 6–11). As regions below the gaps tend to be also gaps, a vertical filling process is further performed for gaps detected in the estimated gap regions (blocks 1–3), in which the lowest positioned gap pixels are filled vertically downwards to their corresponding region boundaries. If no gaps are detected, all body blocks will be aligned to \( H_i \) with equal width of \( \phi w/2 \).

<table>
<thead>
<tr>
<th>( B_i )</th>
<th>Width</th>
<th>Height</th>
<th>( x_0 )</th>
<th>( y_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( wp(\phi - 1) )</td>
<td>( 3\phi^2 w/2 )</td>
<td>( H_i - wp(\phi - \phi) )</td>
<td>( S_y )</td>
</tr>
<tr>
<td>2</td>
<td>( wp(\phi - 1) )</td>
<td>( 3\phi^2 w/2 )</td>
<td>( wp(\phi - \phi) )</td>
<td>( S_y )</td>
</tr>
<tr>
<td>3</td>
<td>( \phi w )</td>
<td>( \phi^2 w/2 )</td>
<td>( H_i - \phi w/2 )</td>
<td>( S_{min} + \phi w^2/2 )</td>
</tr>
<tr>
<td>4</td>
<td>( w/2 )</td>
<td>( S_y - H_y )</td>
<td>( H_{min} )</td>
<td>( H_y )</td>
</tr>
<tr>
<td>5</td>
<td>( w/2 )</td>
<td>( S_y - H_y )</td>
<td>( H_y )</td>
<td>( H_y )</td>
</tr>
<tr>
<td>6</td>
<td>( (G_{10} - S_{min} - S_y) ) ( G_o/2 )</td>
<td>( + \phi^2 w )</td>
<td>( G_y )</td>
<td>( S_y )</td>
</tr>
<tr>
<td>7</td>
<td>( (G_{10} - S_{min} - S_y) ) ( G_o/2 )</td>
<td>( + \phi^2 w )</td>
<td>( G_o/2 )</td>
<td>( G_y )</td>
</tr>
<tr>
<td>8</td>
<td>( (G_{10} - G_y) ) ( + \phi^2 w )</td>
<td>( + \phi^2 w )</td>
<td>( + \phi^2 w^2/2 )</td>
<td>( G_{11}/2 )</td>
</tr>
<tr>
<td>9</td>
<td>( (G_{10} - G_y) ) ( + \phi^2 w )</td>
<td>( + \phi^2 w )</td>
<td>( + G_{11}/4 )</td>
<td>( S_{min} + \phi^2 w/2 )</td>
</tr>
<tr>
<td>10</td>
<td>( \phi^2 w/2 )</td>
<td>( \phi^2 w )</td>
<td>( G_{11} )</td>
<td>( S_{min} + \phi^2 w^2/2 )</td>
</tr>
<tr>
<td>11</td>
<td>( \phi^2 w/2 )</td>
<td>( \phi^2 w )</td>
<td>( G_{11} )</td>
<td>( S_{min} + \phi^2 w^2/2 )</td>
</tr>
</tbody>
</table>

The final step is the horizontal filling in which the inter-space between the outlines are filled. The blue blocks are filled per block while the red blocks are filled per gap region. The final human figure is generated by adding the head and shoulder blobs detected earlier with the filled body regions (blue) and removing the filled gap regions (red).

### 4 Experimental Results

Numerous images have been tested to examine the effectiveness of the proposed method, due to space limitations, only some results are shown in Figure 6. The images used are collected from various sources and have different sizes and quality. Six of these images have been presented in the results, the first two images (from the top) are produced by us using a camera of a mobile robot. The next two images (in the middle) are downloaded from INRIA Person Dataset (http://pascal.inrialpes.fr/data/human/) and the last two images are picked from Google images. These results indicate the high effectiveness of our method as the figures of the human body are clearly extracted.

In the figure, the original input images with body estimation regions indicated by the coloured rectangles are depicted in Figure 6(a). The unfilled body outlines generated from the human body regions (blue blocks) are shown in Figure 6(b), while the filled gap outlines generated from the gap regions (red blocks) are shown in Figure 6(c). The final results indicated by the blue coloured regions are presented in Figure 6(d).
5 Conclusions

Current research in human detection is lacking of training-less approaches, therefore in this paper, a novel human detection method is introduced using gradient maps and the golden ratio. Compared to existing methods, the advantages of our methods are that firstly rather than an approximated region based human detection, the figure of the human body is completely extracted with high accuracy, secondly the algorithm only need one input image and no training images are required. Currently, the algorithm can only detect a single person with front or back view by selecting the pair of blobs with the highest probability ratio, future works are underway to address this issue.

References


Autocorrelation Statistics-Based Algorithms for Automatic Ground and Non-ground Classification of Lidar Data

Sara Shirowzhan\textsuperscript{a} and Samsung Lim\textsuperscript{b}

\textsuperscript{a,b} School of Civil and Environmental Engineering, University of New South Wales, Australia
E-mail: S.Shirowzhan@student.unsw.edu.au, S.lim@unsw.edu.au

Abstract –

Classification of lidar data to ground and non-ground points is important for accurate topography mapping and reliable estimation of slope, volume and buildings’ geometry over urban areas. Manual or semi-automatic classification provides relatively good results, however, automatic classification in complex areas with diverse object sizes is still challenging.

This research aims to propose two novel algorithms based on Getis-Ord $G^i*$ (or $G^i*$ for short) and Local Moran’s $I$ (LMI) statistics to classify a lidar point cloud into a set of points representing ground and another set of points reflected from non-ground e.g. buildings and vegetation. The two statistics, $G^i*$ and LMI, have been widely used in cluster analysis to identify clustered features of high $z$-scores and low $z$-scores. Based on the two statistics, we proposed two classification algorithms that allow varying window sizes e.g. 100 m, 150 m and 200 m, and applied the algorithms to the lidar data in order to obtain optimal classification results. The results show that the $G^i*$-based algorithm decreases omission errors but increases commission errors when compared to the LMI-based approach. Overall the 100-m window size outperforms than the other window sizes in terms of feature extraction in slant areas, whereas the 150-m window size provides slightly better results in a complex scene of high-rise buildings and dense vegetation, and the 200-m window size is more efficient if large buildings are present in the study area. This feasibility study indicates that autocorrelation statistics such as $G^i*$ and LMI can be effectively used to classify a lidar point cloud.

Keywords -

Automatic classification, Bare earth extraction, Spatial association, DSM, DEM, NDSM

1 Introduction

A Digital Elevation Model (DEM) is a grid containing elevation information of ground which can be derived from lidar point clouds. The DEM can be used for topography mapping and extraction of non-ground points including man-made and natural objects such as buildings and trees. Light detection and ranging (lidar) is a proven technology for generating highly accurate Digital Elevation Models (DEMs). However, Pingel et al. [1] reported that terrain classification is a challenging problem for DEM production since classification errors, commission errors in particular, are always inevitable.

Although conventional filtering algorithms e.g. morphological filters for classification in general areas are performing well, accurate classification of lidar point clouds in complex urban scenes including large and small objects with diverse elevation [2-4] is challenging. Classification results often depend on the assumptions made to the neighbours of each lidar point [5]. One of the assumptions is that the points representing an artificial object are clustered with elevations above the mean value of the neighbours, and the points representing ground are also clustered but have elevations below the mean value [6]. This assumption can be formulated as a criterion for lidar data classification by utilising some spatial autocorrelation statistics. For example, Getis and Ord [7] claimed that using Moran’s $I$ (MI) in conjunction with the $G$ statistic would improve the understanding of a given spatial data set. Hence this study aims to apply two local statistics, namely, Local Moran’s $I$ (LMI) and Getis-Ord $G^i*$ (or $G^i*$ for short), for effective classifications of a lidar point cloud into points representing ground and points reflected from non-ground such as buildings and vegetation.

Previously, MI has been applied to images as a texture measurement tool for object classification [8-11]. Su et al. [10] claimed that a higher accuracy for object classification is achieved by applying MI, compared to other spectral and textural measurement methods. They achieved 87% of Cohen’s kappa index as an accuracy indicator of the MI-based classification. However, the calculation of MI for lidar data is challenging because of
the high point density. On the other hand, LMI is one of the spatial autocorrelation statistics that Roggero [12] utilised for deriving information from spectral images. The potential of LMI comes from the nature of its spatial autocorrelation functionality that compares a local variation with the global variation [12, p. 58]. In this paper, a local variation of elevation is compared with the overall variance of elevation in a pre-defined window.

The Global Moran’s I (GMI) is defined by [13] as follows:

\[ I = \frac{N \sum_i \sum_j W_{ij} (Z_i - \mu)(Z_j - \mu)}{\sum_i \sum_j W_{ij}} = 1 \]  

where \( Z_i \) and \( Z_j \) refer to the deviation of attribute values (i.e. elevation information of the lidar points in this study) of features \( i \) and \( j \) from the mean value within the distance threshold of a neighborhood and \( N \) is the total number of features within the neighborhood. In addition, \( W_{ij} \)'s are spatial weights between the pairs \( i \) and \( j \) e.g. inverse distance weights between the pairs. While GMI as a spatial autocorrelation shows the overall similarities or dissimilarities, LMI is known as a more effective spatial autocorrelation indicator which can be used for a microscopic analysis of GMI. Indeed, LMI exhibits the contribution from each observation to the global indicator microscopically. Indeed, LMI exhibits the contribution from each observation to the global indicator microscopically. Anselin [13] defined LMI as:

\[ I_i = Z_i \sum_j W_{ij} Z_j \]  

where the weight was conveniently allocated as the inverse distance between the features \( i \) and \( j \).

A positive LMI indicates that the lidar point belongs to a cluster of similar values, and a negative LMI indicates that the lidar point is dissimilar to the neighbours. As LMI depends strongly on the neighbours, it is a relative measure [14] and therefore cannot be regarded as a z-score for lidar point classification. According to [15], the z-score of LMI is defined as follows.

\[ z_i = \frac{I_i - E[I]}{\sqrt{V[I]}} \]  

where \( E \) and \( V \) represent the expected mean and variance, respectively.

All classification algorithms suffer from two types of errors: Type 1 errors (i.e. omission errors), and Type 2 errors (i.e. commission errors) [5]. In this paper we argue that, when it comes to ground-point classification, omission errors are not so critical for deriving a DEM as long as commission errors remain small because the DEM is interpolated from the correctly classified ground points. Consequently, omission errors for non-ground points can be less important in an initial classification step because the non-ground points can be reclassified by applying the DEM. For this reason, the main aim of this study is to minimize commission errors, even though omission errors will be assessed as well. \( G_* \) can be defined as [16]:

\[ G_* = \frac{\sum_i W_{ij} (Z_i - \mu)(Z_j - \mu)}{\sqrt{V[Z] N \sum_i W_{ij}^2}} \]  

The meaning of a significance level of \( G_* \) and LMI was explained by Getis and Ord [7]. That is, clusters of high or low values are obtained by choosing a significant LMI z-score. In general, a significant z-score is chosen to be higher than 2.58 or lower than -2.58 where the corresponding \( p \)-value is smaller than 0.05. As for the significance level of z-scores, Ebdon stated that “The probability that the null hypothesis is correct is referred as the significance level. The null hypothesis can be rejected if this probability is acceptably low. Significance levels of 0.05 or even 0.1 are possibly adequate in many geographical applications” [17, p.16].

In this paper, two significant levels of the LMI z-scores and \( G_* \) values are used for lidar point classification because it is observed that significant positive values are obtained in the clusters of high-elevation lidar points and significant negative values are found in the clusters of low-elevation lidar points. The significance level of the LMI z-scores is ±2.58, that is, ground points are supposed to have z-scores less than -2.58, and non-ground points have z-scores greater than 2.58. However, the significance level of \( G_* \) is chosen differently in order to improve the classification results, that is, less than -1.65 for ground, or greater than 1.65 for non-ground. This paper presents and tests the two algorithms based on the statistics of \( G_* \) and LMI, then discusses advantages and disadvantages of the two algorithms by investigating omission errors and commission errors.

2 Methodology

In this study we use the lidar data set over the University of New South Wales (UNSW) with 20-cm horizontal accuracy and 12-cm vertical accuracy. The study extent is illustrated in Figure 1.
elevation values (see the vertical profile along the line segment between Point A and Point B in Figure 1). The study area contains 2 challenging areas for object classification: a steep slope area and a complex area. The complex area includes high-rise buildings, large-size buildings, and tall trees. In this study two algorithms are proposed: one based on LMI and the other based on $G_i^*$. With the two algorithms, three different window sizes of 100 m, 150 m and 200 m are tested. For each window size, ground and non-ground points are classified by applying the two algorithms. The $G_i^*$-based algorithm is defined as:

1. Set a window size. For each window,
2. Calculate $G_i^*$ with a pre-defined neighbourhood and weights
3. Assign the clusters of low values ($G_i^* < -1.65$ and $p$-value < 0.05) to ground
4. Assign the clusters of high values ($G_i^* > 1.65$ and $p$-value < 0.05) to non-ground
5. Verify the results against the reference.

The LMI-based algorithm is similar to the $G_i^*$-based except that the LMI $z$-score is used instead of $G_i^*$, and the significance level is ±2.58 instead of ±1.65. Test results of ground and non-ground classification using the proposed algorithms will be compared to the reference ground and the reference non-ground, respectively. Two datasets of ground and non-ground shown in Figures 2a and 2b, respectively, are the classification results from a commercial software package known as Terrasolid, which can be used as a reference for comparison purposes as they are independently obtained. Low vegetation is outside our study scope, hence points of less than 1-m height are removed from non-ground data. Low vegetation thresholds of 0.5, 1, 1.5, 2 and 2.5 m were tested by Estornell et al. [18] who concluded that the larger the thresholds they chose, the smaller error levels they obtained. In our case, any thresholds greater than 1 m are not appropriate because substantial vegetation of 1-m height or higher were observed in the study area.

### 3 Validation

For validation of the results, two types of errors are calculated: Type 1 errors and Type 2 errors. Type 1 occurs when the algorithm “rejects the null hypothesis even though it is true” and Type 2 occurs when the algorithm “fails to reject the null hypothesis even though it is false” [19, p. 147]. Our null hypothesis for lidar data classification is that ground points have low $z$-scores and non-ground points have high $z$-scores. Therefore, Type 1 in ground classification occurs when the lidar point is actually reflected from ground but is not classified as a ground point, hence it belongs to an omission error. Type 2 in ground classification occurs when the lidar point is reflected from non-ground e.g. a building or a tree but is classified as a ground point i.e. it can be categorized as a commission error. Omission errors and commission errors in non-ground classification can be defined similarly.

### 4 Results

One of the objectives of this paper is to show that, by applying $G_i^*$ and MI, clusters of low $z$-scores can be identified from ground points and clusters of high $z$-scores can be identified from non-ground points. Figures 3a and 3b show the extracted ground and non-ground points from the LMI-based algorithm using window sizes of 100 m, 150 m and 200 m. Comparing reference ground points (Figure 2a) and ground points extracted by the LMI-based algorithm (Figure 3a), it shows that a slant area is the subject of commission error (i.e. non-ground points are classified as ground points). On the other hand, low vegetation and man-made objects are the main areas of omission error in non-ground points.
non-ground points with 100-m, 150-m and 200-m window sizes using the $G_i*$-based algorithm. Comparison of the overall results from the LMI-based algorithm and the $G_i*$-based algorithm with reference ground and non-ground points indicates that the level of commission error increases while the level of omission error decreases. This is confirmed if we compare Tables 1 and 2.

The results from the LMI-based algorithm and the $G_i*$-based algorithm are demonstrated in Tables 1 and 2. Ground classification is improved by $G_i*$ regardless of the window sizes. However, improvement in non-ground classification is not so obvious. Omission errors in ground classification decrease from the LMI-based 25-29% to the $G_i*$-based 16-20% (see Tables 1 and 2). On the other hand omission errors in non-ground classification decrease slightly from the LMI-based 28-35% to the $G_i*$-based 23-31%. Therefore it can be concluded that overall omission errors are higher than expected. This problem will be discussed in Section 5.

Commission errors in ground classification increase from the LMI-based 1-7% to the $G_i*$-based 10-12%, and commission errors in non-ground classification also increase from the LMI-based 2-7% to the $G_i*$-based 7-10%. Hence it is obvious that the LMI-based algorithm provides a more reliable result in terms of commission errors. From Tables 1 and 2, it can be seen that the least commission error in ground classification is obtained by using the window size of 100 m. In addition, the least commission error in non-ground classification is obtained by the window size of 150 m if the LMI-based is used (Table 1), and by 100 m in case of the $G_i*$-based (Table 2).

### Table 1. Level of errors in rasters of ground and non-ground points.

<table>
<thead>
<tr>
<th>Window Sizes (m)</th>
<th>LMI Omission Errors (%)</th>
<th>LMI Commission Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRD*</td>
<td>NGRD**</td>
<td>GRD</td>
</tr>
<tr>
<td>100</td>
<td>29</td>
<td>28</td>
</tr>
<tr>
<td>150</td>
<td>25</td>
<td>31</td>
</tr>
<tr>
<td>200</td>
<td>28</td>
<td>35</td>
</tr>
</tbody>
</table>

*GRD: Ground Classification, **NGRD: Non-ground Classification

### Table 2. Level of errors in rasters of ground and non-ground points

<table>
<thead>
<tr>
<th>Window Sizes (m)</th>
<th>$G_i*$ Omission Errors (%)</th>
<th>$G_i*$ Commission Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRD*</td>
<td>NGRD**</td>
<td>GRD</td>
</tr>
<tr>
<td>100</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>150</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>200</td>
<td>20</td>
<td>31</td>
</tr>
</tbody>
</table>

A significant reduction of omission errors in ground classification by the $G_i*$ based algorithm can be realized by comparing the ground points extracted from the two proposed methods. The results show that $G_i*$ with the 100-m window size provides the least omission error and this is also confirmed by the error calculation given in Table 2. One of the areas subject to a high level of omission errors in non-ground classification is a slant area. From a visual inspection, it is likely that the least omission error by LMI seems to be obtained by the 150-m window size, however, Table 1 indicates otherwise i.e. the least error comes actually from the 100-m window size. From another visual inspection, it is seen that the omitted points in the 200-m window size are clustered in the slant and complex areas but the omitted points in the 150-m window size are dispersed.

Tables 1 and 2 represent that the commission error in non-ground classification is overall lower than that in ground classification, except the commission error in ground classification by LMI with the 100-m window size is the lowest. From the results, it can be concluded that the problematic areas of high commission errors are slant areas and object boundaries. Visual inspections of the errors demonstrate that the commission errors by both methods are lower if the 100-m window size is used. An interesting result is that the commission errors in non-ground classification are not affected much by the slope. In addition, the omission error for ground classification decreases to 23% for the LMI-based algorithm and to 15% for the $G_i*$-based algorithm when we combine (see Figure 3) the ground points extracted by 100-m, 150-m and 200-m window sizes. While the omission error reduces by combining the results from different window sizes, the commission error increases.
All in all, it is observed that, in the cluster of low values ($G_i^* < -1.65$ or $|z_i| < -2.58$ and $p$-value < 0.05), 93% of the points are ground points when the window size of 100 m is used. Larger windows tend to decrease this rate to 82.5% (for the 150-m window size) and 75.5% (for the 200-m window size). On the other hand, in the cluster of high values ($G_i^* > +1.65$ or $|z_i| < -2.58$ and $p$-value < 0.05), a high portion of the points are non-ground points: 99.9% for the window sizes of 200 m and 150 m, and 99.8% for the 100-m window size. Therefore it is reasonable to conclude that the window size of 100 m outperforms the other choices and the corresponding commission error is very small. However, it should be noted that a substantial amount of points still do not belong to either the low LMI clusters or the high LMI clusters, hence the omission error can increase.

5 Discussion

The main aim of this research was to apply LMI and $G_i^*$ to the lidar data classification and find one that outperforms the other. The results, however, indicate that the two algorithms are comparable to each other i.e. the $G_i^*$-based algorithm produces lower omission errors but higher commission errors than the LMI-based does. In addition, it is observed that most of the commission errors in ground classification occurred in slope areas. This problem is also reported by Meng et al. [20] that slant areas are challenging for cluster-based classifiers. Therefore, a slope-based algorithm is suggested to reduce the commission errors.

It should be noted that the results of the proposed algorithms depend on window sizes and neighbours per lidar point (i.e. the lidar points within the given window size). In that sense, the study area for this research is very challenging because it contains high-rise buildings, large-area buildings, complex scenes and slant areas. Therefore the omission errors in ground classification were higher than expected. This problem can be minor if an accurate DEM can be generated from the classified ground points, provided small commission errors. That is, the LMI-based algorithm can be applied to the DEM generation, and then the $G_i^*$-based algorithm can be applied to the non-ground classification.

It is noticed that the level of error is significantly influenced by the method of calculation. For example, point-wise error calculation decreases considerably when compared to rasterisation-based calculation. Moreover, it is shown that the low vegetation threshold for DEM generation affects both Type I and II errors [18]. Future work is to determine an optimal threshold to lower the level of omission and commission errors.

6 Concluding Remarks

It was shown in this paper that LMI and $G_i^*$ can be used effectively as the automatic classifiers of lidar points to ground and non-ground points. The test results from the two algorithms exhibit that each has advantages and disadvantages. The LMI-based algorithm provides lower-level commission errors, and the $G_i^*$ based algorithm produces lower-level omission errors. Therefore, the LMI-based algorithm is suitable for the DEM generation, and the $G_i^*$-based algorithm is preferable for the feature extraction.

As for ground classification, slant areas are identified as the main source of the commission errors from both classifiers, and of the omission errors from the LMI-based classifier, regardless of the window sizes. However, the slant areas do not cause significant commission errors in non-ground classification. It is also observed that the commission errors in non-ground classification are dispersed, whereas the commission errors in ground classification are clustered. In plain areas, the commission errors occurred mainly on the boundaries of buildings and vegetation. It is suggested that the proposed algorithms with a moving window or an ad-hoc window size can reduce such commission errors.


AUTOMATION, CONSTRUCTION AND ENVIRONMENT

POSTER PAPERS
A Framework for Supporting Planning and Development of Infrastructure Projects from a Societal Perspective

H. Doloi a

aFaculty of Architecture, Building and Planning, The University of Melbourne, Australia
E-mail: hdoloi@unimelb.edu.au

Abstract

Capital projects are the backbones of our society and they must go through an evolutionary process of repeated planning and analysis cycles. Ensuring a building is structurally sound, or that a tallest building will add a new image to a city is not enough: new infrastructure and urban development must be rationalised in their planning, design and operations ensuring fitness-for-purpose and adequate societal value contributions (e.g. local resource utilisation, community wealth creation etc.) within the community. In order to ensure value creation and complete social acceptance, projects must be planned and developed by aligning the needs and requirements of the wider community proactively. This research develops an experimental framework aiming to facilitate holistic decision making on capital projects. Adopting the Social Network Analysis (SNA) based innovative methodology, a framework for quantifying social value in infrastructure projects has been demonstrated. Based on an Australia case study, the process of integration of representative community views of extended stakeholders over planning, development and operation phases in capital projects has been highlighted. Social Network Analysis (SNA) is considered to be one of the highly appropriate research methods for enquiring the complex patterns of interactions between stakeholders associated with the capital projects. The new framework will potentially assist in replacing the traditional top-down planning processes and promoting sophistication in developing socially responsive infrastructure within built environment.

Keywords: Built Infrastructure and Human Factors, project planning, social value, infrastructure

1 Introduction

In the next one hundred years, cities will rapidly grow in size and complexity with soaring population growth. Melbourne population is predicted to reach 7 million in 2050 contributing to a total of 35.9 million nationally. Now, more than ever, there is a need to assess the long term needs, performance and social acceptance of infrastructures before they are built. One major weakness in the traditional approach is the poor inclusion of the stakeholders’ concerns at the early stage of project planning process (Clarkson 1995).

In order to ensure value creation and complete social acceptance, projects must be planned and developed by aligning the needs and requirements of the wider community proactively (Carrington 2005; Galaskiewcz 1979). One major weakness in the traditional approach is the poor inclusion of the stakeholders’ concerns at the early stage of project planning process. Stakeholders are the individual or groups of people who have an interest in the project or are being affected directly or indirectly by the project outcomes, irrespective of their direct contractual relationships with the project sponsors (Doloi et al. 2011, Doloi 2012). While the community consultation processes are used for development of public projects in Australia, the decision making process usually does not reflect any clear integration of the feedback received and the project-specific knowledge gained is often wasted. Current practice seriously lacks any model for allowing consolidation of the consultation feedback in objective decision making process (Elkington 1998, Littig and Griessler 2005). The absence of such capability clearly hinders the use of knowledge gained in community consultation from one project to the planning of the next project. Instead, the stakeholders are often considered as nuisance and their interests treated as constraints in achieving the project objectives. Consequently, projects have suffered serious set-backs due to political, social, environmental and community challenges and fail through statutory processes. (Doloi 2010a, 2010b, 2008)

2 Objectives of the Research

The aim of this research is to develop an experimental smart framework that will facilitate holistic decision making from a stakeholder perspective on capital projects. The objectives to be achieved are:

- To conduct fundamental research for constructing Social Network Models that integrates representatives of the entire stakeholder community, such as businesses, end-users and neighbourhood communities, for assessment of their interactions and relative stakes over planning, development and operation phases (e.g. life cycle) of capital projects (Doloi 2012, Doloi2011, Doloi and Almahmoud 2012, Prell et al. 2009, Pyrke 2012);
- Develop a hierarchical decision model that places benefits to stakeholders and community interests at the core of decision making, in addition to the usual resource and financial considerations (Doloi
2007; Grimble and Wellard 1997, Maignan and Ferrell 2004);
• To develop fuzzy preference models and intelligent agents for facilitating learning the functional relationships between different level of variables in the decision model, monitoring project information and associating patterns of data with opportunities/risks, determining optimality and forecasting outcomes of proposed decisions based on trends in the project data (Kumar and Ghoshal 1998, Kuo and Xue 1999, Khosla and Dillon 1997, Nwana and Azarni 1997, Festinger 1949, Gunaratnam and Gero 1994, Fodor and Roubens 1994); and

To address these issues, a number of existing research will be integrated into the information model with added capability for benchmarking of social performance and life cycle management of built environment projects.

3 Background Review

Project failures are not just related to experiencing delays or cost overruns as typically portrayed in the project management literature or the media. More significantly, failures tend to affect the relevant commercial and operational fundamentals, project fitness for purpose, adaptability, acceptance by the community and users groups and other vital characteristics. Numerous experiences of expensive but failed projects (e.g. School projects from Government’s Building Education Revolution (BER) program, Ararat Prison project, East Link project in Melbourne, Southern Cross Railway Station in Melbourne, Docklands Development, Cross City Tunnel in Sydney etc.) show the complete lack of any mechanism to incorporate these factors for ensuring the target outcomes. While development of all these projects may have been rationalised based on at least the traditional cost-benefit analysis, none of these projects have evidently resulted meeting the perceptions of the community at large and hence perceived to have failed to create appropriate value in a societal context. In contrast, however, a project such as the Sydney Opera House with significant problems in design and development did not end up being a failed project in a societal context. Instead, this iconic project added enormous value in Australian social fabric. This is clearly a good example of how perceptions of the end users (internationally in the case of the Sydney Opera House) contribute to the long term social value creation over project lifecycle.

Figure 1: Broad view of input and output variables in project development

Figure 1 shows a broad view of an ideal input-output based Information Model associated with project planning and development. As the input variables are both quantitative and qualitative, these are prone to change with time. Thus the criteria for decision optimisation must not only consist of hard functions based on well-established theories, but also evolutionary soft functions based on community preferences. The processing and learning the entire information in real time should be the key components in the model.

3 Model Development

In order to develop the model for analysing stakeholder’s needs and requirements and thereby performing relevant analysis on the available feasible solutions, an information model hierarchy of underlying variables needs to be developed. For real time execution of the model on a particular case project, the following steps need to be performed:

1. Determination of the criteria (target values set) for assessment of the project, i.e. financial targets (based on investment policy and market conditions), overall stakeholders satisfaction target and overall social value performance target (both have to be derived for each project) that must be achieved (Table 1 and Figure 2).
2. Estimation of the expected financial outlay based on market conditions, expected satisfaction levels based on stakeholders’ network, their position within the network, relative stakes and potential influence and expected total social value performance on each feasible solution is developed (current research capability already permits conduct of this part);
3. Development of ANN models using the indicators in step 2 and train the network using similar data from case study projects (empirical data collected from case studies);
4. Derivation of the expected overall financial status, overall stakeholders’ satisfaction level and the overall social value performance points respectively for each of the feasible solutions, using the results from Step 3 and preference rules (Figure 2);
5. Constructing a Decision Table with the results from all the aforementioned analyses (including financial estimates) as per Table 1 format;
6. Locating all solutions that meet the target criteria; and
7. Locating the most favourable single/hybrid solution by applying Fuzzy preference models (multi-criteria decision modelling).

The multi-criteria decision modeller will be developed based on variables derived in Social Network Models on case studies and augmented by a Project Reference Group (PRG). The PRG is formed by inviting the representatives of the various stakeholders directly or indirectly involved in the project. Fuzzy preference modelling techniques would be used to determine the relative weighting for these variables. Neural networks, of the Multi-layer Perceptron (MLP) type, would be used for learning the required set of relationships between the high level variables and the associated indicators, and for monitoring and predicting the values of these variables (Lorterapong and Moselhi 1996).

4 Application of Social Network Analysis (SNA)

SNA is a powerful method for collective representation of group of actors (such as individuals or organisations) and their dyadic ties between these actors. The social network perspective provides a clear way of assessing interactions of social entities in order for representing global patterns, locate influential actors relative to others and examine the overall network dynamics. The social network is a multidisciplinary science emerged from social psychology, sociology, statistics, and graph theory where human interactions impacting the outcome of a phenomena is quite pivotal. The first true social network concept in the field of sociology was formulated by Moreno in 1953 for conceptualising the structures of small groups of people produced through friendship patterns and information interactions (Moreno 1953). Extending the concept, Bavelas (1948) and Festinger (1949) laid a solid foundation of ‘group dynamics’ which later form the backbone of the American Social Psychology in 1960s. While the development of the concept was rapidly spreading out across numerous fields of scientific studies, the most powerful anthropological developments were associated with the work on African societies carried out at the University of Manchester which was codified as one of the first books on Social Networks in Urban Situation (Mitchell 1969).

In recent years, Social Network Analysis (SNA) is becoming increasingly popular as a general methodology for understanding complex patterns of interaction (Carrington, 2005). The network perspective examines actors that are connected directly or indirectly by one or more different relationships. Any theoretically meaningful unit of analysis may be treated as actors such as individuals, groups, organizations, communities, states, or countries. Regardless of unit level, social network analysis describes structure and patterns of relationships, and seeks to understand both their causes and consequences (Haythornthwaite, 1996).

5 Integration of Stakeholders in the Model

In complex projects with many stakeholders, application of Social Network Analysis (SNA) provides a platform for integrating their perceptions holistically in the project development process. Conceptually, there are three key elements required to understand for social network analysis. In the network, the ‘Nodes’ or ‘Actors’ are entities, persons, organisations, or events. The ‘Links’ between the nodes represent the relationships of any kind such as transfer of money, communications, friendships, exchange of resources or information etc. (Galaskiewicz, 1979). One of the key characteristics of the network is ‘centrality’. Networks may have one or several or even no central actor(s) with links from many actors directed to it, which represents high or low network centrality. A central position within the network indicates the amount of power obtained through the structure and capacity to access information and the other members (Wasserman and Faust, 1994). Thus, SNA is concerned with the structural positions (such as central, isolate, bridging etc.) of actors. If an actor has many links to others in the system, then it has different network characteristics than an actor with fewer links within the system.

In order to assess the social performance in projects, the underlying social needs and necessities within the broader communities involved (directly or indirectly) over the project lifecycle must be accurately understood. For instance, the idea of what really constitutes a decent life for a particular community in relation to a project is highly subjective and it depends on the perceptions of that community. Thus, if personal needs of a community, such as food, housing, health care, freedom and liberty etc. are combined with institutional needs such as education, recreation/leisure, social relationships etc. a much broader of action and opportunity is required in the social value creation process. The major driving force behind society and socialisation in broadest sense is the creation of opportunities to meeting one’s need and for that purpose, Malinowski (1988) suggests that categorisation of societal needs across meaningful
6 Analysis and Results

Making holistic decisions requires balancing the financial gains flowing to promoters and investors on the one hand and the benefits accruable to the community, environment and stakeholders on the other. Table 1 shows how a holistic analysis may be structured and used.

Considering a case project such as construction of a new technology park or a new highway, the solutions in Table 1 are all assumed feasible (generated by the project planners), each has an estimated incremental cost and a corresponding overall stakeholders’ satisfaction level and the total social value performance. Both 3 and 4 solutions satisfy all the expected target values set as criteria, though Solution 4 has a higher incremental cost (6.75%) while 3 has a higher return on investment (27%). Solution 1 is rejected as it does not meet any of the criteria, while Solution 2 is financially superior but will not meet the target set for the total social value performance. Selection is therefore confined to either 3 or 4. The question is whether or not the additional investment of 1.20 percentage points (6.75-5.55) of total life cycle cost is worth it, or whether or not the drop in the IRR value by 2 percentage points in Solution 4 relative to Solution 3 is acceptable, given that the latter has an increased satisfaction level by 0.5 and additional total social performance (accruable to the stakeholders and the community) of 5 percentage points of the project base value. Currently, this type of decision is normally made by the sponsors considering mainly the financial merits of each case. However, integrating the input from the extended stakeholders and community alike, the proposed model will facilitate such trade-offs to locate the optimal solution among those which meet the target criteria.

Table 1: Holistic analysis of alternative plans for a given project

<table>
<thead>
<tr>
<th>Solution (plan) number</th>
<th>Incremental cost*</th>
<th>Return on investment</th>
<th>Overall Stakeholders satisfaction level</th>
<th>Total Social Value Performance (SVP)**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target</td>
<td>Expected</td>
<td>Target</td>
<td>Expected</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>22</td>
<td>40</td>
<td>5.4</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>25</td>
<td>40</td>
<td>7.0</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>25</td>
<td>40</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>0.75</td>
<td>25</td>
<td>40</td>
<td>5.5</td>
</tr>
</tbody>
</table>

As seen in Table 1, the incremental cost is the additional cost in each solution (% of total project life cycle cost). Financial measure includes key indicators such as IRR, NPV, Cost-Benefit Analysis, Life Cycle Costing etc. Stakeholders’ Satisfaction level usually is an index from a scale of 0 for no satisfaction to 6 for best practice solution in the management of non-quantitative variables, derived through the application of Social Network Analysis and the Preference Rules to the indices for all individual variables. The Total Social Value Performance (SVP) is expressed as % of the overall (consolidated)

value contributed by the project in societal context. SVP for individual social issues can be computed at a lower level in the system (Hossain and Silva 2009, Hossain 2009).

Given the dynamic project environment, the input variables will certainly change over time and thus repeated cycles of analysis and replanning will be required. So the real benefit is the feedback that the model (SIM) will provide, and the clarity of how solutions meet the relevant targets and whether or not a hybrid solution that embodies elements of all the aforementioned solutions can be synthesised with superior performance all around. Whole of life financial modelling and part of stakeholders’ satisfaction modelling have already been successfully researched and developed by me at the University of Melbourne. Derivation of the overall social value performance as well as research and development of the proposed smart model with required intelligent functions with predictive power remain the challenge being addressed in my fellowship project.

7 Future Direction

The research is currently at a very early stage. However, given the uniqueness of the research methods and adopted approach with a clear focus of the stakeholder’s needs and requirements in decision making context, the output promises to contribute towards a sophisticated and efficient framework replacing the traditional practices. Thus, the novelty of the research is clearly evident across the concepts, aims, objectives, activities and methodologies. For the very first time, SNA will be used for objective assessment the stakeholders’ influence in the project decision process. Thus the innovation in this project is realised in number of fronts. Firstly, the adaptation of Stakeholder Theory for identifying the variables, indicators and criteria associated with project development and operational environment. Secondly, the integration of collection of powerful research methods namely fuzzy preference modelling, artificial neural network and intelligent agents, multi-criteria decision modelling for smart use of information and thereby developing strong predictive models for supporting optimal decision making in development of capital projects. For the first time, the resulting Smart Information Model (SIM) will provide a unique capability for empowering the project owners, policy makers, sponsors alike to developing socially responsive projects.

8 Conclusions

The research is aiming to tackle a very important issue of ensuring optimal value creation of capital intensive infrastructure projects in societal context. The unique mechanism of integrating stakeholders’ perceptions over project life cycle for optimal planning and development decisions in capital projects is considered to be at the
forefront in construction research. Quantification of stakeholders’ influence with respect to their relative stakes in projects and responsive planning for complete social acceptance will establish a new benchmark on success and failure of capital projects both nationally and internationally. Application of the resulting Smart Information Models in planning, development and operations of capital projects will advance the knowledge base across numerous dimensions including construction, architecture, urban planning and urban renewal and regeneration.

The SNA based approach developed in this research will help the policy makers and industry practitioners better understand the current status of development planning in relation to its capital budgeting and investments programs. The application of research in real life projects will provide convincing evidence to the policy makers and government bureaucrats on the shortfalls of current project development practices and thereby adaptation to the changes. The development of the scientific framework will add significant intuition among the researchers for further development of knowledge in the field.

References


[16] Doloi, H. (2010b) Balancing stakeholder’s requirements in project procurement – have we learnt anything yet?, CIB World Congress 2010, 10-14 May, Salford University.


Lessons Learnt from the Vernacular Architecture of

Bedouins in Siwa Oasis, Egypt

R.M. Ahmed

The Bartlett School of Graduate Studies, University College London, United Kingdom

E-mail: riham.ahmed.10@ucl.ac.uk

Abstract -

In the last few decades, climatic changes have become very important issue to investigate when dealing with sustainable structures. This paper presents the challenges made by the Bedouin residence of Siwa oasis in Egypt through the case study of three vernacular buildings, in order to identify the best practice and the most appropriate systems for climatic responsive low-carbon buildings. This paper investigates thermal insulation, best utilisation of building with the local materials, passive cooling techniques, natural sun lighting, all with the concept of best utilization of the available resources. The paper is also provides lessons learnt from the environmentally friendly case study buildings in Siwa for education, training and employment of people in rural areas with enriched environments, thus, the outcome of the paper will be used to achieve a surplus for investment and innovation.

Keywords -

Vernacular Architecture; Neo-vernacular Architecture; Sustainable Architecture; Environmental social enterprises; Siwa oasis; Egypt

1 Introduction

In the shadows of evolution of buildings’ design strategies, philosophies and techniques, the world is facing environment crises and huge energy demands. Thus, emerging carbon and greenhouse gas emissions. In this notion of environmental pollution, climate has been linked directly to the “vernacular responses” where materials, resources and production is becoming the world’s foremost source for research in the area [1]. Over the past few decades, there has been a greater demand in the field of architecture to achieve design strategies for comfortable living conditions in this challenging area, building on traditions and knowledge obtained from research into sustaining the natural environment. The majority of relevant research is focusing on the natural environment and its relationship with cultural heritage; where we found projects aiming to apply developed approaches of vernacular technologies for preserving cultural vernacular traditions. Some research exists on regionalization of vernacular architecture based on climate and cultural heritage aiming for energy saving, e.g.: [2]. However, limited data is available on the paradigm of the transformation of vernacular architecture from one generation to another with the transmission of technology. Nevertheless, by looking at the rich environments of Matmata in Tunisia, various lessons could be learnt from the ancient technique adopted of digging their houses in the caves as a mean by which they limit direct exposure to the harsh sun radiations [3]. Not only a remarkable beautiful traditional architecture have been emerged, but also such techniques comply successfully with the thermal comfort standards used worldwide now a days like ASHRAE-55 (The American Society of Heating, Refrigerating and Air-conditioning Engineers) [4] or ISO-7730 [5]. Likewise, many lessons could be learnt from the challenges met by Rajput Village in India where we can see the best utilisation of the landscape and agriculture for facing the prevailing monsoon rains, sudden flooding, rapid erosion and extremes of temperature and tropical storms [6]. From the same perspective, a recent example of neo-vernacular architecture has emerged from the wise use of technology in achieving the attractive luxuries vernacular Gouna resort in Hurghada, Egypt [7]. For this reason, this study is particularly interested in providing a review on three case study buildings in Siwa oasis in Egypt have made a major contribution to the sustainable architecture. Three case study buildings will be discussed for taking some lessons to be learnt from how the Siwan local people managed to incorporate their social enterprise dimension with the local architectural practices.

2 Siwa Oasis

Firstly, Siwa oasis is considered to be one of the ancient oases since pharaohs’ days, called the oasis “mercy islands” as they represented the resting place for travelling tribes in the desert. The ancient Egyptian name of Siwa was “Sekht-am” which means the palm land. Siwa is located in Egypt's remote western desert, about 60 feet below the sea level. The total area of Siwa is 1088 kilo meters and it contains more than 300 fresh water streams and springs, populated by eleven traditional tribes totalling 20,000 people [8]. Siwa oasis has a dry hot summer reaching 39°C and cool winter reaching 5°C [9]. Siwa is popular for its palm and olive trees due to its location and the presence of hundreds of fresh water streams and springs, thus considered to be an agricultural oasis. It is one of the few Egyptian oasis communities
that have managed to retain most of its traditional characteristics. Shali’ the ancient salt-mud brick ‘kershef’ fortified dwellings was built in the 13th century. Dwellings were built side-by-side along steep, narrow and winding dirt roads, yet largely abandoned and left to collapse. Recently, heavy unusual rains damaged the dwellings, leading the population to abandon Shali searching for more space, dismantling any building materials and fixtures they could rescue to erect new. At the same time, the newly developed initiatives did not adequately consider the impact on the environment. Land was rapidly purchased by outside investors, the social fabric of the region started to change and traditional methods of sustainable use of resources were no longer practiced [8].

As a natural response to adapt to the harsh conditions of Siwa’s arid desert environment, dwellings are generally characterised by being compact in shape for minimizing the amount of building surface exposed to the direct radiation of the sun and the alleys in between are narrow and often covered and shaded streets to avoid the heat of the sun and extreme brightness and provide ventilation shaft. According to interviews with the Siwans regarding their traditional architectural vocabularies used, another natural technique was used for cooling the air during the hot summer is the use of vegetation beside the openings and hence improving cooling the air before passing through the windows. Moreover, wind towers and atriums were used inside the houses. Also the windows were oriented opposite to each other’s for creating cross ventilation. With regards to the construction technique with Kershef blocks, salt from the salt rich soli was dehydrated via leaving it in a direct sun exposure. Then ‘Tlakht’ is used as the filling material which is again fermented wet mud from the salty soil left for one week or two to dry.

Buildings normally don’t exceed 5 meters high. Thickness of walls is normally 50cm starting from the first row in the ground reaching 30 cm thick in the last row. Currently, Kershef building technique is abundant owing to the high moisture content in the soil, so it determined insulation before the workers immense in the building. Nevertheless, during the construction phase, the builders had to build layer by layer after making sure that the sand Kershef’ blocks get dried. Hence, recently Siwans have started to replace their traditional Kershef buildings with typical white blocks and cement as a mean by they save time.

3 Case study buildings

3.1 House of a Siwan: Haj Ali

The house of Haj Ali (Figure 2), occupying around 350m²/ floor and 10 meters height. The building has been selected as it is one of the typical Siwans’ Kershef built houses which are in a good condition up to date.

![Figure 2. External view of Haj Ali’s house](image)

The building has two storeys connected by a central staircase also serving as a ventilation shaft, and a backyard sitting area for men ‘khos’ (Figure 3), a guest room ‘almarbauaa’ for visitors who visit the family on more frequent basis located close to the main entrance with a separate door from outside to insure privacy for the family member. The external walls are painted for emitting the solar radiations.

Inside the house, a small entrance lobby welcomes the visitors; family members pass through this area to access their private area upstairs. Storage room is located in the area between the entrance lobby and the kitchen, followed by the family living area so that there is no need
for workers or outsiders to penetrate the house (Figure 4).

Interestingly enough, the stair case tower is used as lighting pipe where mirrors have been installed on the walls to reflect the sun light inside the house as shown on the diagram below (Figure 5), beside acting as a cooling tower as well.

On the upper floor, bedrooms are located with a central living area in which the family member gather, eat and discuss the family issues. Besides, the bedrooms are separated in to roofed rooms for sleeping in winter and non-roofed for sleeping in summer to replace the use of any mechanical modes for cooling during the hot summer. On a similar basis, another un-roofed kitchen is located on the upper floor in addition to the one in the ground floor. The upper kitchen ‘Tabent’ is actually the one sued for cooking, equipped with a built-in Kershef cooker for minimising the use of electricity as well. The

house was built first with 50 cm high concrete wall on the ground floor level, which is not costly to build for isolating the ground water from the Kershef blocks. Then they determined the areas of the room and afterwards they started building with the thick Kershef blocks until they reached the desired height. They repeated this for each single room. Following, they supported their walls with palm wood trunks connections on roofs to achieve straight endings of the building’s walls, which was used as decorative element for interior design as well (Figures 6, 7).
As shown from this example, various low cost climatic responsive building techniques have been adopted inspired by the traditional Siwans’ houses, summarised as follows:

1- Best utilisation of the local material for climatic responsive zero carbon emission building, while using the local material for building the cooker to minimise the electrical devices.

2- Using the stair case tower as an atrium for passive ventilation.

3- Creating an affordable lighting pipe for maximum the use of sunlight shed on the house for long hours during the day.

4- Orienting the windows in a way to promote cross ventilation as a replacement for air conditioners.

5- Designing the main entrance in L-shape and locating the private rooms at the back of the house for conserving the local culture and traditions of keeping the privacy of the family and segregation between females and males.

6- Designing a beautiful yet robust palm tree trunk ceilings, matching with the layout of Kershef building and also as a material for climatic responsive zero carbon emission material.

7- Encouraging social interaction among the family members by providing central gathering areas.

8- Encouraging social interaction among the members of the society by providing outdoor sitting area for men ‘khos’.

All these combined techniques helped achieving a healthy, low-cost, thermally comfortable and well-designed house, externally and internally.

3.2 Shali Lodge and its extension Al Baben Shali

The second building proved a role example in a community sustainable designed project which helped in providing a better quality of life for the Siwans.

In an aim to achieve sustainable development in Siwa, Mounir Neamatalla (president of Environmental Quality International (EQI) in Egypt), decided to invest in projects for preserving Siwa’s wealth of natural assets and cultural heritage which in turn could reduce poverty. Thus, Shali Lodge and its extension Al Baben Shali were constructed as the first eco-tourism lodges in Siwa. Shali lodge is situated few meters from the village in Siwa and built with the local Kershef. Interestingly, no electricity was used, unlike the other eco-lodges, which promoted the awareness of less energy consumption among the society. The building technique adopted relied mostly on best utilizing the available resources. For instance; both buildings were built with Kershef and the interior walls exposed to the sun were painted in light colors for emitting solar radiations (Figures 8, 9). Also, open air atrium between the rooms has been created for passive ventilation, where we can see wide opening arches are surrounding the atrium for enhancing the wind speed. Again, the windows were orient in an opposite direction for achieving cross ventilation for replacing air conditioners. In addition, different forms of palm trees’ layout were designed for ceiling construction, inspired by the traditional Siwan technique adopted in the previous example (Figure 11).

Figure 8. An atrium in Shali lodge between the hotel rooms influencing wind movement via passive cooling and creating cross ventilation as well.

Figure 9. An open roof gathering area in Shali lodge, showing the Siwan bedwin spirit reflected on the interior façades design and integration of palm trees.

As seen on Figure 9, alcoves were grooved on the walls for storage or decorative purposes inspired by the old Siwan vocabularies. Most important, since the project determined construction using the local materials, Neamatalla had to establish a contracting company for kershef to carry out the building work where not many Siwans were skilled for the task. As a result, further demand in kershef building techniques occurred afterwards; creating more job opportunities for many Siwans. The number of workers who obtained training in building with kershef increased as well. Moreover, part of the project was dealing with proper management of Siwa’s assets, which attributed to raising the
awareness of the local people to the importance of protecting the non-renewable resources of the oasis. As a result, land reclamation was significantly limited by agriculturalists to become 25,000 acres after it was from 250,000 while also water depletion was prevented and agriculturalists started to grow organic pesticide-free crops.

Figure 11. A unique pattern of the wooden palm tree ceiling with a lamp made from salt, hung from the ceiling.

The unique wooden designs inspired by the traditional Siwan vocabularies are also applied in arches’ designs (Figure 12).

Figure 12. Decorative arches using palm trees’ trunks

Lighting lamps, furniture items using dehydrated sand. A unique pattern of the wooden palm tree ceiling with a lamp made from salt, hung from the ceiling.

Figure 13. A center piece table made from dehydrated salt blocks

Figure 14. Natural technique adopted for passive cooling and the design of narrow alleys with narrow windows from one side and wider ones from the other side for creating cross ventilation.
The project helped employment of over 45 Siwan which also helped them not only to earn wages but to gain the abilities and techniques of long-established building systems.

In addition to the best utilization of Kershef in buildings’ design, both Shali and Albabin Shali lodges helped achieving the following:
1- Enhancing passive cooling via separating the building with narrow alleys for creating wind currents, besides cross ventilation replaced the use of air conditioners.
2- Promoted the awareness of less energy consumption among the society.
3- Workshops for the community to raise the awareness for climatic responsive architecture, as a result land reclamation was significantly limited for instance.
4- Encouraging local trades inspired by the traditional building techniques like constructing a Kershef factory, and palm tree roof construction for creating more job opportunities to the society.
5- Creating a new and unique practice of dehydrating the salt furniture and lighting units inspired by the best utilization of the natural environment.

3.3 Adrere Amellal eco-lodge

Another EQI sustainable driven project is the Adrere Amellal eco-lodge. It has been selected for the remarkable blend of the building with the surrounding natural environment and landscape.

Amellal eco-lodge is considered one of the main touristic eco-lodges in Siwa. It is located in a unique place surrounded by sculpted limestone, 75 acres of palm and olive trees, salt rocks and clay, aimed to best use the available local building materials available in the surrounding environment. The hotel is at the edge of the lake Siwa in the western desert. The building has 39 environmentally friendly rooms, the rooms are designed with a style of comfort and respect to the environment at the same time. Windows were designed to be very small in size and imbedded in the very thick Kershef walls. Similar to the previous buildings discussed, climatic responsive materials used and passive cooling and cross ventilation were achieved for better climatic conditions.

Figure 15. Adrere Amellal eco-lodge blended within the mountain and within the surrounding landscape

Figure 16. Left: The small windows imbedded in the thick walls.
Figure 17. Right: Bathrooms built with stones on walls and floors

Figure 18. Sky light roof made of palm tree trunks blended with sand blocks
Figure 19. An example of the dehydrated Kershef textures which added beauty to the interior design in addition to acting as good insulators

According to a report by Hatem T. (2007), the EQI projects enabled the local people to create economic opportunities for themselves while restoring the physical environment, promoting gender equity, marketing local products to the international market, and helping position Siwa on the global stage [11]. Also, it has been reported by a Siwan that the EQI projects have respected Siwa’s culture, norms, and nature. Moreover, it used modern technology to enhance the past; referring to not using electricity which was done intentionally to make travelers experience night time and day time to allow them to go back to the natural rhythm of life, and feel harmony with nature. A better quality of life was provided to the people, including: simple, clean, good food, and fusion of state of being, that makes Siwa a unique enterprise. Furthermore, according to Abou Adel, marketing coordinator of the Siwa Initiative, Siwa was one of the poorest communities in Egypt where over 90 percent of its people were involved in agriculture, and the rest in tourism. Not until the EQI eco-lodges encouraged the local people to be engaged in other activities related to tourism, like Kershef construction, palm trees’ roof construction, sand furniture and products, the Siwans. At the same time, Siwan women were given opportunities to gain employment, a prospect that was previously unheard of in the male-dominated society [12].

A primary lesson that could be learnt from this example is the remarkable blend of the building with the natural environment using the traditional vocabularies, not only on the exterior façades, but also through the interior vocabularies used.

4 Lessons learnt

According to the case studies illustrated, it could be concluded that the local material in Siwa is cheap, naturally available in large quantities in the surroundings, durable, when applied correctly, having good thermal mass capacity for insulation and having beautiful surfaces as a finishing materials. Most important, the building technique adopting the local materials is more climatic responsive and adds the uniqueness of design and traditional spirit versus other techniques for a more sustained environment. Thus, owing to Siwa’s rich natural environment and the use of vernacular architecture in transforming Siwa’s traditional building techniques, the following lessons could be learnt from the previous examples:

1- Using the local availabilities in the building site yield the most appropriate building isolative properties for achieving thermal comfort

2- Using technology in adopting vernacular architecture vocabularies helps achieving climatic responsive innovations in the building construction filed.

3- Explaining to the local people the actions that benefit the earth compared to what may be damaging to the environment enhances climatic responsive building techniques, best utilization of the local available materials, and the creation of low-cost zero carbon building techniques.

4- Engaging the community in environment awareness programs promotes a better quality of living from all aspects of life.

5- Transforming the sustainable development project scope in the low income communities in to a business model would significantly enhance the economic status of the poor.

6- Limiting the use of air conditioning to be replaced by natural ventilation methods achieves cheaper and a more energy-saving conventional alternative for climate control.

Overall, one can say that these lessons could be taken for a comprehensive sustainable development model that can inspire other communities around the world and would significantly increase positive impacts on the poor.

References


Efficient Method of An Optimum Construction Company Supplier Selection Supported by Software

Jozef Gašparík and Peter Bažík

Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Slovakia
E-mail:jozef.gasparik@stuba.sk, peter.bazik@stuba.sk

Abstract-
An optimum selection of construction company suppliers is one of the most important processes of the top management. The most of construction processes are realised by many suppliers all over the world and during the process of construction planning it is necessary to find an optimum solution which would cover several criteria. Ensuring control over the outsourced processes does not absolve the organisation of the responsibility to meet all the customer, statutory and regulatory requirements. According to ISO 9001:2008 developed for Quality Management System the type and range of control to be applied to the outsourced process can be influenced by many factors such as a potential impact of the outsourced process on the organisation’s capability to provide a product that would conform to the customer requirements and a degree to which the control for the process is shared.

There are several factors and criteria for the efficient selection of company suppliers. Our paper analyses these: quality management level of suppliers, an offered price of the construction process or product, time of construction process realisation and other specific criteria. Using the multi-criterion optimising method and scientific synthesis a method of the efficient selection of suppliers for construction processes is proposed: it relies on selected criteria and their importance.

The next part of our research results is a proposal of our original software which allows for an optimum selection of suppliers over a short course of time. This efficient method and software is implemented and verified on a real example from a real construction practice. Application of this method and software will increase the efficiency of the construction company supplier selection from the viewpoint of the key criteria of optimising: quality, time, cost and others. This efficient method and software can be applied in any construction company which cooperates with one or more suppliers in a building process. The optimum selection of suppliers can help construction companies save time and money which can be used for a useful company development process.

Keywords-
Effective selection; Multi-criteria optimization

1. Introduction

The question of efficient management of external processes for a construction organisation and an optimum selection of suppliers of construction materials, elements and processes is highly up-to-date virtually in every organisation of main construction contractors. An external subcontractor temporarily becomes part of the main contractor and may convey either positive or negative image. Moreover, with an optimum supplier selection it is possible to save considerable financial resources which might be utilised for the development of the organisation and improvement of the employees' living conditions. The issue of an efficient supplier selection is part of ISO 9001:2008 international standard dealing with implementation and certification of quality management system, section 7.3 and of practically each philosophy focused on quality management, i. e. total quality management (TQM) [1,2], KAIZEN methods [2], re-engineering methods [3] which deals with a radical reconstruction of company processes. External processes play an extraordinary role in the respective philosophy.

Many authors in the world deal with the selection of contractors. Ekambaram Palaneeswaran, Mohan M. Kumaraswamy were focused on developing a model for a contractor prequalification and bid evaluation in design-build projects. A study written by Bo Xia, Albert Chan, Jian Zuo, Keith Molenaar has summarized twenty-six selection criteria and it has shown that although price still remains an important category, its importance has declined in the last decade as other criteria have become more important. Also, we can find the application of selection based on fuzzy theory as described by D. Singh, Robert L. K. Tiong. Charles A. Weber, John R. Current and W.C. Benton describe vendor selection criteria and methods. These include increased quality guidelines, improved computer communications, and increased technical capabilities. Specific attention is given to the criteria and analytical methods used in the vendor selection process. R.
Duolmin and V. Mininno analyzed supplier selection using a multi-criteria decision aid method and analyse how to allow for a simultaneous change of the weights (importance of performance criteria), generating results that can be easily analysed statistically, performing an innovative sensitivity analysis. The whole suppliers selection model presented (promethee/gaia techniques plus high-dimensional sensitivity analysis) can be a useful additional tool inside the final choice phase of a supplier selection process. According to the De Boer et. al. taxonomy, many decision models have been suggested for supporting the supplier selection process along its main steps (problem definition/formulation of the criteria, pre-qualification of suitable suppliers, final choice).

In the paper, we aim for a complex approach which addresses the question by listing specific real results and simultaneously gives a contractor an option to choose suitable criteria and determine their value according to a specific situation and significance of the construction.

2. A Proposal of a Method of a Building Processes Supplier

Main construction contractors ensure several external processes in the course of the construction preparation and performance. Mostly these include a supply of building materials and products or building processes. The paper is focused on the selection of building processes suppliers.

When selecting a building processes supplier, we propose a following method in accordance with the Figure 1:

- defining the building processes to be provided by an external supplier
- preparing the input data of the construction contractor for the needs of potential external suppliers
  o planned construction time-schedule
  o construction project,
  o bill of quantities etc.
- a call for bids (on the Internet and in other media)
- collecting bids, analysing them from the viewpoint of pre-defined criteria (see the following chapters)
- setting the most favourable bid for realisation of building processes
- signing the main construction contractor’s contract with an external building processes supplier based on the supplier’s bid
- monitoring the works of an external supplier in correspondence with the contract; and continuous invoicing of the works

What plays a significant role in the selection of a building processes supplier is a selection of criteria and their values. The first suggested criterion is a supplier’s capability to perform the work in time as defined in the construction time schedule. The suppliers who are not able to start performing in the required time shall be excluded from the tender. Eligible suppliers are evaluated according to a certain point system where the bid with the shortest time schedule (yet, not at the expense of quality and technological requirements) is awarded with the highest amount of points. At this point it is vital that the main contractor prepare an optimum time schedule of building processes performance and set their minimum as well as maximum requirements regarding the provision of quality of the processes.

The second suggested selection criterion is the price of the works. In this case, the main contractor and their costing clerk should establish an optimum price and set a minimum and maximum limit, e.g. 20 % from the optimal price using suitable software (CENKROS, CENECOM, CONTEC). The bids exceeding these limits would be excluded; the rest of bids would be evaluated in points and mathematically in the way that the lowest accepted price gains the highest number of points, while the highest accepted price gets the lowest number of points. Other price offers between these extreme limits are to be determined by interpolation.

The third selection criterion is the quality level of a particular building processes supplier. The most suitable method appears to be the assessment of the level of quality management, e. g. certified system of quality management, application of the total quality management (TQM), KAIZEN system, re-engineering methods, EFQM model, the best building of the year award or previous experience with realising similar constructions etc. The next suitable criterion appears to be the invoice due date. The later the invoice due date, the more favourable the situation for the construction contractor from a financial point of view.

The software to be described in the next part of the paper enables contractors to suggest more criteria deemed significant by them. Setting values for the criteria may allow considering the priorities which are important for individual contractors. Thus, contractors may objectively evaluate the best supplier for selected. As a result, negative assessment will lead to the supplier’s exclusion in future tenders.

Thus, the contractor has an option to create their own database of the most suitable suppliers on the basis of optimal selection and supplier’s assessment which may lead to future offers of cooperation.
3. Structure of Criteria and System of Their Evaluation

It is necessary to set the key criteria for every activity, hence for the selection of building processes suppliers, too. Among them are undoubtedly the quality of performed works, duration of the processes and their price. These basic criteria may be extended by additional criteria such as an invoice due date which may play a significant role in the financial management of the construction. Within the framework of a model application example this paper is focused on the abovementioned four criteria.

The scoring system of the criteria is based on the possibility or impossibility to assign points regardless of knowing the offers of other potential suppliers.

Hence, the scoring system stems from two approaches of assigning points to individual criteria. In both cases a 0 – 5 scoring scale was used where 0 stands for the worst and 5 for the best variant.

3.1. Determined Scoring System

Determined scoring system is based on knowing the content of each criterion and points assigned to it.

The scoring system in question was used with the criteria of quality and invoice due date. Table 1 and table 2 visualise the scoring system chosen for the model example:

<table>
<thead>
<tr>
<th>No.</th>
<th>A criterion related to the quality level of suppliers</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The organisation has no QMS in accordance with ISO 9001</td>
<td>0 pts</td>
</tr>
<tr>
<td>2.</td>
<td>The organisation has QMS in accordance with ISO 9001 without the certificate, yet they have positive references</td>
<td>1 pts</td>
</tr>
<tr>
<td>3.</td>
<td>The organisation has the QMS certificate in accordance with ISO 9001</td>
<td>2 pts</td>
</tr>
<tr>
<td>4.</td>
<td>The organisation has the QMS certificate in accordance with ISO 9001. Also, they have implemented an Integrated Management System (IMS) – quality, environmental, occupational safety and health management system in accordance with ISO 9001, ISO 14001 and OHSAS</td>
<td>3 pts</td>
</tr>
<tr>
<td>5.</td>
<td>The organisation has the IMS certificate in accordance with ISO 9001, ISO 14001 and OHSAS.</td>
<td>4 pts</td>
</tr>
<tr>
<td>6.</td>
<td>The organisation has either QMS or IMS certificate and applies higher forms of quality management (TQM? KAIZEN, business process re-engineering, EFQM model etc.</td>
<td>5 pts</td>
</tr>
</tbody>
</table>

3.2. Dynamic Scoring System

Dynamic scoring system lies in changing values which are unknown until their registration which derive from the bids of the rest of suppliers. The scoring system in question applies mainly to the criteria of price and time. In other words, we know both maximum and
minimum value we would assign the most or least points, yet it is unknown how many points would be assigned to other values so that proportional point distribution is kept. This point distribution may be reached with the help of an equation of a straight line.

Table 2. Criteria related to the invoice due date

<table>
<thead>
<tr>
<th>No.</th>
<th>A criterion related to the invoice due date</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A supplier requests payment in advance</td>
<td>0 pts</td>
</tr>
<tr>
<td>2.</td>
<td>Payment on the product delivery</td>
<td>1 pts</td>
</tr>
<tr>
<td>3.</td>
<td>Payment within 3 days from the invoice issuance</td>
<td>2 pts</td>
</tr>
<tr>
<td>4.</td>
<td>Payment within 14 days from the invoice issuance date</td>
<td>3 pts</td>
</tr>
<tr>
<td>5.</td>
<td>Payment within 1 month from the invoice issuance date</td>
<td>4 pts</td>
</tr>
<tr>
<td>6.</td>
<td>Payment after the 1st month, in instalments based on the contract</td>
<td>5 pts</td>
</tr>
</tbody>
</table>

Prices translated into the equation will result in prices lying on the x-axis, points lying on the y-axis. Thus the equation of a straight line will look like this:

\[
y = kx + q
\]  

(1)

Subsequently, for the scale of 1 – 5 points we get the following equation:

\[
y = 4((\max - x_i)/(\max - \min)) + 1
\]  

(2)

Modifying the equation leads to the following:

\[
y = (1 + (4\max/(\max - \min)) - (4x_i/(\max - \min))
\]  

(3)

where

\(\max\) – the bid with the highest price,
\(\min\) – the bid with the lowest price,
\(x_i\) – the price of the bid we seek the points for,
\(y\) – the number of points of each bid.

3.3. Values of Individual Criteria

Values express the importance of individual criteria and subsequently influence the final number of gained points. For the need of supplier selection a percentage expression of values was used. A percentage scale is dynamic and depends on the provider's choice of a supplier.

4. Software for Selection of Building Processes Suppliers

The software in question operates in the Microsoft Excel environment and was created by means of VBA (Visual Basic for Applications). At the beginning of the software development, it was necessary to define basic functionalities of the whole system. Therefore four basic system parts were created, namely:

- setting criteria,
- setting suppliers,
- showing results,
- printing results.

The main parts were created in object-oriented programming. It means that objects were created to which commands were programmed. The software contains various text fields which had to be programmed in a way that the set data remained in the program memory and could be used in computations at the same time. Besides, it was necessary to find a way which would enable certain data to be highlighted.

Part 1: Setting criteria

This part contains units for setting the information needed for the supplier selection. These units are (Fig.2):

- highlighting criteria,
- names of criteria,
- values of criteria,
- description of point assignment to individual criteria.
Figure 2. Part of program for setting criteria

This part is secretly linked to a calculation part where the calculations which use the set data take place.

One of the curious properties of the programme is its ability to warn a user about the failure to fulfil the needed 100 % value of criteria. The programme warns a user by colouring certain fields in red; this indicates that the sum of all the values does not equal 100 %. It is possible to write information about the names of criteria and description of how points are assigned to given criteria into empty fields in this part. The description will be consecutively shown in the information section in the ‘setting suppliers’ part. The user's advantage is always having an up-to-date description of point assignment to any criterion.

Part 2: Setting suppliers

This part includes eight tabs. Each tab is meant for one company. The required data are then written in the tabs. These data are related to the supplier selection and they are written into the ‘evaluating criteria’ part. Also, company’s identification data such as name and contact details can be found here.

In the ‘evaluating criteria’ part, the companies price offers and time needed for a service or product delivery are put in.

There are also 10 criteria which a user has set in advance with an option of point selection on a 0 – 5 point scale. When the button next to each criterion is pressed, a piece of information about the point assignment for a particular criterion is displayed. Figure 3 shows the bids of five suppliers which were analysed by applying the software.

Part 3: Showing results

This part contains a total evaluation of individual suppliers showing the number of gained points. The most favourable offer is in green colour (Fig.3).

Part 4: Printing results

Printing serves as a well-organised and well-documented evaluation of suppliers with a complex report on individual suppliers; it also serves for a printed output.

Part 5: Calculation

The invisible part comprises computations into which the data from the ‘setting criteria’ and ‘setting suppliers’ sections are entered. All the mathematical formulas are translated into computations. Among significant elements in the computational part are functions seeking maxima and minima, computational model creating dynamic scoring system and computational operations providing final number of gained points.

5. Application of the Method and Software on a Particular Model Example

Stage 1:
Input definition:
– type of construction: multi-function building,
– stage of works: processes of structural works,
– expected price: 30 mil. EUR,
– required construction time-schedule 1st March 2014 – 1st March 2015 (1 year).

Stage 2:
Publishing the call for bids and providing the main contractor’s documents for the competing suppliers:
– construction project,
– overview of building processes of structural works,
– bill of quantities etc.,
Stage 3:
Presenting five bids of the competing suppliers (see Table 3).

Table 3. The bids of suppliers for software evaluation

<table>
<thead>
<tr>
<th>S</th>
<th>Performance time (dd.mm.yy)</th>
<th>Price</th>
<th>Quality</th>
<th>Invoice due date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1st March 2014 – 1st December 2014, 276 days</td>
<td>28</td>
<td>ISO 9001 Certificate</td>
<td>14 days</td>
</tr>
<tr>
<td>B</td>
<td>1st March 2014 – 1st March 2015, 365 days</td>
<td>31</td>
<td>ISO 9001 Certificate + EFQM Model</td>
<td>1 month</td>
</tr>
<tr>
<td>C</td>
<td>1st March 2014 – 1st February 2015, 358 days</td>
<td>35</td>
<td>Technical standards</td>
<td>3 months</td>
</tr>
<tr>
<td>D</td>
<td>1st March 2014 – 15th February 2015, 352 days</td>
<td>33</td>
<td>15 years of experience</td>
<td>Payment in advance</td>
</tr>
<tr>
<td>E</td>
<td>1st March 2014 – 23rd February 2015, 362 days</td>
<td>29</td>
<td>10 years of experience</td>
<td>2 months</td>
</tr>
</tbody>
</table>

Figure 4 shows the suppliers' point values based on the analysis of their bids, the offered price and performance time schedule. Figure 4 depicts a graphical evaluation of the suppliers’ bids. One of them was excluded due to the undervalued price, the other four suppliers were analysed according to the criteria based on the input data stated in Figure 4. Figure 5 shows final results and implies that the best building processes supplier for the need of the order appears to be the supplier B.

4. Stage:
Evaluating the bids through the software according to the defined criteria, setting the winning bid and conclusion of the contract with the winning supplier.

In our model example the values of criteria were used in the following manner:
- price 40%,
- time 10%,
- quality 35%,
- due date 15%.

Figure 5. Final results

Conclusion

The research paper was focused on the increase in efficiency of external processes and optimum supplier selection according to defined criteria. The method of supplier selection and automatized system of bids evaluation through the proposed software (author: P. Bazik) leads to a transparent and objective supplier selection in the short time. Many construction companies approach this process spontaneously and mostly consider only the price. Their records on the selection process are partially kept in secret, which sometimes leads to corruption and biased evaluation. The method of supplier selection supported by the software was met with great reception and interest in the programme in the organisations with an implemented quality management system. The results of our research work are also applicable to public procurement tenders regarding buildings. In the future research work, it might be possible to develop the question of extending the selection criteria and accurately defined scoring system which would best reflect the quality of bids.

The paper was elaborated within the framework of VEGA project No. 1/0184/12
References:


SOH Estimation of Lithium-ion Batteries for Electric Vehicles

D. Qing, J. Huang and W. Sun

a,b,c School of Mechanical Engineering, University of Chongqing, China
E-mail: dtqin@cqu.edu.cn, jingyingh@qq.com and fangfines@qq.com

Abstract
Accurate estimates of the state-of-health (SOH) for rechargeable batteries provide a significant value to the management of any operation involving electrical systems. This is especially important for transportation systems, where unexpected battery performance may lead to catastrophic failures. This paper performs experiments aiming at analyzing Lithium-ion battery performances with aging due to different temperatures and charging-discharging rates, and the optimum working areas of temperature and charging-discharging current are determined. In addition, the cycle life tests of battery are launched based on the simulations of battery performances under typical urban driving cycles using ADVISOR, and after the inspection of the results, a new SOH prediction model is proposed. Finally, in comparison with the experimental results, it is shown that the proposed method could be valid and effective in estimating battery SOH.

Keywords
State-of-health, Capacity attenuation, Charging-discharging rate, Electric Vehicle

1 Introduction
Batteries are core part of many important devices and always critical to the performance of the overall system. Failure of the battery often results in catastrophic effects, especially in Electric Vehicle (EV) and aerospace systems. Due to complex vehicle operating conditions, frequent charging and discharging and high working temperature will intensify the attenuation of battery. An efficient method for battery cycle life prediction would greatly improve reliability of the power system.

In the electrified domain, researches have looked at various failure modes of the battery systems, different diagnostic methods have been evaluated [1]. Other works have concentrated more on the prognostic perspective [2]. Impedance spectrometry has been used to build battery models for cranking capability prognosis [3]. State estimation technics, like the Extended Kalman Filter (EKF), have been applied for real-time prediction of (State-of-Charge) SOC and SOH (State-of-History) of automotive batteries [4]. Automated reasoning schemes based on neural fuzzy and decision theoretic methods have been applied to fused feature vectors derived battery sensor to arrive at estimate of SOC and SOH [5].

The following sections will expand more on the chosen battery cycle life prediction model, the experimental setup, results analysis and finally the conclusions presented.

2 Overview of the experiment
2.1 Used lithium-ion cells and equipment
The cells used are lithium-ion power cells: each cell outputs a nominal capacity of 4.8 Ah and has a full charge voltage of 4V. Its chemistry is based on a natural graphite negative, a LiFePO4 positive and LiFL6 electrolyte.

The equipment applied is NBT battery test system, temperature sensors and a climatic chamber with a temperature fluctuating between -30 ℃ and 60 ℃, as illustrated in Figure 1.

Figure 1. Equipment connection diagram

2.2 Experiment procedure
The battery performance is influenced by the charging (ch) and discharging (disch) history and it can also drift dramatically with temperature variation.
Moreover, in practice load current of the battery changes when an EV (electric vehicle) accelerates or decelerates. Thus, different temperatures and charging-discharging \((ch\text{-}disch)\) rates, as well as the cycle life tests must be concerned in the experiment.

2.1.1 Different temperature test

In this paper, NEDC (New European Driving Cycle) is selected as the test driving condition, from which we can obtain the rate distribution of \(ch\text{-}disch\) current, as shown in Figure 2. On this basis, the experiment temperature is set to be 20°C, 30°C and 40°C; the \(ch\) current is 5C.

![Figure 2. NEDC ch-disch rate distribution](image)

And the overall measure procedure is described by the following steps:

1. Charge the cell with 0.5C current till the voltage reduces from nominal value to 3.0V, standing for 30min;
2. Charge the cell with 5C current till the voltage comes up to 4.95V, standing for 30min;
3. Discharge the cell with 1C current till the voltage decreases to 3.0V, standing for 30min;
4. Repeat step 2 and 3 of the experiment till the cell capacity accounts for 80% of the nominal value.

2.1.2 Different current test

In practice, the utmost work temperature of EVs is around 40°C, which hereby is set as the experiment temperature. And according to the NEDC current rate distribution, \(ch\) current are determined to be 5C and 10C, while \(disch\) current are 10C and 20C.

The charging procedure is similar to the above process except for adopting different current, the main difference lies in the discharging procedure, where the third step here is discharging the cell with 1C, 10C and 20C respectively till the voltage attains a certain value (1C vs. 3.0V whereas 10C and 20C vs. 2.4V).

3 Battery failure prediction model

3.1 Theoretical analysis of failure prediction model

Battery Life is defined as: the quantity of cycle times when battery capacity attenuates to a certain percentage of the nominal value at certain \(ch\text{-}disch\) rate, namely, battery capacity attenuation rate. Generally, the two main factors which cause the battery capacity attenuation are temperature and \(ch\text{-}disch\) current. In addition, fault model can be concluded to cumulative damage-reaction theory life model, including Arrhenius, Inverse Power Law and Eyring as well as damage cumulative model. Based on the Arrhenius model, the battery attenuation can be expressed as:

\[
\frac{dX}{dt} = f(I)e^{-\frac{AE}{RT}} = f(I)f(T)
\]  

Where \(f(I)\) and \(f(T)\) are functions of current and temperature respectively, \(AE\) is activation energy, \(eV\); \(K\) is Boltzmann constant, \(0.8617\times10^{-4}\ eV/K\); \(T\) is the absolute temperature, K.

On the integration of Equation (1), can deserve:

\[
\int_{X_0}^{X_L} dX = \int_{t_0}^{t_L} f(I)f(T)dt \Rightarrow X_L - X_0 = f(I)f(T)(t_L - t_0) \quad (3)
\]

Let \(C_t = X_L - X_0\), \(n = t_L - t_0\); then

\[
C_t = f(I)f(T)N \quad (4)
\]

In which, \(C_t\) is battery capacity attenuation rate, \(N\) is the cycle times.

Furthermore, due to damage cumulative model, the relationship between capacity attenuation rate and cycle times can be described in power function of:

\[
C_t = mN^n \quad (5)
\]

Where \(m, n\) is the model parameters, which will be determined later.

3.2 Data processing and analysis

With the aid of the testing data, the analysis of the influence of temperature and charging-discharging current on battery capacity attenuation rate \((C_t)\) is carried out through curve fitting method, as shown in Figure3 (a)-(d).
From Figure 3 (a) and (b), we can observe that temperature is a critical factor affecting \( C_r \), which has a relatively low level around 30°C, and \( C_r \) increases either the temperature arises or declines, however higher temperature has a more significant influence on battery attenuation.

And Figure 3 (a), (c) and (d) depict the impact of \( ch \) and \( disch \) current on \( C_r \), the influence of \( ch \) current is far greater than \( disch \), as illustrated in Figure(c) and (d); and \( C_r \) rises with the increases of \( ch \)-disch current.

### 3.3 Model parameters determination

On the basis of the cumulative damage-reaction theory, the parameter \( m \) and \( n \) can be obtained through fitting method.

**Table 1** Curve fitting values of \( m \) and \( n \)

<table>
<thead>
<tr>
<th></th>
<th>PT = 293K</th>
<th>T = 313K</th>
<th>T = 313K</th>
<th>T = 313K</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I = 5C )</td>
<td>( I = -10C )</td>
<td>( I = -20C )</td>
<td>( I = 10C )</td>
<td></td>
</tr>
<tr>
<td>( m )</td>
<td>0.3334</td>
<td>0.2548</td>
<td>0.6361</td>
<td>4.604</td>
</tr>
<tr>
<td>( n )</td>
<td>0.6045</td>
<td>0.6298</td>
<td>0.5885</td>
<td>0.3745</td>
</tr>
</tbody>
</table>

Table 1 shows that \( m \) and \( n \) are closely related to temperature \( T \) and current \( I \), so Equation (5) can be expressed as:

\[
C_r = m(I_{ch}, I_{disch}, T) N^n(I_{ch}, I_{disch}, T)
\]  

(6)

Afterwards, three groups of curve fitting are carried out, of which each one of the three variables \( (I_{ch}, I_{disch}, T) \) is taken as the only variable respectively with the other two don’t change.

And finally, through further amendment with the experimental data, Equation (6) can be expressed as:

\[
C_r = 0.01656 I_{ch}^{0.3428} I_{disch}^{0.1905} N^{0.1595} e^{0.94267} \cdot e^{10.149963 N^{0.0257}}
\]  

(7)

### 4 Model and Actual comparison

As described in the section, battery attenuation can be evaluated through Equation (7), and this prediction model is confirmed to be effective and feasible compared with the experimental observed results, as shown in Figure 4.
5 Conclusions

The combined curve fitting and cumulative damage-reaction theory life model method has significant advantages over conventional methods of battery capacity attenuation monitoring.

Ageing of lithium-ion cells involve in complex electro-chemical reactions internal, its state variables are hard to detect and its performance is highly ambient environment and load current related. The battery capacity attenuation model presented provides a simplified method to estimate the capacity attenuation rate, not only for the failure prediction of battery in EVs but also suitable for other battery mounted machines.

Further studies are planned to be performed, showing how more extreme conditions of temperatures and charging-discharging current may impact on the cycle life of lithium-ion batteries.

References


Robust Method for Detecting the HRI Device using RANSAC

M.S. Gil\textsuperscript{a}, M.S. Kang\textsuperscript{a}, Y.S. Lee\textsuperscript{b}, S.H. Lee\textsuperscript{b}, S.H. Kim\textsuperscript{b} and C.S. Han\textsuperscript{c}

\textsuperscript{a}Dept. of Mechatronic, Eng., Hanyang University, Republic of Korea
\textsuperscript{b}Dept. of Mechanical, Eng., Hanyang University, Republic of Korea
\textsuperscript{c}Dept. of Robot, Eng., Hanyang University, Republic of Korea

E-mail: msgill798@gmail.com, wowmecha@gmail.com, pacheon76@hanyang.ac.kr, hopezic@gmail.com, shkim83@hanyang.ac.kr, cshan@hanyang.ac.kr

Abstract -

Installation work of large inner/outer wall panel glasses increases the labor load and stress of workers and causes the danger of such accidents as falling and crane overturning. To solve this problem, Gil designed the easy handling robot system which is composed of mobile system, 5-DOF manipulator system and HRI (Human-Robot Interface) device included HRC (Human-Robot Cooperation) algorithm. During glass installation work using this robot system, detecting the position of the HRI device has great influence on the work efficiency. While working there, the construction worker placed the HRI device on panel glass randomly. Also there are many factors that disturb the detecting the HRI device. In this paper describes a method for detecting the HRI device robustly. To robustly find the HRI device, in this paper, applied the RANSAC (RANdom SAmple Consensus) and LSF (Least Square Fit) algorithm. And sensor module for detect the HRI device is composed of IR (Infrared Ray) sensor and RC servo motor. The distance between the sensor module and the circle-shaped HRI device is utilized for detecting this device. In this paper, to verify the robust method, performed laboratory experiment: for finding the HRI device put with a square-shaped disturbance.

Keywords -
RANSAC; LSF; Human-Robot Interface device

1 Introduction

Installation of large inner/outer wall panel glasses such as curtain wall installation work in construction sites is carried out at the border between the inside and outside of a building. This kind of work increases the labor load and stress of workers and the danger of such accidents as falling and crane overturning. Until now, such large glass finishing materials have been installed by workers. Except for the movement of glasses, all the work processes including assembly and installation have depended on workers [1]. This is because there is a limitation in completely automating glass installation works (access to the installation position and fitting) in the changeable environment of construction sites. Furthermore, as the high-rise buildings become larger, it is an inevitable trend that such construction materials become larger and heavier, and installations of these materials have naturally increased. Therefore, as a measure to guarantee the safety of construction workers and shorten construction period, semi-automation systems based on human-robot cooperation (HRC) techniques instead of full automation are being developed and applied [2, 3].

OKTOPUS by Materials Handling Corp. in Australia, Mobile Ergonomic Handler by Arlington Equipment Corp. in the U.S., Geko & Glass Robot Hire by GGR Corp. in the U.K., and KS Robot 280 by K.Schulten GmbH & Co.KG. In Germany are representative glazing robot systems based on human-robot cooperation technology [4-7]. Similarly, Yu et al. in Korea developed a curtain wall robot with a 3-DOF manipulator attached to the end-effector of a mini-excavator [8]. For these robots, the operator manipulates the joints of the robots. If unskilled operator uses this robot manipulation method, however, it takes long time to work and has a high likelihood of accidents. To address this problem, Lee et al. developed a ceiling glass installation robot with 6-DOF F/T sensors so that operator can directly operate the robot from the robot end-effector [9]. Furthermore, Lee applied force control to promote the safety of workers and protect glasses when operator works from the robot end-effector [10]. Force control transmits the repulsive power of materials against the environment to the operator so that the operator can safely operate the robot for such tasks as assembly. However, additional problems occurred in actual application such as sensor cost, calibration of the weight of heavy objects, and the separation of force and torque. Thus, Gil et al. proposed a robot control method in which the operator defines the rotational axis of the
panel glass and the motion of glasses is determined by the resultant force applied to this axis and a point of action on the panel glass [11]. And Gil et al. designed an easy handling robot (EHR) system for installing the panel glass such as a curtain-wall. EHR is composed of mobile system, 5-DOF manipulator system and HRC (human-robot cooperation) algorithm. This robot system is available to move and rotate in a confined space. Also an operator, using a HRI (human-robot interface) device and robot manipulator, is available to work easily [12].

In this paper describes a robust method for detecting this HRI (human robot interface) device. Chapter 2 in this paper shows the summary of the HRC (human-robot cooperation) algorithm and the HRI (human-robot interface) device. Chapter 3 describes detecting sensor module and robust detecting method. Chapter 4 shows the experimental result of this method.

2 Human-robot interface device

2.1 Human-robot cooperation algorithm

The HRC algorithm proposed by Gil et al. is described below. In general, the fundamental study of the HRC algorithm is the motion of a rigid body in Cartesian space. To make it identical to the actual glass installation method by construction workers, the instantaneous axis of rotation (IoR) and the point of action (PoA) of the glass are defined. During the installation work by construction worker, the panel glass has an IoR and PoA and the position of these axis change continuously. Gil et al. analysed the motion of panel glass during installation work by construction worker and typically deduced the 3 position of IoR. Figure 1 (a) shows the position of IoR which is located on center of panel glass. At this time, the position of PoA is determined by a resultant force of two workers. When firstly point contact is occurred between panel glass and installation plane, such as figure 1 (b), the position of IoR is located on this point. And PoA is randomly positioned on panel glass except this point. Figure 1 (c) shows the line contact by two point contact. When occurring line contact between panel glass and installation plane, the position of IoR is located on edge of panel glass. And PoA is randomly positioned on panel glass except this edge line.

When the IoR is located at infinity, the panel glass has a linear motion. But when occurring the contact between the panel glass and installation plane (point contact, line contact), it has a rotational motion. Because the operator, for maintaining the point and line contact, push or pull the panel glass.

HRC algorithm is a control method that panel glass has a linear and rotational motion by a predefined rotational axis and PoA. And a control input is two wor-
Figure 2 shows the HRI devices for HRC algorithm. In this figure, two HRI devices put on panel glass. These devices define the position of rotational axis and PoA, also it measure the worker’s force and transmit to the robot manipulator by electronic signal. This figure presents when panel glass has a rotational motion.

2.2 Detailed design of human-robot interface device

The HRI device, which measure the operator’s force, is composed of a handle, a 3-axis load cell and a suction device (figure 3). To grip the HRI device, a handle part is considered. And 3-axis load cell measure the operator’s force. So, to easily put the HRI device on the panel glass, a suction device is made up of a lever and a sucking disk. The sucking method is a screw. That is, while the lever rotates, the sucking disk expands. As a result, the IMD is attached to the surface of the panel glass. To find the IMD easily, a detecting section is designed in the circle shape with a 100mm diameter. The height of the IMD is 150mm and the allowed weight of the sucking disk is 80kg.

3 Robust detecting method

3.1 Detecting sensor module and method

In this paper, to detect the HRI device which is randomly placed on the panel glass, detecting sensor module is applied. This sensor module is composed of IR sensor and RC servo motor. IR sensor measures the distance between robot end-effector and contour of HRI device. Figure 4 presents the detecting sensor module and method.

3.2 Robust detecting method

In case of scanning the contour of HRI device using IR sensor, unexpected disturbance can be detected such as operator’s hand and reflected light. Thus, it must be robustly found the contour by detecting sensor module. In this paper, using the RANSAC (RANdom SAmple Consensus) and LSF (Least Square Fit) algorithm, it can be found the IMD’s contour by sensor module robustly. RANSAC algorithm is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers or disturbance [13]. In this paper, it estimate the mathematical model of HRI device from the observed data by sensor module. The robust detecting method is following.

1. Gathering the observed data in twice (clockwise and anticlockwise direction)
2. Apply the RANSAC algorithm to sampled data and select the data group \((G_d)\)
3. Apply the LSF algorithm to these data group \((G_d)\) and estimate the circle equation.
4. Estimate the center position from this circle equation.
5. Deduce the center position of the HRI device by mean value.

Firstly, the method for gathering the sample data is following. In figure 5, RC servo motor is controlled that it turns the IR sensor from 0(degree) to 160(degree) in clockwise (①) direction and anticlockwise (②) direction by 1 degree. And it gathers the IR sensor data in 12 times each degree and calculate the average. At th-
-is time, considering the timing delay between first output and first measurement, it gathers the data after the RC motor is rotated in 1 degree. Also it is defined that the HRI device is located from 200mm to 700mm because of IR sensor’s linearity.

In gathered data, this set contains both inliers (points which approximately can be fitted to a parameterized model) and outliers. To select the inliers from gathered data, RANSAC algorithm is applied. RANSAC achieves its goal by iteratively selecting a random subset of the original data. RANSAC is divided from hypothetical generation and evaluation. To achieve the goal, firstly, it selects the sample data randomly. And it estimate the model (circle model) using sampled data. Then it calculates the error of data with respect to estimation. After it count number of inlier candidate, if sufficiently many points have been classified as part of the set then keep this set. After enough number of iteration, it return the final estimation. Figure 6 shows RANSAC algorithm. For applying RANSAC, N (iteration number) and T (boundary line between inlier and outlier) parameter must be defined. In this paper, for detecting the HRI device, it is defined that N is 50 sample and T is 10mm.

After applying the RANSAC algorithm, applying the LSF algorithm, estimate the center position of IMD. The LSF algorithm is based on minimizing the mean square distance from the fitting curve to data points [14]. Given n points \((x_i, y_i), 1 \leq i \leq n\), the objective function is defined by

\[ F = \sum_{i=1}^{n} d_i^2 \]

Where, \(d_i\) is the Euclidean (geometric) distance from the points \((x_i, y_i)\) to the curve. When fitting circles, one parameterizes those by the equation

\[(x - a)^2 + (y - b)^2 = R^2\]

where \((a, b)\) is the center and \(R\) is the radius.

Figure 7 presents the position of HRI device’s contour and center. Where, the red line and points are the inlier data group during clockwise direction and the blue line and points are during anticlockwise direction. In figure 7, \(\circ\) is center of circle (contain blue and red), \(x\) is inlier data group in clockwise direction and \(\bullet\) is inlier data group in anticlockwise direction.

4 Experiment

To verify the robust detecting method, this chapter shows laboratory experiment. Two test is performed. First, the HRI device is placed on specific position. After-

![Figure 6. RANSAC algorithm](image)

![Figure 7. Applied the LSF algorithm](image)

Finally, using the average of two center points in figure 7, it estimate the center of HRI device.

(a) The system configuration for experimental test
The result of test - er many times scanning, it is to verify the performance of detecting algorithm. In second test, the HRI device is scanned with disturbance which is placed on different position.

Test – 1: Detecting sensor module is located on the origin point. And the HRI device is located in \( x = 300 \text{mm}, \ y = 400 \text{mm} \). the sensing range of detecting sensor module is between 200mm and 700mm. (figure 8 (a))

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>clock-wise</th>
<th>anticlock-wise</th>
<th>clock-wise</th>
<th>anticlock-wise</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>283.39</td>
<td>276.37</td>
<td>417.04</td>
<td>437.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>270.45</td>
<td>269.75</td>
<td>418.18</td>
<td>424.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>292.35</td>
<td>269.79</td>
<td>419.85</td>
<td>426.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>275.44</td>
<td>275.87</td>
<td>415.85</td>
<td>427.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>296.52</td>
<td>269.25</td>
<td>419.75</td>
<td>428.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>282.01</td>
<td>272.20</td>
<td>418.13</td>
<td>429.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In figure 8 (b), x red points are outliers and o points are inliers. The result of test the average error for X axis is about within 18mm in clockwise direction and 28mm in anticlockwise. And the average error for Y axis is about within 20mm in clockwise direction and 30mm in anticlockwise (table 1).

Test – 2: In figure 9 (a), this case shows that the circle shape and the square shape are touched with each other, then the sensor data is continuous. In figure 9 (b), this case present that the circle shape is hidden behind the square shape. In the last case (c) shows that the circle shape is away from the square shape.

In figure 10, each case presents test result in clockwise direction and anticlockwise. In figure, the blue circle is contour of the HRI device. And red points are raw data from detecting sensor module.

In case of figure 9 (a) and (c), the HRI device is not covered by disturbance. Therefore, the data for device’s contour can be gathered wholly. After applying the algorithm, it can be found the contour of the HRI device with small error such as figure 10 (a) and (c). But in case of figure 9 (b), the HRI device is hidden behind the dis-
-urbance because the detecting sensor module is in Polar coordinate. After applying the algorithm, it cannot be found the contour of the HRI device such as figure 10 (b). The result of the test, the detecting algorithm applied RANSAC and LSF can robustly find the HRI device with small error but in some cases cannot find such as hidden. Also though it is applied to same sample data group, the result can be different. This is because the RANSAC selects the sample data randomly.

5 Conclusion

Installation work of large inner/outer wall panel glasses increases the labor load and stress of workers and the danger of such accidents as falling and crane overturning. To solve this problem, Gil designed the easy handling robot system which is composed of mobile system, 5-DOF manipulator system and HRI device included HRC algorithm.

In this paper describes a method for detecting the HRI device robustly. To robustly find the HRI device, in this paper, propose the algorithm containing RANSAC and LSF. And the detecting sensor module for the HRI device is composed of IR sensor and RC servo motor. The IR sensor measures the distance between robot end-effector and contour of the HRI device. The RC servo motor is controlled that it turns the IR sensor from 0(degree) to 160(degree) in clockwise direction and anticlockwise direction by 1 degree.

To verify the robust detecting method, this chapter performed laboratory experiment. Two test is performed. First test is aimed to performance of algorithm and second test is aimed to robustness. Test result, the performance of algorithm is that it has a small error (clockwise: within 20mm, anticlockwise: within 30mm) and the robustness is that in case the HRI device is not hidden by disturbance, it can be found the contour. Also though it is applied to same sample data group, the result can be different. This is because the RANSAC selects the sample data randomly.

Acknowledgement

This research was funded by Building-façade Maintenance Robot research Center, supported by Korea Institute of construction and Transportation Technology Evaluation and Planning under the Ministry of Land, Transport, and Maritime Affairs (MLTM), the MKE (The Ministry of Knowledge Economy), Korea, Technology Innovation Program (10040180).

References


Knowledge-based Building Information Modeling (K-BIM) for Facilities Management

V. Paul C. Charlesraj
Assistant Professor, RICS School of Built Environment, Amity University, India
Email: vpcharlesraj@rics.org

Abstract -
Knowledge Management (KM) is a business management technique that promises improved competitive advantage among other benefits for an organization. The application of KM in construction is fairly researched and reported. Although Building Information Modeling (BIM) is gaining wide acceptance among Architects and Project Managers for efficient and effective design and construction management, the adoption of BIM in operations during post-construction phase such as Facilities Management (FM) is in its normative stage. As FM is knowledge and information intensive and relies heavily on historical information, a Knowledge-based BIM (K-BIM) that is developed on the basis of as-constructed information of the facility has the capability for effective and efficient FM and thereby enhance the competitive advantage of a FM organisation. Ontologies have the potential to represent the body of knowledge of the various domains involved such as KM, FM and BIM. Integration of KM, FM and BIM can also be achieved through ontologies. The proposed conceptual framework, K-BIM is an attempt to advance BIM by way of integrating KM, FM and BIM using ontologies, rather than building a system on a model built using BIM.

Keywords -
IT Applications; Knowledge Management; Building Information Modeling; Facilities Management; Knowledge-based BIM; Ontology

1 Introduction

Of late, in the quest of sustainable competitive advantage, organisations have come to realise that their competitiveness edge is mostly the brainpower or intellectual capital of their employees and management. To be more specific, an organisation’s competitive advantage depends on what it knows – how it uses what it knows – and how fast it can know something new. In order to stay ahead of the pack, organisations must leverage their knowledge internally and externally to survive. Specific KM activities help focus the organisation on acquiring, storing and utilising knowledge for problem solving, dynamic learning, strategic planning and decision making. It also protects intellectual assets from decay, adds to organisation intelligence and provides increased flexibility. The emergence of KM may be explained by the confluence and natural evolution of several factors. KM is a necessity, driven by the market forces of competition, market place demands, new operating and management practices, and the availability of KM approaches and Information & Communication Technologies (ICT).

The application of KM in construction is fairly researched and reported. Although Building Information Modeling (BIM) is gaining wide acceptance among Architects and Project Managers for efficient and effective design and construction management, the adoption of BIM in operations during post-construction phase is in its normative stage. Facilities Management (FM) is one of the major tasks involved in the post-construction phase. As FM is knowledge and information intensive and relies heavily on historical information, a Knowledge-based BIM (K-BIM) that is developed on the basis of as-constructed information of the facility has the potential for effective and efficient FM and thereby enhance the competitive advantage of a FM organisation. An integrated ontology-based KM framework for FM facilitated by BIM has the potential to promote the efficiency and effectiveness of a FM system. K-BIM is such a framework that attempts to advance BIM by way of integrating KM, FM and BIM using ontologies.

2 Literature Review

2.1 Knowledge Management (KM)

KM can be defined as the systematic and explicit management of knowledge-related activities involving knowledge-workers in an organisation to improve organisational knowledge-related efficiency and effectiveness, thereby to achieve specified organisational goals and objectives.
There has been a quite extensive research reported on the role/application of KM in construction [1,2]. The applicability and usefulness of KM in construction has been researched in strategic management of construction [3], general construction project management [4,5,6], knowledge discovery from construction databases [7], design management [8] and corporate memory for construction [9]. Carrillo and Chinowsky investigated the implementation of KM initiatives in major engineering design and construction organisations in United States of America [10]. Chen and Mohamed studied the impact of the internal business environment on KM within construction organizations in Hong Kong [11] and also the strategic importance of tacit knowledge management activities in construction [12]. It has been reported that the changes in organizational culture are critical to successful KM [13].

### 2.1.1 Ontologies in KM

Ontology is an explicit specification of a conceptualisation [14]. Ontologies can be effectively used in solutions for many KM processes, especially for knowledge representation [15]. Maedche et al. proposed an integrated enterprise-KM architecture for implementing an Ontology-based KM System (OKMS) [16]. Saito et al. described the KM technologies according to their support for strategy through an ontology development method and categorised the KM technologies based on their relationship with KM strategy [17]. Ontology-based KM frameworks have been reported for engineering design management, risk management in construction projects and competency development of construction project managers [18,19,20].

### 2.2 Building Information Modeling (BIM)

Essentially, BIM combines technology with new working practices to improve the quality of the delivered product and also improve the reliability, timeliness and consistency of the process. It is equally applicable to asset and facilities management as it is to construction. BIM provides a common single and coordinated source of structured information to support all parties involved in the delivery process, whether that be to design, construct, and/or operate. Because all parties involved with a BIM project have access to the same data, the information loss associated with handing a project over from design team to construction team and to building owner/operator is kept to a minimum. It has been reported that BIM is a suitable facilitator for KM in construction for various applications such as knowledge sharing [21], construction supply chain management [22], sustainability & asset management [23] building maintenance [24] construction defect management [25] and lean architectural practice [26].

### 2.3 Facilities Management (FM)

Facility management is a profession that encompasses multiple disciplines to ensure functionality of the built environment by integrating people, place, process and technology [27]. It has been reported that the potential benefits of using BIM in FM seem to be such as valuable ‘as-built’ (heritage) documentation, maintenance of warranty and service information, quality control, assessment and monitoring, energy and space management, emergency management or retrofit planning [23,28,29,30]. Decontamination or deconstruction processes could also benefit from structured up-to-date building information to reduce errors and financial risk, e.g. through deconstruction scheduling and sequencing, cost calculation, rubble management, optimization of deconstruction progress tracking or data management [28].

### 2.4 Summary of Literature Review

It has been observed that the application/role of KM in construction is well researched and in the other phases of the built environment projects is limited. Ontology is a potential technique for the solutions involving KM systems. Research efforts on the role of BIM in KM as well as in FM are in their normative stage. It would be interesting to investigate how KM and BIM contribute together for better FM.

### 3 K-BIM for FM Framework

The objective of this framework is to effectively facilitate the FM through the strengths of KM & BIM. K-BIM attempts to incorporate the best practices of the three domains viz. Knowledge Management, Building Information Modeling and Facilities Management as shown in Figure 1.

In 2009, a global job task analysis (GJTA) of International Facility Management Association (IFMA) defined 11 core competencies of facility managers. The GJTA included responses from facility managers in 62 countries. It is the most comprehensive to date and the first truly global survey and analysis [27]. Those core competencies are:

- **Communication** - Communication plans and processes for both internal and external stakeholders
- **Emergency Preparedness and Business Continuity** - Emergency and risk management plans and procedures
- **Environmental Stewardship and Sustainability** - Sustainable management of built and natural environments
At the very outset, development of all the above mentioned core competencies of facility management can be facilitated by adopting the best practices of KM, BIM and FM. The concept map of the proposed K-BIM framework (as shown in Figure 2) is based on the premise that BIM models of the constructed facilities would facilitate the FM processes. Hence, it is expected that BIM would play an important role in the development of the core competencies of FM as BIM is the interface between the facility managers and the knowledgebase.

There are three major components in the proposed K-BIM framework. They are (i) Knowledgebase, (ii) K-BIM Layer, and (iii) Stakeholder Interface.

3.1 Knowledgebase

The knowledgebase of the K-BIM framework is primarily consists of two components. They are: Ontology of KM and Ontology of FM.

3.1.1 Ontology of KM

The entire body of knowledge of the KM domain is represented in this ontology. This include the KM processes, knowledge domains, KM tools (techniques & technology) and knowledge mapping [20].

3.1.2 Ontology of FM

The domain knowledge of the FM is represented ontology of FM. This include the FM processes, historical information, and best practices in FM. It would also have the definition, assessment guidelines, and performance criteria of the various core competencies of FM as well as the competency mapping.

These ontologies can also interact between them based on the context.

3.2 K-BIM Layer

The BIM layer consists of the BIM models of the constructed facilities and the associated standards such as National BIM Standard - United States (NBIMS-US) [31]. These are managed by the BIM managers. The data/information present in the models are depends heavily on the knowledgebase as they are context-specific. Whenever there is a request for information (RFI) from any stakeholder, the knowledgebase is referred before returning a result. Also, any updates shall be applied to the BIM models/standards. In this way, the capability of the BIM is enriched. The data/information present in the BIM models are primarily knowledge-driven rather than information-dependent. Hence, this becomes the crucial layer of the proposed K-BIM.

3.3 Stakeholder Interface

Various stakeholders involved in the project/facility, especially facility managers, shall use this interface to interact with the K-BIM for problem-solving or decision making.
Figure 2. Concept Map of K-BIM Framework for Facilities Management
Based on the context of the query (may be related to the core competencies of FM) the I-BIM returns appropriate information derived from the knowledgebase that would facilitate the competency development of the manager as well as better facility management. This in turn, would enhance the competitive advantage of the FM organisation.

3.4 Suggested Methodology for Implementation

The proposed framework can be implemented by first building the knowledgebase. Ontologies of KM & FM shall be developed using Knowledge-Engineering methodology [32] or use some existing ontologies. BIM models shall be integrated with the existing standards & protocols to achieve the intended goals such as coordination, collaboration & communication across various stakeholders and this needs to be moderated by the BIM managers. The stakeholder interface shall be integrated with the workflow management systems to maximise the operational efficiency.

4 Summary

The proposed K-BIM framework for FM is an attempt to harness the power of KM and BIM to facilitate the FM processes. As discussed, the framework consists of three major components viz. knowledgebase, K-BIM layer and stakeholder interface. The entire body of knowledge of the KM & FM domain are modelled in the knowledgebase. Ontologies are the potential technique for design, development and update of the knowledgebase. The middle layer, K-BIM is an enhanced BIM (models/standards) that is driven by the knowledgebase. The stakeholder interface is the platform through which all the stakeholders of the project/facility would interact. This framework would also be helpful in the competency development of the facilities managers.

References


Algorithm for Economic Assessment of Infrastructure Adaptation to Climate Change

Hoyoung Jeong 1*, Hyounkyu Lee 2, Hongjo Kim 3 and Hyoungkwan Kim 4

1 Graduate Research Assistant, Dept. of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Republic of Korea E-mail: jhy0@yonsei.ac.kr
2 Graduate Research Assistant, Dept. of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Republic of Korea E-mail: sophistkyu@yonsei.ac.kr
3 Graduate Research Assistant, Dept. of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Republic of Korea E-mail: hongjo@yonsei.ac.kr
4 Professor, Dept. of Civil and Environmental Engineering, Yonsei Univ., Seoul 120-749, Republic of Korea Corresponding author (hyoungkwan@yonsei.ac.kr)

Abstract
Climate change, along with the increase of severe weathers and natural disasters, is becoming an important factor to consider for infrastructure investments. To adapt infrastructure to the effects of climate change, new design, construction, or rehabilitation methods – so-called adaptation methods – can be deployed. However, it is crucial to understand the impact of adaptation methods on infrastructure before they are actually implemented. When the economic benefit and cost are clear, asset managers can confidently make informed decisions about the priority of investment alternatives. This paper proposes an integrated algorithm to assess the benefit and cost of adaptation methods. The “integrated” aspect of the algorithm is derived from the fact that climate change effects on infrastructure can be divided into two categories. One is sudden extreme weather events caused by climate change; this sudden event leads to swift and disruptive damages to infrastructure. The other is a gradual climate change of which effects are shown over a long period of time. The algorithm combines the two different aspects of climate change to estimate the net benefit of adaption methods in an integrated manner. Future climate scenarios are first assumed and their input variables are determined for further procedures. With extreme events such as super-typhoon, the procedure for sudden failure of infrastructure is used to estimate the cost and benefit of the rehabilitation effort. Maintenance cost under gradual climate change is also estimated with the climate change adjusted deterioration curve for the infrastructure of interest. Finally, the above three steps are repeated for each year to estimate the life cycle cost infrastructure adaptation to climate change for the comparison of the costs with and without adopting the adaptation method.

Keywords: Adaptation, climate change, infrastructure, future climate scenario

1 Introduction
Previous studies have indicated that climate change increases the frequency and strength of the severe natural disasters and its effect is global (Easterling et al. 2000). Climate change would continue and intensify in the future (Solomon et al. 2007). There are two strategies to minimize influences of climate change; one is adaptation and the other is mitigation. Mitigation is an effort to reduce or eliminate causes of climate change. Greenhouse gas (GHG) is pointed out as a main source of climate change, so mitigation efforts, such as certified emission reduction (CER) policy by IPCC, are executed to minimize GHG emission.

Even though GHG emission can stay at the 1990s level, climate change’s effect would last for next generations by its inertia (UNFCCC 2009). Thus, as a practical way of responding to climate change, adaptation is a viable choice to take. Adaptation recognizes climate change to overcome or adjust and it helps to reduce adverse impact and vulnerability regardless of the scale of mitigation efforts undertaken (IPCC 2007).

Adaptation appears on different spatial and societal scales (Adger et al. 2005) including coastal, agriculture, energy and infrastructure fields. In this research, adaptation is focused on management of infrastructure for its economic assessment. Cost is one of key factors - with others such as social acceptability, ease of implementation, and long term viability - to consider when infrastructure’s adaptation methods are implemented (Agrawala and Fankhauser 2008). So this paper proposes a methodology to evaluate adaptation methods’ net benefit and damages from climate change during life-cycle of infrastructure. Due to the
infrastructure’s characteristics, this study divides climate change into two categories to properly reflect its effect on a life of infrastructure: extreme weather events and gradual climate change. For the two types of climate change, damage and benefit calculation methods are separated. At the end, net benefit (benefits - costs) of adaptation is computed through the integrating simulation algorithm. With this result, this study can offer a guide for decision making processes in monetary units.

2 Literature Reviews

Global climate change has accelerated natural disasters for the last 35 years according to the Emergency Events Database (EM-DAT)’s statistics. This statistics reported that natural disaster frequency has been increased to 400 cases in 2010 from 120 cases of 1975. Population who were affected by these disasters in 2010 is grown up 2.5 times compared with 1975 and damage cost in 2010 has become 8 times more than that of 1975. This comparison supports that it is high time to make efforts about climate change issues.

In previous studies, an assessment of adaptation to climate change is identified in three steps (IPCC 2007, Fussel and Klein 2006, 2007). They are climate impact assessment, vulnerability assessment, and adaptation policy assessment. Climate impact assessment analyzes potential damages or harms from climate change. Vulnerability assessment estimates the level of climate change risk that a region is exposed to, considering its adaptive capacity. Adaptation policy assessment evaluates adaptation method’s capacity to reduce impacts from climate change. In the view of this adaptation steps, this study is in line with the adaptation policy assessment and impact assessment in the infrastructure field.

Studies about adaptation cost assessment have been conducted on the global, regional and sectoral level (Stern Review 2006; Agrawala and Fankhauser 2008; UNFCC 2009; Richards and Nicholls 2009; Hulse et al. 2010; Hinkel et al. 2010). These studies expanded the boundary of knowledge about estimation of adaptation costs considering climate change. However, their methods were mainly top-down approach showing a picture of large scale; there was a limit in producing economic results on a smaller scale such as individual construction project or infrastructure. In this study, a bottom-up analysis is proposed to assess the cost of infrastructure adaptation to climate change.

3 Methodology

3.1 Classification of the climate change

As previously mentioned, climate change is divided into two categories in a view of infrastructure damage patterns: hard strike (sudden damage) resulting in instant rehabilitation of infrastructure and acceleration of infrastructure’s deterioration affecting maintenance parts. When damage is occurred by climate change, climate factors produce damage on infrastructure either in short term or long term. Summing up, infrastructure gets stressed by sudden extreme climate change, it leads to swift and disruptive damages, whereas gradual climate change that occurs over a long period of time causes increase on maintenance costs.

3.2 Adaptation method evaluation under extreme climate change

Adaptation cost, adaptation benefit and damages from climate change are calculated in the context of asset value. Adaptation cost is to increase capacity to climate change when infrastructure is vulnerable to the new climate. Adaptation benefit means potential cost savings from expected damage that can be caused by climate change. Damages are defined as reduction of infrastructure performance due to climate change. Adaptation cost reduces a probability of infrastructure performance failure; decrease of the probability is linked to mitigation of damages and rehabilitation costs.

3.3 Adaptation method evaluation under gradual climate change

Since gradual climate change affects performance decline of infrastructure in a long term, it is related with investments for maintenance. In other words, it has strong relations with infrastructure deterioration. With the introduction of adaptation methods, the rehabilitation cycle becomes longer and annual maintenance costs decrease. However, the adaptation method incurs a cost increase. With this trade-off concept, the net benefit of adaptation to gradual climate change can be determined.

In the first body of this simulation algorithm, information about target infrastructure and adaptation method alternatives should be summoned to set-up variables for simulation. At the following section, called climate factor simulation I, extreme climate change occurrence frequency is determined and its strength is predicted. If infrastructure is damaged by extreme climate change, the algorithm leads to adaptation decision making part under extreme weather events. In this part, the procedure for sudden failure of infrastructure is used to estimate rehabilitation costs with or without the adaptation method. It is assumed
that infrastructure condition is restored to its initial status by the rehabilitation.

After performing the above parts, the year-\(n\)'s simulation is terminated and the yearly simulation is repeated for every year in the analysis period. If \(n\) meets the total analysis period, the yearly simulation stops and the aggregated damages, rehabilitation and maintenance cost are calculated. Finally, this simulation algorithm allows estimating the life cycle cost of infrastructure adaptation to climate change for the comparison of the costs with and without adopting the adaptation method.

4 Conclusions

This paper proposed dividing the climate change effect into two categories by infrastructure's characteristics: one is extreme weather events that cause sudden damage to infrastructure and the other is gradual climate changes that bring long-term effects on infrastructure. For each climate case, a methodology for assessing damage and adaptation benefit of infrastructure was suggested. The simulation algorithm was used to assess the integrated net benefit of infrastructure adaptation to climate change. The benefit figures from the algorithm can provide a basis for asset managers to make informed decisions about infrastructure adaptation. Future study is still required to research relationships between climate change factors and infrastructure performance. It is also important to evaluate adaptation methods' effects on adaptive capacity of infrastructure.

5 References


[2] Solomon, S., Qin, D., Manning, M. et al. (2007). Climate change 2007: the physical science basis—contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate change. Cambridge University Press.


Intergovernmental Panel on Climate Change, IPCC 4th Assessment Report.


Figure 1. A simulation algorithm for estimating net benefits of climate change adaptations
Introduction of adaptation method in maintenance?

Yes

No

Evaluation of Maintenance Cost I
- With previous maintenance & repair method
  \[ M = M + M_a \]
  Here, \( M_a \) stands for maintenance cost in year \( n \).

Evaluation of Maintenance Cost II
- With adaptation method
  \[ X = X + X_n \]
  Here, \( X_n \) stands for increase of maintenance cost because of adaptation method introduction in year \( n \).
  \[ M = M + M_n \]
  Here, \( M_n \) stands for maintenance cost in year \( n \).

Initialization of Structure Condition II
- Initialization of structure condition (C) by maintenance

Climate Factor Simulation II (Gradual Climate Change)
- Determining climate factors that affect condition rating
- Determining probability distribution of climate factors
- Prediction of yearly occurrence frequency

Evaluation of Structure Condition Rating
- \[ C = C + C_n \]
  Here, \( C_n \) stands for structure condition rating change by climate situation in year \( n \) (Best, Moderate, Worst)
- \( C \leq \) Minimum acceptance level? (Maintenance criteria)
  Yes
  No

Assessment Termination of Year \( n \) and Forward to Next Year
- \( n = n + 1 \)

No

Yes

Printing out Results
- \[ D = R + L = ? \]
- \( M, X = ? \)

End

Figure 2. A simulation algorithm for estimating net benefits of climate change adaptations (continued)
Development of the BIM based process for Dredging Works

R. Heikkilä, T. Leinonen, H. Paukkeri and H. Virtanen

aConstruction Technology Research Center, University of Oulu, Finland
bTerramare Oy, Finland
cMeritaito Oy, Finland

e-mail: rauno.heikkila@oulu.fi, tapio.leinonen@boskalis.com, heikki.paukkeri@meritaito.fi, henvirtan@gmail.com

Abstract - Development of building information modelling for dredging process has been studied. In dredging area different 3-D guidance and control systems have been used already over 15 years. However, the information transfer process from early initial data acquirement phase through design and further to construction phase has suffered from information delivery gaps, undetermined modelling and information delivery methods. In the paper, a new specification for initial data modelling, BIM based design process and BIM based as-built measurement and quality control procedures are to be introduced. Most essential in the process model is to transfer the information, which is binding the parties. This will be done electrically using 3D information models throughout the total operation chain. An important result of the project is also the documentation of the current dredging work process in Finland that has not been done in the past. The new models were tested using Novapoint software as a BIM design tool and two different construction projects (Rauma and Helsinki, Finland) were used for the experiments. Future work will be also described.

Keywords - BIM, Automation, Dredging

1 Introduction

Dredging monitoring systems (DMS) have been widely used for a long time (more than 15 years). As well, survey methods and modelling applications are relatively advanced. The whole dredging project still needs improvement. There have been a lot of problems in split of work and communication. For example the official data transfer has been document-based and the electronic data provided has remained to be only unofficial additional information and thus not transferred and shared with the later work phases.

In the fairway and dredging projects ordered by Finnish Transport Agency, the undeveloped data transmission and data utilization are today aimed to be improved by applying and utilizing BIM (Building Information Modelling). BIM or Infra BIM on the area of civil engineering, means the whole of the parsed digital data of the product, in its whole life-cycle. In the model-based process, the current situation of the planning area is described in the initial model which becomes to the design model by adding the design data into it. The design model is edited to production model which is utilized in dredging monitoring systems. Furthermore, the as-designed model is compared to the as-built model, the measured data of the fairway is saved in the maintenance model, and the final situation of the fairway is described in the residual model.

Figure 1. From planning to construction – current process: approximately 200 drawings in 5 folders (length about 10 km) and a map for background from the client, in DWG format. Bathymetry data is given to contractor in ASCII format.

The aim of the research project called Dredging BIM was to investigate and develop for dredging works a new process utilizing information modelling and automation, and to publish it for the whole branch along the instruction of the client. As the sub-purposes, the following tasks were set: to research the possibilities to develop the source data, to develop the construction planning by defining the construction model, to research the possibilities to develop the gathering of as-built information, utilizing information gathered by dredging
monitoring systems, continuous sounding methods and barsweeping, and to develop the methods for digital data transfer utilizing open formats when possible.

2 Methods

A new BIM based process model (Fig. 1) has been developed for dredging works and industry. In the process model, special initial information model is created based on 3D underwater surveys and other investigations. The initial information model will be transferred to designer. The design of new waterway will be performed using specialized modelling method, which creates a design model. For production purposes a specified production model will be transferred and utilized in the dredging automation systems used. During the production, specific as-built models based on measured data will be created, and transferred to the utilization in maintenance of the waterways and further in the control of water traffic. As-built and maintenance models are based on real measured data. During the operation and maintenance period continuous 3D control surveys will be performed, these survey results are further saved to the maintenance model.

Figure 2. New BIM based model for dredging works.

One of the purposes of this first study was to develop the specification of the as-designed model for dredging, of which the usability was found out by processing and modelling the data of Rauma fairway. Moreover, the possibilities to utilize and develop the source data of dredging, the open format data transmission with IM (LandXML-based InfraModel) and infraBIM guidelines, were studied. The work focused on the phases of construction design and construction of the fairway.

Finnish InfraBIM guidelines were applied on the modelling of the source data of Rauma fairway. Details from the view of fairways were developed for the specification of these guidelines. The development work of the specification in the as-designed model was based on the InfraBIM guidelines on the as-designed modelling of roads and highways. Also, the dredging contractors were listened in order to gather the information for the development work. The as-designed model contained definitions for the digital data needed for dredging works, including fairway model, bathymetric and subsurface data, boulders and pits for buoy weights.

Figure 3. Rauma dredging project was selected as a pilot project for the experiments (The 4th largest port in Finland, deepening the current 7.5 m/10 m fairway to 12 m, costs app. 14 million EUR, estimated benefits app. 33 million EUR (30 years, 5%).

Figure 4. Production model – current process done by contractor. The contractor creates the 3D model:
edge lines, slopes, slope inclinations, buoy pits, triangulation in correct way, triangulation into matrix, crane monitoring system (three matrixes).

Figure 5. Operator user interface for 3D dredging work (University of Oulu, Terramare Oy).

Figure 6. Example of measured (echo sounding) 3D as-built model (University of Oulu).

3 Results

Concerning the source data of dredging and planning works, needs for the development its coverage, the utilization of the measurements and the modelling of investigations were found. Also the development of dredging monitoring systems is needed for commissioning the model-based process and the IM-based data transmission. Also, the specification of the Inframodel format was still insufficient unambiguous, which turned out as the computer applications processed the Inframodel data differently. The ensemble of the source data of the dredging process and its usability, quality and manageability is estimated to be improved in the future by processing and aggregating the data in the model-based way.

Figure 7. A scene to the waterway model in the design software (Novapoint). The sweep level (blue surface), the rock cuts (grey) and the soil cuts (light green) of the waterway were modelled.

Figure 8. A scene to the as-designed string line model using 3D-Win software. The file was imported using open Inframodel format.

Figure 9. A scene to the as-designed model (3D-Win software).

In the machine monitoring test the production (as-planned) definition was detected functional. Machine monitoring worked flawlessly as planned. The need for improvement was though detected in the naming of the conveyable files. In the original production definition the production data meant all the data needed for
execution. To clarify naming in the future it’s suggested that production model only consists of the structural surfaces of the fairway or harbour. The rest of the model-based and other materials should be named as production material.

It was detected that when naming the as-built model the same rules should be followed as when naming the as-planned model. In road and railway construction (as well as in building sector) as-built model is at this moment defined as a design model which is corrected with the design mistakes. At the dredging industry this is not accurate and overall as-built model is defined as the actualized seafloor model detected by Multibeam or other acoustic surveying. The as built materials should consists of as-built model, detected soil type information, boulder information, the locations of blasted areas, experiential information and other files. The maintenance material is needed when keeping up the fairways and as start-up information for later dredging works in the same area. Maintenance information consists of the seafloor model detected by Multibeam or other acoustic surveying, soil type information, boulder information, the locations of blasted areas and experiential information.

Advanced dredging contractors are able to use their existing systems and know-how for utilizing information models proficiently and effectively. The subscribers’ needs and opportunities in the model-based execution should be developed further for example using a suitable cloud service. At this moment it’s not easy to storage the information models needed by the maintenance in the current data registries of the subscriber.

When comparing to the traditional design process, BIM based design needs a little bit more modelling work. In addition of the surface models every single pit should be modelled including also feature classified information. In the dredging process used in Finland these types of surface models have been used in design process for the needs of the volume calculations and to create cross-sections as a part of traditional design process. By developing more accurate specifications and guidelines for modelling methods, the quality of design models as well as the accuracy of design work can be improved.

The open LandXML standard and the Finnish extension Inframodel (IM) are quite unknown among the dredging industry in Finland and worldwide. To enable and facilitate the change and move to the new BIM-based process, more efforts should be aimed to the information communication between BIM developers and dredging industry.

In this work, suggestions were given to improve the source data of dredging projects and the data transmission. By increasing the amount of low-frequent soundings and the soil interpretations made of them, with utilizing the other investigations and already in the early phases of the projects, the source data will be improved. As well, the MBES data is to utilize also for habitat mapping in order to improve the soil co-interpretations. Also, the IM-specification and its usability will be improved by changing the cross section parameters of the fairways and by adding the definition for point objects, such as boulders, with feature data in the definition.

4 Conclusions

A new unique Infra BIM based process model has been developed and documented for dredging works and industry. The process model can further be utilized worldwide after some additional tests. Most essential in the process model is to transfer the information, which is binding the parties. This will be done electrically using 3D information models throughout the total operation chain. An important result of the project is also the documentation of the current dredging work process in Finland that has not been done in the past.

In addition, detailed modelling specifications as well as the nomenclature and numbering of part models have been developed. Also an extension for the next open inframodel schema has been suggested adding the needed parts and features to the schema from the dredging side.

The new Dredging BIM process model has been practically tested, but not yet in the whole scope of the purpose. New experiments will be needed for the feasibility evaluation and additional development of the inframodel extension part as well as in-as-built and maintenance models. In that way confidence will be created to ensure that all necessary information is transferred in different project phases. For the wide continuation development and utilization, a common development and piloting plan would be valuable.

Open information transfer using the newest Inframodel extension will be a new challenge also for dredging industry. The saving and transferring of measured information needs to be studied more due to the typical extent of measurement data and information content. The transferring of production model to dredging monitoring systems needs software development and programming work from the industry. The final form and production of maintenance information model needs to be focused with more precision.

References


The Applicability of a Geomagnetic Field based Positioning Technique with Mobile Phone to Underground Tunnels

T. Makkonen*, R. Heikkilä*, A. Kaaranka*

*Construction Technology Research Center, University of Oulu, Finland
E-mail: Tomi.Makkonen@oulu.fi, Rauno.Heikkila@oulu.fi, Annemari.Kaaranka@oulu.fi

Abstract
In underground mines and tunnels accurate and continuous positioning of moving machines and humans has been studied extensively. Geomagnetic field based positioning is an interesting possibility for this measurement need. Using a commercial smart phone and a new calculation service provided by a Finnish company, a feasibility and accuracy assertion tests were done in Finnish underground mine. First time a standard mobile phone, as far as we know, was used in underground tunnel environment to produce continuous position data and an accuracy assessment was performed. An L-shaped tunnel area, two tunnels crossing, was selected for the experiments. A total station was used for reference measurements. Accuracy of measurement in different static and dynamic situations was studied.

The found RMSE of 7.32 meters should satisfy many needs. The use of standard smart phone for underground positioning worked fine, but more development work is needed it to be usable in everyday mining situations.

Keywords - Magnetic field; Underground positioning; Tunnel

1 Introduction
The need of continuous positioning in underground mines is needed among other things for safety, logistics, machine operations and machine automation [1], [2]. Position technologies are widely applied to above ground level tasks mainly because GPS is so effortless to use. There has not been commercial service which could position underground effortlessly.

The best position method currently is obviously a total station and they are widely used, however they need to be set up and moved according to the positioning needs. The narrow corridors add challenge for setting up the device and also the requirement of direct view complicates measurements.

Several positioning solutions have been proposed for the underground mines, such as radio signal based positioning [3], [4] and the use of RFID [5], [4]. CSIRO has made lot of work with their Inertial Navigation system to advance the field of automated continuous mining equipment [6].

A geomagnetic field based positioning technique for the underground mines was suggested and tested by Haverinen and Kemppainen [7]. The magnetic field of Earth is not constant, but has a number of anomalies inside the buildings [8] as well as in mines when magnetic minerals are present. Measuring these anomalies and generating a magnetic field map, one can use this data to successful positioning.

Haverinen and Kemppainen made a positioning test underground with an accuracy of about 1.5 meters. They used a magnetometer sensor array at approximate length of 1 meter. [7] Their work with geomagnetic positioning technique lead to company called IndoorAtlas which has a business model to provide web based service for position calculations using Monte-Carlo Localization from recorded geomagnetic maps and tools in mobile phone to do the mapping inside the buildings. In most mobile phones there already is a three-axis magnetometer costing less than one dollar making the device a magnetometer [9].

Magnetic field positioning indoors has been studied increasingly in recent years [10]–[15].There were not prior test with geomagnetic positioning underground using mobile phone thus the results are quite unique.

The measurements were done in the Outokumpu Kemi mine, a chromium mine producing 2.7 million tons of ore per year in northern Finland, about 500 meters below the surface.

2 Measurements and setup

Main tools for measurements were:
1. IndoorAtlas web service for creating and maintaining maps
2. IndoorAtlas Mobile APK for creating magnetic maps and visual testing on the site
3. KemiMineDataAquisition APK mobile software for recording the location data
4. Android Nexus 4 phone
5. Kemi Mine wireless network

Because of the novel use of cell phone to get position data undergrounds a group tests were performed using setup seen in Figure 1. Performed tests where: A. Visual inspection of usability B. Positioning error on the central line of the tunnel, where the mapping was made in 10 separate locations. C. Positioning error on the central line of the tunnel, when walking distance was increased by including every second point. D. Positioning error on the side of the tunnel in 10 separate locations. E. Initial tests with vehicles for future use.

Figure 1. Schematic map about the experiment

2.1 Principle of magnetic field based positioning

The magnetic field is usually presented as a constant field as in Figure 2. Magnetic field has local variation as well as time based variation. Figure 3 shows a more localized map of the size of 5 km x 5 km, where variations are clearly seen. Finally on the mine tunnel level (Figure 4) we see the same variations on the magnetic field. Using this data as a map a position can be calculated. Fluctuations on the magnetic field can be both natural or generated by steel and reinforced concrete structures, electric power systems, electric and electronic appliances [8]. In the mines sufficient variations in magnetic fields to enable magnetic field positioning is still a question mark as the man made magnetic constructions and local environmental variations are different around the globe but it has been shown to work now at least in two mines in Finland.

It has to be noted that the magnetic map is not unique in given location, but algorithms are used to derive position when object moves through the map. More there is anomalies the easier the location is to define [10]. Widely used algorithm in robotics to this task is Monte Carlo localization.

Figure 2. Earth’s magnetic field. Source: Wikimedia Commons. Image by Zurek [16]

Figure 3. Magnetic map [μT] of the Pyhäsämi mining area in Finland. Image courtesy of J. Haverinen and A. Kemppainen [7]

Figure 4. Magnetic field map [μT] of the tunnel. Image courtesy of J. Haverinen and A. Kemppainen [7]
2.2 Data acquisition preparations

Pre data acquisition a map of the studied tunnel was imported to the IndoorAtlas service. Three points from the map were chosen as wide apart as possible (600 meters) and were mapped to correspond to known locations. Points in the picture were chosen as carefully as possible using image viewer with full zoom to get xy-pixel coordinates. Then these points were referenced to known reference coordinate points (EUREF89) using the IndoorAtlas web service, see Figure 5. The image was automatically resized when moved to the cloud by the web service and thus scaling with image size was calculated to three reference picture points.

Two sources of error are present here. Firstly, selecting the correct pixels contains an error and secondly, reference points are transformed from mine specific coordinate system to KKJ (an older Finnish standard) and then again to EUREF89. Further the IndoorAtlas uses WGS84-coordinate system and there is an error element between the two of magnitude of 30 cm in the year 2009 [17].

![Figure 5. A map of the tunnel used in the measurements as seen in the IndoorAtlas web service. Three reference points to fix map to correct location was used.](image)

A program, KemiMineDataMain, was created to Android phone by modifying IndoorAtlas API Example 0.53, see Figure 6. Data saving, time stamp and measurement running index were added. The programming language JAVA was used with Eclipse IDE. The program will list current time, number of measurement, roundtrip time to the servers, Lat and Lon in WGS84, X and Y in meters from the map’s upper corner, I and J pixel coordinates again from the map’s upper corner, heading in degrees, and server side estimation of probability about the correctness of the position.

On the site initial mapping was performed using Android 4.4.2 LG Nexus 4 phone and software IndoorAtlas Mobile 0.6.2.1382 APK. Figure 1. shows the area mapped. The area was divided to two lines AB-BC and magnetic field recording and validation was made.

Two lines of reference points were measured. Middle line was as close of the recorded and validated line (Figure 1) as possible and the second reference line was measured to the very side of the tunnel, of where we assume the largest errors are due the measurement and validation method.

![Figure 6. KemiMineDataMain. Android data collection software running on Nexus 4 phone.](image)

2.2.1 Measurement A

Visual inspection of usability was performed using IndoorAtlas Mobile APK where researchers and mine worker walked around the mapped area and made their observations (Figure 9). In this experiment a map of the area was presented and a person’s location was moving just like it does in everyday GPS-based localization software.

2.2.2 Measurement B

Positioning error on the central line of the tunnel was measured next. Ten reference points were measured with Leica Viva R1000 robotic total station (Figure 8.) created in an L-shape (Figure 1) with 4 points in the cross tunnel and 6 in the main tunnel. These points were measured in the Kemi mine coordinate system. A total of 40 measurements were made, every point four times,
along the path 1 to 10 - 10 to 1 - 1 to 10 - 10 to 1. First half of the data measurements were made using static so that the phone stays totally still for one minute. After realizing the results do not change if the movement of the phone is small, the rest 20 points were measured keeping the phone still by hands only the time that was need to take down the measurements.

2.2.3 Measurement C

Positioning error on the central line with increased walking distance was performed using the same 10 reference points than in measurement B. The distance was increased because we made an assumption that system will perform better this way due its location principles. A total of 20 measurements were made in order (see Figure 1.) 10-8-6-4-2-1-3-5-7-9.

2.2.4 Measurement D

Positioning error on the side of the tunnel was performed along the path 11 to 20, see Figure 1. The points were measured as close to the tunnel wall as reasonable. Procedure was same than what was used in the Measurement B. A total of 10 reference points were used with two measurement each.

2.2.5 Measurement E

Vehicle tests were performed as an initial preparation for full tests in the future. A magnetic field magnitude was measured at the distances of 15 m, 10 m, 5 m and 0 m away from the operational mining drill (in Figure 8.) using mobile application on Iphone4, with ten measurements on each point. Then a magnetic field was measured using the same application when mine van car drove from the 10 meters to 10 cm from the device and also by visually checking the position data. Also xy-coordinates were recorded 3 times when the mine van car drove from about 5 meters to the close proximity of magnetic field positioning device.

2.3 Coordinate transformations and error analysis

Error calculations were performed in the Kemi mine coordinate system (KMCS) and are presented in meters. The procedure to calculate errors was same in Measurements B,C and D.

For calculating coordinate transformation to KMCS from WGS84 three known points were used as far apart as possible on the mine map (Figure 5) and mapping was done using six parameter affine transformation, further, ETRS89 and WGS84 were treated as a same coordinate system. Kemi Mine uses wide net of reference points for total stations which are known in KMCS and KKJ. These points were used to check the calculated transformation with the standard transformation KKJ to ETRS89.

Measurement reference points (Figure 1) are known in KMCS, but measured positions with the phone are acquired in WGS84. Again using assumption given earlier and a six parameter affine transformation a transformation chain WGS84-ETRS89-KMCS is ready.
These values are then compared to the values measured with a total station in the KMCS. Root-mean-square-errors (RMSE) were calculated, Equation (1) where \(\hat{x}_i\) is predicted value and \(x_i\) reference value.

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}}
\]

(1)

Average error, also in this case know as mean bias error (MBE) [18], is calculated using equation 2.

\[
MBE = \frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)}{n}
\]

(2)

For precision evaluation standard deviation is calculated. Because mapping was done on the central line of the tunnel, the results on the Measurement D will suffer a bias as the measurements are projected to this mapped line. There seems to be no difference in Measurement B and C, thus combining these measurements for a larger measurement group is used as a main data set to study error.

Additionally errors for distance \(d\) from \(\Delta x\) and \(\Delta y\) was calculated, defined in Equation (3).

\[
d = \sqrt{\Delta x^2 + \Delta y^2}
\]

(3)

For errors in vehicle test average magnitude was used when measurements were made. The visual inspections were performed as subjective opinions.

### 3 Results

Visual inspection of the usability of the geomagnetic field based positioning technique with mobile phone 500 meters underground gave good results. The position presented graphically in the screen of the smart phone was deemed satisfactory for positioning a person underground by all four testers. In Figure 9 a positioning view is shown, where position of the phone is shown inside a blue circle. Notice how predicted position circle gets smaller, when convergence error is estimated to get smaller.

Root-mean-square error was calculated for Measurements B, C, D and combined B+C and also total distance error \(d\) (Table 1). Other measurements include: standard deviation (Table 2), Maximum errors (Table 3) and MBE (Table 4).

<table>
<thead>
<tr>
<th>Measurement</th>
<th>(x(m))</th>
<th>(y(m))</th>
<th>(d(m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>6.27</td>
<td>3.12</td>
<td>7.00</td>
</tr>
<tr>
<td>C</td>
<td>7.51</td>
<td>2.47</td>
<td>7.90</td>
</tr>
<tr>
<td>B+C</td>
<td>6.71</td>
<td>2.92</td>
<td>7.32</td>
</tr>
<tr>
<td>D</td>
<td>8.83</td>
<td>3.56</td>
<td>9.52</td>
</tr>
<tr>
<td><strong>Table 1. Calculated RMSE values</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement</th>
<th>(x(m))</th>
<th>(y(m))</th>
<th>(d(m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>4.88</td>
<td>2.46</td>
<td>4.23</td>
</tr>
<tr>
<td>C</td>
<td>4.57</td>
<td>1.47</td>
<td>4.32</td>
</tr>
<tr>
<td>B+C</td>
<td>6.70</td>
<td>2.89</td>
<td>4.22</td>
</tr>
<tr>
<td>D</td>
<td>4.57</td>
<td>1.47</td>
<td>4.32</td>
</tr>
<tr>
<td><strong>Table 2. Calculated standard deviation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement</th>
<th>(x(m))</th>
<th>(y(m))</th>
<th>(d(m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>19.98</td>
<td>7.20</td>
<td>20.62</td>
</tr>
<tr>
<td>C</td>
<td>19.15</td>
<td>4.84</td>
<td>19.48</td>
</tr>
<tr>
<td>B+C</td>
<td>19.97</td>
<td>7.20</td>
<td>20.62</td>
</tr>
<tr>
<td>D</td>
<td>16.45</td>
<td>5.01</td>
<td>17.20</td>
</tr>
<tr>
<td><strong>Table 3. Calculated maximum errors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement</th>
<th>(x(m))</th>
<th>(y(m))</th>
<th>(d(m))</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>3.39</td>
<td>3.09</td>
<td>4.17</td>
</tr>
<tr>
<td>C</td>
<td>4.88</td>
<td>2.46</td>
<td>4.23</td>
</tr>
<tr>
<td>B+C</td>
<td>0.36</td>
<td>0.36</td>
<td>5.98</td>
</tr>
<tr>
<td>D</td>
<td>-7.56</td>
<td>3.25</td>
<td>8.48</td>
</tr>
<tr>
<td><strong>Table 4. Calculated MBE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The transformation error on the measurement area between KMCS and WGS84 was on average 0.04 cm and -2.58 cm for \(x\) and \(y\) directions, with standard deviation 0.03 cm and 0.21 cm.

The averages for magnetic field magnitudes when approaching drilling machine were: 54.2 µT (15 m), 49.6 µT (10 m), 53.7 µT (5 m, another vehicle close) and 44.3 µT (0 m).

Mine van approaching gave average errors 1.73 m and 1.06 m, for \(x\) and \(y\) directions.

### 4 Conclusion

The experiments and measurements for magnetic field based positioning error using mobile phone showed great promise. Talks with the Kemi mine staff revealed that first encounter with this positioning technique was pleasant. RMSE value of 7.32 m on the central line of the tunnel was calculated and for unmapped area (side of the tunnel) 9.52 m. The standard deviation behaved nicely between 4.22 - 4.23 meters between the different measurements. This accuracy is
enough for safety and logistic operations. However, when we compare the accuracy to earlier work [7] we see a quite large difference of almost 6 meters. The obvious difference between the two is the measurement device, in this measurement was done using standard smart phone, in earlier work a specially designed sensor array was used.

The max distance error is quite large and reason for that is unknown, it can be a glitch or Monte-Carlo Localization was performing badly due magnetic field being too monotone or a convergence error is possible too. Also one possible source for it is a break in a data connection.

The vehicle experiments showed also great promise to the future work. There seemed to be no problems when vehicle was near, although we do have to remember that these tests were not performed in-depth enough to answer if vehicles are a problem to the method. There seem to be some changes in the magnetic field near the vehicles.

To work the measurement device needs a constant internet access; this is possible in the Kemi mine but not necessary in all other mines. To make system work in all possible scenarios underground we feel that the magnetic field positioning calculations should be performed on the phone. This also would make system safer for disturbances like fire and landslide not to mention everyday problems every wireless network goes through time to time.

The positioning system is also, at the moment, only optimized for people walking. When the distances are great, like in mines, this is not convenient. However, when it’s now proven that geomagnetic positioning with mobile phone works underground, there shouldn’t be any technical or financial obstacles to overcome this limitation. There could be for example a fixed unit on the vehicle completed with odometry.

Currently there is no support for automatic floor selection, this could be solved in future versions and should be solvable already by combining positioning techniques like wireless base station selection for the map.

We hope that this positioning technique will be seen in mines in the near future in daily usage.

References


Geotechnical Monitoring for Safe Excavation of Large Rock Cavern: A Case Study

A. Mandal\textsuperscript{a}, C. Kumar\textsuperscript{b}, A. Usmani\textsuperscript{b} and A. Nanda\textsuperscript{b}

\textsuperscript{a}Department of Civil Engineering, VNIT-Nagpur, India
\textsuperscript{b}Engineers India Ltd., New Delhi, India

E-mail: amandalthesis@yahoo.com, chandan@eil.co.in, altaf.usmani@eil.co.in, ananda@eil.co.in.

Abstract - The present study highlight the methods used to monitor the geotechnical behaviour and performance of the cavern during excavation in order to verify and, if necessary, adjust the rock support to ensure safe construction at every stage. The present geotechnical monitoring practice for underground structures involve convergence monitoring with the help of optical targets and rock mass displacement with help of bore hole extensometers. More importantly it involve detailed planning to finalize the position of each monitoring instruments based on location and orientation of geological features. Further it is equally important to ensure proper recording of monitoring data in order to analyze and take immediate action in case of any adverse situation. This requires a dedicated high end automated software, which can record, analyse and produce significant results out of large quantity of recorded data which otherwise turns out to be only ravage. Initially geotechnical monitoring was carried out in recently excavated zone of cavern on daily basis. Further based on continuous monitoring data for at least one week, frequency of monitoring is further decided. In most of the cases the deformation of rock mass was quite less compare to the alarming values which were evaluated based on detailed design for different rock classes.

Keywords - Underground excavation, Cavern, Tunnel, Geotechnical Monitoring

1 Introduction

This paper discusses in detail geotechnical monitoring activity carried out during construction of an underground unlined rock cavern complex for storage of crude oil at western part of India. The total storage capacity was 2.5MMT of crude oil. As four numbers of rock cavern units were to be built, for the ease of construction, the caverns were divided into two parts. Each part consists of two U shaped storage units with two legs. Each leg was extended almost 700 m, thus total length of cavern was around six kilometres. This project involve the excavation of vertical shaft for crude intake and out let, access tunnel to facilitate excavation at required depth, water curtain tunnel for continuous ground water recharge and the large cavern for storage space. However, present study only concentrates on the geotechnical monitoring activity implemented inside the cavern as it was most critical due to its large size and extent. Each U shaped cavern unit (in plan) is designed with 30.0 m high and 20.0 m wide (in cross section) to create the required storage space. On and average cavern crown was 60.0 m below the ground surface with a trend of 60\textdegree/240\textdegree.

As this construction involves lot of excavation activity at different depth and direction, it was very essential to ensure global stability of this zone along with the stability of each structure. Any underground excavation involves certain amount of geological uncertainties. The more the extent of such excavation more the uncertainties have to deal with. So, the available option to handle such situation was continuous geotechnical monitoring, Geological assessment and suitable design provision to incorporate different geological conditions during excavation.

The support philosophy of all underground openings within the project is based on staged excavation, incremental installation of rock support measures and verification by monitoring. In addition robust design approaches was followed to
accommodate modification of rock supports based on actual characteristic of rock mass, thus leading to cost effective and practical rock support.

2 Geological Assessment of Site

Geological and geotechnical investigation was carried out in three phases during the period 2005 to 2010. In total 20 nos of bore holes were drilled in vertical and some of them in inclined orientation in each phases. Initially six numbers of boreholes were drilled in the year 2005 at some specific locations based on the information available from geophysical survey as well as out crops observed during site visit. The data obtained from first phase survey was used to develop tentative Geological map of this site projecting the extension and orientation of different features. Based on this map, second phase of investigation was planned which includes another six bore holes to ensure the possibility of projected geological features. Finally detailed geological map was developed after completion of third phase investigation which includes additional drilling of eight numbers of bore holes. This geological map kept on updating as excavation progress incorporating the geological information obtained from excavated zone. Further this map was used to plan the location of optical targets and bore hole extensometer along cavern section for continuous monitoring of critical area.

2.1 Geology

This area was located within 'Peninsular Gneisses' (M.S. Krishman, 1953). The Peninsular Gneisses are characterized by heterogeneous mixture of different types of granite intrusive into the schistose rocks after the latter were folded and crumpled. Within the project area, Peninsular Gneiss is by far the most represented rock type. However, types of granite are also found in some places. The details of the structural discontinuities within the rock mass as determined from the mapping reveal 3 significant joint sets, 1 minor joint set and the foliation as show in Table 1.

Generally it was expected that very poor to poor quality rock will be encountered up to a depth of approximately 3 to 5 m below ground level, whilst fair and good rock is anticipated greater than 20 to 30 m below ground level. The local geological conditions of the project site are characterized by three geotechnical relevant rock mass layers of varying thickness.

Table 1: Observed Geological features

<table>
<thead>
<tr>
<th>Features</th>
<th>Dip Direction</th>
<th>Dip.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint set 1</td>
<td>070</td>
<td>85</td>
</tr>
<tr>
<td>Joint set 2</td>
<td>100</td>
<td>85</td>
</tr>
<tr>
<td>Joint set 3</td>
<td>200</td>
<td>85</td>
</tr>
<tr>
<td>Minor Joint set</td>
<td>045</td>
<td>45</td>
</tr>
<tr>
<td>Foliation</td>
<td>020</td>
<td>85</td>
</tr>
</tbody>
</table>

The top layer is varying in thickness and consists of residual soil and lateritic. Immediate second layer is characterized by weather, jointed and fractured rock with boulders with varying thickness and bottom layer is hard fresh and massive rock. Most of the project component is located in the bottom layer (bed rock) except initial portion of access tunnel and shaft. No major tectonic event has occurred in this zone subsequent to formation of these rocks.

2.2 Geotechnical Assessment

The main geotechnical failure modes expected in the caverns of this storage complex are sliding of blocks/wedges along discontinuities, toppling and rotational failures and the development of overbreak in and around zones of fractured and weather rock. Weathering and ground water are deemed to have had little influence on the regional rock mass condition. The main influencing factor is the quantity and quality of discontinuities. A number of small fault like features were recorded during the investigation. However, it was expected to have little impact of these faults on rock mass behaviour in most cases. No major pre existing inactive and active mass movements were detected. However, this area is classified as Zone-III with seismic coefficient of about 0.04g as per Seismic Zone map of India (IS 1893).

Relatively high horizontal stress ($S_h$) in the range of app. 6 to 9 MPa was envisaged along the cavern roof. For the typical vertical stress at the site of about 4 MPa, corresponding approximately 50 to 80m of overburden, the typical values of the in-situ stress ratios was in a range of 2.2 MPa.

3 Geotechnical Monitoring Plan

In this project two differ types of instrumentations were used to measure the deformation i.e.
convergence monitoring by optical targets and rock mass deformation by extensometer. Convergence monitoring was carried out along the cavern section. However, deformation measurement by extensometer was only limited to geological hot spot as observed during investigation and subsequently substantiated during the excavation of other components (access tunnel and water curtain tunnel) of this project, specifically from the excavation of water curtain tunnel which was located just 20.0 m above the cavern crown.

3.1 Optical Targets

The deformation monitoring of excavation surfaces of cavern using 3D measurement of optical targets (in x-y and z direction) was the main geotechnical monitoring method in order to assess the rock mass behaviour during excavation.

Optical targets were fixed to reference points through bolting to the rock surface. The bolt was installed through a drill hole of diameter 25 mm up to a depth of 220 mm from rock surface. Then bi reflex target was fitted to the installed bolt with the help of the screw. Monitoring section within the caverns are typically arranged at every 25.0 m of interval with five targets within the top heading (one in crown and two at each side) and 4 targets along each side wall as shown in Fig. 1. However, additional section of optical targets was installed based on observed geological features in updated geological map.

All three optical targets at the crown location were installed immediately after the completion of necessary support activity in pilot tunnel and other two targets at heading were installed after completion of first bench excavation. The targets at the wall were installed after completion of subsequent bench excavation. Initially at few locations, the damage of optical targets was reported during first bench excavation. Therefore, necessary precaution was taken to protect the targets by removing it from the bolt during blasting activity and placed thereafter.

3.2 Extensometer

Extensometer was installed in drilled boreholes from the invert of water curtain tunnel at some geologically critical locations as identified from updated geological map. This water curtain tunnel is located just 20.0 m above the crown of the cavern which will be filled with water during operation to ensure continuous flow of water around caverns to maintain require hydrostatic pressure around the cavern for confinement of crude oil. This tunnel is extended along the cavern length.

At each location two different extensometers were installed toward the cavern on either side of the water curtain tunnel as shown in Fig. 2. Extensometer was installed in drilled boreholes with three point instrument at 5.0 m interval thus maximum length of 20.0 m and furthest point is approximately 5.0 m away from cavern roof. This will monitor the rock mass movement at crown of the cavern which was envisaged as most critical part from stability point of view.

Each extensometer rod was grouted where the rod end is located in rock to develop strong bonding between the rod and surrounding rock mass. Thus minimize the chances of slippage between extensometer rod and surrounding rock. The head of the instrument was housed within a steel protection cap to avoid disturbance and damage from different construction activity. The sole purpose of these extensometers was to monitor the rock mass behaviour during construction.

Once the cavern was fully excavated and its stability was further ascertained through the stabilized extensometer monitoring data, the instrument was abandoned and all void within the instrument was grouted with cement grout to avoid any potential passageway of water towards the cavern crown.

It is usually works based on the assumption that the deepest anchor is in stable ground and any change in anchor spacing is interpreted as rock mass
movement. Here, the topmost anchor point (close to water curtain tunnel) was considered as in stable ground and will indicate any relative movement of rock in cavern crown during cavern excavation which will be recorded in it’s head assembly.

Fig. 2: Location and extent of Extensometer from Water Curtain Tunnel

4 Geotechnical Monitoring Procedure

Monitoring team comprises Project manager, Shift Manager, Site Engineer and Surveyor worked till the end of construction which completed in the year 2013 involving the underground excavation of 34.9 Lakh cum rock without any accident related to stability of rock.

4.1 Convergence Reading

Reading from optical targets was taken with the help of total station survey. The control points were fixed with reference to the original project co-ordinates at every section where the targets were fixed. First set of readings shall be zero readings of the targets. Data of all the readings were stored in the memory module of total station automatically and then down loaded it to the computer for further processing.

4.2 Extensometer Reading

Grouted anchored type borehole extensometer composed of head assembly, vibrating wire displacement sensors, connecting rod assembly and anchor or reinforcement bar were installed. Extension rods were installed in to the bore hole up to the required depths. Then the extensometer sensors were fixed after four days of grouting and initial reading were taken from read out unit. Data were stored and plotted to find the trend of rock mass behaviour.

4.3 Monitoring Frequency

The data manually obtained from optical targets and extensometers were entered and saved electronically in spreadsheet and plotted with time (frequency).

Table 2: Monitoring frequencies for convergence monitoring by Optical targets.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Frequency</th>
<th>Condition-1</th>
<th>Condition-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical Targets</td>
<td>Once per day</td>
<td>Within 30.0m from the excavation face</td>
<td>The differential deformation between the current and the previous reading is more than 5.0 mm.</td>
</tr>
<tr>
<td></td>
<td>every other day</td>
<td>within 30.0m from the excavation face</td>
<td>The differential deformation between the current and the previous reading is more than 2.0 mm.</td>
</tr>
<tr>
<td></td>
<td>once every week</td>
<td>between 30.0m and 60.0m from the excavation face</td>
<td>The differential deformation between the current and the previous reading is larger than monitoring accuracy</td>
</tr>
<tr>
<td></td>
<td>once per month</td>
<td>excavation face is more than 60.0 m away</td>
<td>The differential deformation between the current and the previous reading is within monitoring accuracy</td>
</tr>
</tbody>
</table>
These frequencies were mentioned in Table 2 & 3 with different condition. However, this frequency was further modified based on actual behaviour of rock mass observed after plotting and interpreting the recorded data. Observed data as per given frequency was plotted and uploaded to the centralise server and all concern person could have the access to the server whenever required. This was followed by the interpretation of all monitoring data by design engineer to assess the overall stability at certain stage to decide and verify further excavation strategy.

Table 3: Monitoring frequencies for Extensometers.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Frequency</th>
<th>Condition-1</th>
<th>Condition-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensometer</td>
<td>Once per day</td>
<td>Excavation face in cavern gallery is within 40.0 m of the instrument</td>
<td>The differential deformation between the current and the previous reading is still more than 5.0 mm</td>
</tr>
<tr>
<td></td>
<td>Once every week</td>
<td>Excavation face in cavern gallery is more than 40.0 m away from the instrument.</td>
<td>The differential deformation between the current and the previous reading is larger than monitoring accuracy.</td>
</tr>
<tr>
<td></td>
<td>Once per month</td>
<td>Excavation face in cavern gallery is more than 40 m away from the instrument.</td>
<td>The differential deformation between the current and the previous reading is within monitoring accuracy.</td>
</tr>
</tbody>
</table>

5 Monitoring Results and Discussion

In such large projects, compilation and recording of monitoring data is an important aspect and cannot be left to simple practice of sheets and tables. Presently with the advent of technology, this filed of data management has also witnessed lot of upgradation and automation. Therefore in this project also, all monitoring data were managed through a dedicated geo monitoring software, which works as a tunnel information system for continuous monitoring by concerned person at site as well as design office. Deformation value as observed in optical targets was plotted in all three direction (two horizontal and vertical) with respect to time of observation. Similarly, extensometer reading of three anchor points was also plotted with respect to time of observation. All these results were compared regularly in reference with different limit value.

5.1 Deformation Limit

Monitoring values were used to evaluate the rock mass condition by comparing measurement results against expected soil/rock deformation. Different monitoring limit values were introduced to categorise measurements in terms of their severity. This limit was developed from the expected deformation in detailed design calculation and functioning and maintenance of the structure. In the detailed design calculation, rock mass as observed in this area was classified broadly in three classes as per their Q values i.e. Class 1. Good and Very Good Rock, 2. Fair Rock and 3. Poor Rock and expected deformation were defined for each of them. For all these three classes two limits were fixed namely “trigger level” and “allowable level”. Trigger level is defined as the permitted maximum displacements of a particular structure or area, which is critical to the safety, functioning of the structure. It was considered as 80% of design value.

Table 4: Monitoring limit values for different rock class

<table>
<thead>
<tr>
<th>Rock Class (Q-value)</th>
<th>Top Heading</th>
<th>Benching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trigger (mm)</td>
<td>Allowable (mm)</td>
</tr>
<tr>
<td>Very Good and Good</td>
<td>5.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Fair</td>
<td>5.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Poor</td>
<td>24.0</td>
<td>42.0</td>
</tr>
</tbody>
</table>

Whereas allowable level, is the maximum limit which should not exceed in any case and this was set at 140% of the design value. Absolute values as considered in this project for different rock class at various stages of excavation are reported in Table 4.
5.2 Interpretation

In most of the case, the observed deformation in geotechnical monitoring data observed from optical targets was well within the trigger limit. There was not a single incident reported where the deformation exceed the allowable limit. However, at few locations sudden increase in deformation was observed due to the major excavation activity in that vicinity. When the deformation value exceed the trigger value at some section of the cavern, that location was visually inspected by engineer to check the possible distress in terms of increased seepage or any crack in the installed shotcrete to allow further excavation.

![Fig. 3: Deformation plot for vertical and horizontal direction upto completion of project](image)

Typical deformation of cavern section plotted in Fig. 3 for poor rock condition where relatively higher deformation was observed.

![Fig. 4: Extensometer reading at poor rock condition](image)

Similar trend of deformation was also observed in Extensometer reading as presented in Fig. 4. In this case increase in the deformation was observed during the pilot tunnel and side drift at heading excavation. Afterward there was no significant increase in deformation due to subsequent bench excavation. In both the measurement, the deformation values in poor rock condition were remarkably low compare to the trigger value at different stage of excavation. This may be due to the conservative design approach as followed in this project for poor rock class. However, the incident of poor rock class encountered was less than 5% as compare to total excavation length.

6 Observation and Comments

This paper was specifically developed to highlight the systematic approach followed for large cavern excavation to minimize the risk of rock fall due to sudden change in geology. Initially, three stage geological investigations were quite efficient to understand the rock mass character more accurately. Particularly it helped to minimize the occasion of geological surprises involved in underground excavation. Moreover, continuous design updation activity employing high end dedicated software’s based on latest geological information along with geotechnical monitoring helps to use rock support system more efficiently and economically. However, geotechnical monitoring method using optical targets need to be improved for such large excavation. It was observed in many occasion during the excavation that the total station survey to take the reading suffer from accuracy due to disturbance from construction activity. The initial deformation which may have occurred immediately after excavation can’t be recorded in optical targets measurement. This need to be addressed in future study by cost effective automatic instrumentation. Thus monitoring data will be more accurate and will definitely minimize the scope of human error involved in recording and uploading such huge data. Also the application of mobile technology can help to keep alert all concerned round the clock.

References

[1] Design Package for Cavern Storage 4923/REP/PUA/0310, Engineers India Ltd., New Delhi, India.
Decision Support Model with Life Cycle Assessment for Building in Design Phase

Chia-Chen Wei a, Chia-Chi Hsiang b*, and Tzu-Chi Shan a

aCivil Engineering Department, Chien Hsin University of Science and Technology, Taiwan
bDepartment of Civil Engineering, Chung Yuan Christian University, Taiwan
E-mail: arnold@uch.edu.tw, chiachi@cycu.edu.tw, kiki800920@gmail.com

Abstract -
The building materials and structural form chosen in design phase not only affect the costs for construction, but also decide the costs for maintenance. However, the large amounts of maintenance costs are not easy to estimate. This article proposes a decision support model adopting life cycle assessment to assist the project owner and architect engineering for selecting an appropriate building structure type and materials is used in design phase via assessment the building’s total costs and CO2 emissions in the periods of build, maintenance, and disposal. The test results of cases show the model is feasible to help owners to decide a properly structural form.

Keywords -
Life Cycle Assessment; Building Information Modelling; Decision Support System

1 Introduction
The costs of a building in design phase mostly are evaluated as only one factor - construction cost. However, the building materials and structural form chosen in design phase not only affect the costs for construction, but also decide the costs for maintenance. Under the trend of buildings management, the researchers pay much attention to life cycle assessment (LCA) of buildings. LCA is a process of evaluating the economic performance of a building over its entire life. [1, 2]

From plan, design, construction, operation, maintenance, demolish, we think about the total costs in different stage. Based upon those considerations, we can step by step achieve the goals - energy saving, litter of building reduce, and environment sustainability. And this study proposes a decision support model adopting life cycle cost assessment concept to assist the project owner and architecture for selecting an appropriate building structure types and materials used in design phase.

2 Literature Review

2.1 Life cycle assessment
Life cycle assessment (LCA) is a technique for assessing various aspects associated with development of a product and its potential impact throughout a product’s life from raw material acquisition, processing, manufacturing, use and finally its disposal. [3] It is used to assess systematically the impact of each material and process.

The LCA process has three major phases: production phase, use phase, and the end of life phase. Each of them includes production, transportation, and distribution. The first studies on environmental impacts date from the 1960s and 1970s, focusing on the evaluation or comparison of consumer goods, with only a small contribution to the use phase [4].

In the beginning of the 1980s, the concept of life cycle adopt to study in the construction sector with focusing on the use of (renewable) resources [5]. The LCA concept is extensive applied by researchers in different aspect of construction sector in recent years, such as structure [6], energy consumption and environmental impacts [7, 8, 9, 10], building materials and components [11, 12].

2.2 Building information modelling (BIM)
BIM is currently the most common denomination for a new way of approaching the design, construction and maintenance of buildings. BIM is a digital representation of physical and functional characteristics of a facility [13]. A BIM is a shared knowledge resource for information about a facility forming a reliable basis for decisions during its life-cycle; defined as existing from earliest conception to demolition. Use of BIM goes beyond the planning and design phase of the project, extending throughout the building life cycle, supporting processes including cost management, construction management, project management and
facility operation.

The most benefits of applying BIM in design phase are cost reduction and control and time saving by improving productivity, better coordination and reduced error, and rework [14, 15].

3 Decision Support Model with LCA

This study proposes a model which was composed with modules in different phase of building lifecycle. The decision support model is shown in Figure 1.

![Figure 1. Main structure of decision support model](image)

3.1 Life cycle cost

In this study, the building’s life cycle cost in study year $y$ can be defined as below:

$$LCC = C_C \times (1+i)^y + \sum_{t=0}^{T} C_M \times (1+e)^{y_m} \times (1+i)^{y-e}$$

while:
- $C_C$ - construction cost in year 0.
- $C_M$ - maintenance cost in year 0.
- $y_m$ - years of performing maintenance incur after construction accomplish.
- $e$ - inflation rate.
- $i$ - interest rate.
- $\sum$ - denotes the summation of all the cost of maintenance.
- $T$ - numbers of maintenance incur.

And in the end of lifespan, the building’s life cycle cost can be defined as below:

$$LCC = C_C \times (1+i)^y + \sum_{t=0}^{T} C_M \times (1+e)^{y_m} \times (1+i)^{y-e} + (C_D - R) \times (1+e)^{y_e}$$

while:
- $C_D$ - disposal cost in year 0.
- $R$ - residual value in year 0.
- $y_e$ - years of lifespan.

3.2 CO₂ emission

The greenhouse gases (GHGs) produced by humans’ daily activities and emission to atmosphere is the main cause of global climatic change. The primary GHGs in the Earth’s atmosphere are CO₂, O₃, CH₄, N₂O, CFCs, PFCs, HFCs, HCFCs and SF₆.

In the studies of GHGs emissions of construction activities show that the emissions of CO₂ hold a major part of GHGs [11]. In this study, the CO₂ in buildings life cycle is the topics for discussion. Therefore, the CO₂ emissions in buildings life cycle is one criterion of the alternative decision in this study.

4 Case Study

This study takes 60 years as a building lifespan, and calculates the life cycle cost and estimates the CO₂ emissions in different study year. Life cycle cost is considered under the interest rate 0.017 and the inflation rate 0.0083.

4.1 Structure type

Due to different construction materials of reinforced concrete structure and steel structure, the structural properties, and the lifespan of buildings are different [16]. This study chooses the above-mentioned two types structure building for LCA calculation. For simplicity, the easy form is taken into account only including beam and column.

The building’s BIM model shown in Figure 2 is drawn by Autodesk Revit. The dimensions and quantity of two type structure’s element of buildings are export by the software and shown in Table 1 and Table 2.
4.2 Costs

4.2.1 Construction cost

The unit price of construction materials used in case study is evaluated from the local construction market and shown below:

<table>
<thead>
<tr>
<th>Material</th>
<th>Unit</th>
<th>Price (NTD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete (210kgf/cm²)</td>
<td>m³</td>
<td>2,300</td>
</tr>
<tr>
<td>Rebar</td>
<td>tonne</td>
<td>20,000</td>
</tr>
<tr>
<td>Formwork</td>
<td>m²</td>
<td>540</td>
</tr>
<tr>
<td>H-Steel</td>
<td>tonne</td>
<td>37,000</td>
</tr>
<tr>
<td>Antirust painting</td>
<td>m²</td>
<td>140</td>
</tr>
</tbody>
</table>

From Table 1, Table 2, and Table 3, the construction costs of different structure types building is calculated and shown in Table 6.

4.2.2 Maintenance cost

The reinforced concrete structure is formed by a combination of concrete and rebar. Due to the neutralization of concrete cover is main cause of corrosion of rebar and damage the structure. This study adopts the action of replace the concrete cover to maintaining the reinforced concrete structure.

Cover with antirust paint is generally used to prevent the steel rusting; this study adopts painting action to maintaining the steel structure. Table 4 shows the maintenance cost of two type structure.

4.2.3 Disposal cost

When a building reaches a predetermined lifespan, the building must be demolished. The following table presents the disposal cost of two different type structures in this study.

4.2.4 Residual value

After the demolition of buildings, the used rebar can be sold as scrap, the unit price adopted in this study is 6,000 NTD/tonne, the used H-steel may reuse so the unit price adopted in this study is 9,000 NTD/tonne.
### 4.3 CO₂ Emissions

From Table 7 and Table 8 shows the CO₂ emissions (including transfer) in building’s life cycle.

#### Table 7. Unit CO₂ emissions of construction material

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit</th>
<th>CO₂ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete (210kgf/cm²)</td>
<td>m³</td>
<td>253.68</td>
</tr>
<tr>
<td>Rebar</td>
<td>tonne</td>
<td>964.75</td>
</tr>
<tr>
<td>Formwork</td>
<td>m²</td>
<td>2.18</td>
</tr>
<tr>
<td>H-Steel</td>
<td>tonne</td>
<td>982.16</td>
</tr>
</tbody>
</table>

#### Table 8. Unit CO₂ emissions of maintenance

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit</th>
<th>CO₂ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:3 cement finish</td>
<td>m²</td>
<td>5.66</td>
</tr>
<tr>
<td>Coating cement paint</td>
<td>m²</td>
<td>0.27</td>
</tr>
<tr>
<td>Antirust paint</td>
<td>m²</td>
<td>0.27</td>
</tr>
</tbody>
</table>

### 4.4 Parameters in Revit components

For processing the LCA in design phase, the Revit components used in case study not only have the basic geometric and material parameters but also contain the aforementioned costs and CO₂ emissions parameters. Table 9 shows the parameters of CC1 column in case study.

#### Figure 3. Parameters of CC1 column

### 4.5 Perform LCA

This study takes 60 years for a building’s lifespan, and calculates total lifecycle cost from construction cost, maintenance cost, disposal cost. With compound interest form, we compute the final cost within period. The figure represents the relationship between time and cost.

#### 4.5.1 Scenario 1

Assumed the two structure type of buildings were not performed any maintenance activity in whole lifespan, the LCC value is shown as Figure 4. The total CO₂ emissions are 117,005 kg for reinforced concrete structure and 362,527 kg for steel structure.

#### Figure 4. LCC of case study - scenario 1

### 4.5.2 Scenario 2

Due to the neutralization phenomenon of the covering of concrete will damaged the rebar of reinforced concrete structure. So a new building in the same form will be constructed in the half of 60 years. The building of steel structure takes maintaining action every 10 years to keep the quality of the building and all services in a safe condition in 60 years. The LCC value is shown as Figure 5, and the total CO₂ emissions are 117,630 kg for reinforced concrete structure and 362,667 kg for steel structure.

#### Figure 5. LCC of case study - scenario 2
4.5.3 Scenario 3

The neutralization phenomenon of the covering of concrete has damaged the rebar of reinforced concrete structure. The owner of the building decides to construct a same form new building in the half of 60 years, and dispose the damaged building at same time. The building of steel structure takes maintaining action every 10 years to keep the quality of the building and all services in a safe condition in 60 years. The LCC value is shown as Figure 6, and the total CO₂ emissions are 234,010 kg for reinforced concrete structure and 362,667 kg for steel structure.

Figure 6. LCC of case study - scenario 3

5 Conclusions

Building maintains are seriously considered in design phase. This study proposes a decision support model with life cycle assessment to help the project owners to select an appropriate building structure type and materials is used in design phase.

This study only tests a model with simple structure including beams and columns. But those results are acceptable for project's owner, designers, and constructors. Based on the results of scenarios can conclude that it is not proper to select structure type only by construction cost.

In scenario 1, although the maintenance costs is not considered, the disposal costs and the residual value will have a significant impact on the LCA value. A good design can prevent to waste more resources if taking into consideration the demolitions and reuse of materials.

In scenario 2, the LCA value of lowest construction cost alternative may be not lowest under considering the maintenance costs, disposal costs and residual value.

In scenario 3 show that, maintenance activities is necessary, because in the same total lifespan condition, the LCA value of rebuilt will higher.

Since CO₂ emissions variation with structure form and maintenance, the decision makers should trade-off between total CO₂ emissions and LCC value to select a proper alternative.

References


A Study on Vertical Zoning Algorithm of Real-Time Construction Lift Control for High-Rise Building

Joonghwan Shin\textsuperscript{a} and Soonwook Kwon\textsuperscript{b} and Donghyun Kim\textsuperscript{c}

\textsuperscript{a}Department of Convergence Engineering for Future City, Sungkyunkwan University, Republic of Korea
\textsuperscript{b}School of Civil & Architectural Engineering, Sungkyunkwan University, Republic of Korea
\textsuperscript{c}Department of Convergence Engineering for Future City, Sungkyunkwan University, Republic of Korea

E-mail: kiwijoa@skku.edu, swkwon@skku.edu, kdh1945@skku.edu

Abstract -
For the past few years, numerous super high-rise buildings have been built around the world. Among them, movement management of materials and labors is closely related to productivity of super-high rise construction. The objective of this study is to develop an algorithm which can increase productivity of lift operation on twin-cage and multicage temporary lift for construction site. The algorithm is expected to enable efficient response to complex movement in a super high-rise construction site by applying principles of the vertical lifting operation. It is expected that the algorithm may reduce working hours and traffic line through operation optimization. The developed algorithm can optimize lift operation time by using lifting cycle-estimating method which is generated based on the fundamental concerns when lift scheduling is planned. Lifting cycle-estimation has become a vital part of an arithmetic computation based on lift selection algorithm which controls factors such as distance between each lift, among passengers, and distances among lifts according to moving direction. This and a follow-up study aim to develop unmanned smart construction hoist with a goal to improve productivity and safety of vertical lifting in super high-rise construction site.

Keywords -
Construction Hoist; Vertical Zoning; ELIS(Embedded Lift Information system)

1 Introduction
Numerous super high-rise buildings have been built around the world, and many more are planned to be built that are often over 100 stories high\cite{1}. Larger, higher buildings are subject to more restrictions in terms of construction planning and operation\cite{2}. Among them, movement management of materials and labors is closely related to productivity of super-high rise construction, and its importance grows as the buildings become higher.

Currently, at super high-rise construction sites, an experienced site supervisor and operator manage the hoist operation for movement of materials and labors\cite{3}. This lowers efficiency of vertical movement in operating construction hoist. The lift user’s queue time increases in higher buildings.

Several construction hoists are planned and built in constructing super high-rise buildings over 100 stories. Unlike elevators installed at the core of the structure, construction hoists are built outside the building, upon the mast, and they are hard to control in an integrated manner. Often construction hoists are operated redundantly, delaying construction schedule in a large project. Given these circumstances, this study conducts a simulation on construction hoist operation, explores how to improve movement of materials and labors in super high-rise construction by developing optimized operation algorithm. The simulation results are assessed based on the cycle time of daily unit work processes according to lifting cycle time calculation (Cho, 2010).

2 Research Trend
Sacks et al. developed an automated lifting equipment monitoring system (Sacks et al, 2005). Cho et al. conducted a study on construction hoist operation planning in terms of lifting height and loading (Cho, 2011). Further, Shin proposed optimal operation of temporary construction hoists in a super high-rise building based on simulation and genetic algorithm. Before them, most studies focused on the use of tower crane or mobile crane, and other studies on construction hoist tended to emphasize lifting planning rather than lifting operation.

As super high-rise construction becomes more popular in Korea, there is a growing need for a
3 Concept of research

3.1 Unmanned Smart Construction Hoist

This study aims to advance construction technology of super high-rise buildings, with a goal to optimize vertical lifting of materials and labors in erecting a super high-rise building. In this study, the control server was connected to construction hoist installed at each mast and to the Zigbee wireless network system. Figure 1 Based on the data received from each construction hoist, the study designed a system to give an operating order. Each construction hoist sends real-time data on lifting speed, direction and present load to the control server. The server collects the data, selects the optimum lift and sends an operating order. The focus is on proposing a model that controls many construction hoists at once and operates them, as a preliminary step to develop unmanned smart lift.

3.2 Twin or Multi-cage Hoist Operation System

The highest building in the world, Burj Khalifa, used 17 construction hoists during the construction. Each mast is put in different places, depending on the site conditions, and currently, work schedule is made to prioritize the order of materials and labors for integrated management. However, with predetermined schedule, it is hard to flexibly respond to the unexpected situations at the site. Thus, an integrated control system is needed to manage construction hoist operation.

![Figure 1 Concept of Unmanned Smart Construction Hoist](image1)

![Figure 2 Algorithm Flow of Unmanned Smart Construction Hoist](image2)

The algorithm flow basically depends on elevator distribution algorithm for labor movement Figure 2. The difference is that the control server analyzes condition of the construction hoists by receiving operation data from the Zigbee network. Data related lift operation includes Velocity, Current Position, Moving Direction, Weight, Electricity Power, Cable tension and Guide roller status. We set up data intervals 200ms for this server system.
4 Embedded Lift Information System

The control server examines algorithm to select the optimum construction hoist, requesting the following data:

1. Velocity of Lift car
2. Position of Lift car
3. Direction of Lift car
4. Real-time available transportation capacity

The system named ELIS (Embedded Lift Information System) stands on basis sensor device data. The sensor module consist of Double Sensor type Encoder, Limit Switch, Separating type Current Transformer, Proximity Sensor, Load Cell. Figure 3 shows text type lift operation database.

In selecting optimum construction hoist, a primary consideration is the present load capacity; if the capacity is already full, the hoist is instantly excluded. The lift direction and the direction to the call floor are considered in terms of minimizing the movement. Third, travel time is calculated for the remaining hoists to select the optimum construction hoist that minimizes queue time of materials and labors.

Figure 4 ELIS (Embedded Lift Information System) Display

Original purpose of ELIS database system designed for lift safety monitoring and sending the out-of-order signal. But this information data has more important value by using computational algorithm.

5 Optimal Algorithm for Construction Hoist Operation

The construction hoist information collected by the sensors is transmitted to the control server through the Zigbee network. The information provides a basis for the algorithm to select optimum construction hoist when the next call comes in. Figure 5 describes the selection flow, how it eliminates unsuitable construction hoists by lift direction, present location and load capacity.

In selecting optimum construction hoist, a primary consideration is the present load capacity; if the capacity is already full, the hoist is instantly excluded. The lift direction and the direction to the call floor are considered in terms of minimizing the movement. Third, travel time is calculated for the remaining hoists to select the optimum construction hoist that minimizes queue time of materials and labors.

Figure 5 Optimum Hoist Selection Evaluation
The pseudo code below shows an algorithm system that yields the order of lift car selection based on cycle time calculation. First round of filtering is done between step 1 and 9, considering the present load capacity and accordance of the direction. Second round of filtering is done between step 10 and 17 to select the hoist with minimum travel time.

After the first round of filtering, travel time is calculated for the remaining hoists; the calculation formula are different for hoists that are presently operating and hoists that are idle at the moment. It is because time for power supply, acceleration and reduction should be considered. For example, acceleration time needs not be considered for currently operating hoist; only deceleration time matters at time of arrival.

1. \( i \leftarrow \text{code number of each construction hoist} \)
2. create arrays \( \text{Hoist}[1 \ldots 4] \)
3. \( \text{Hoist}[i] = -1 \)
4. for \( i \leftarrow 1 \) to 4
5.   do if \( W_i \leq W_{Mi} \times 0.8 \)
6.     if \( n - n_i < 0, k_i = 0 \) or 2
7.     else if \( n - n_i > 0, k_i = 0 \) or 1
8.     then \( \text{Hoist}[i] = 0 \)
9.   return \( i \) that hoist arrays value is 0
10. for \( i \leftarrow 1 \) to 4
11. if \( \text{Hoist}[i] = 0 \)
12. if \( k_i \leftarrow 0 \)
13. do \( T_i = \left( \frac{(n - n_i - h)}{V_i} \right) + S_1 + S_2 \)
14. else \( k \leftarrow 1 \) or \( k \leftarrow 2 \)
15. do \( T_i = \left( \frac{(n - n_i - h)}{V_i} \right) \)
16. let \( z \) be the smallest value in \( T_i \)
17. return \( i \) that is \( T_i = z \)

Here,
- \( i \) : Code number of each construction hoist
- \( W_i \) : Weight of each construction hoist
- \( W_{Mi} \) : Maximum capacity of each construction hoist
- \( n \) : Call floor
- \( n_i \) : Position of each construction hoist
- \( h \) : Distance for construction hoist acceleration and reduction
- \( k_i \) : Direction of previous call of each construction hoist
- \( T_i \) : Lifting time at operation speed
- \( S_1 \) : acceleration time
- \( S_2 \) : reduction time

6 Simulation

6.1 Simulation Method

The following hypotheses were adopted to verify workability of the proposed twin or multi-cage operation algorithm.

First, except for the ground floor, all floors have a same floor height.

Second, the building has two masts and four construction hoists.

Third, when a floor call comes in and a selected construction hoist begins to move, other hoists do not respond.

Fourth, rated velocity, acceleration and reduction time are pre-set for each type of construction hoist. More specifically, velocity of construction hoist A and D is 100m/min; their acceleration and reduction capacity are 0.60m/sec² and 0.57m/sec² respectively. For construction hoist B and C, the velocity is 70m/min, and their acceleration and reduction capacity are 0.55m/sec² and 0.53m/sec² respectively. In other words, operation speed and acceleration speed are given as constant, not variables, in an algorithm for optimum construction hoist selection. Fifth, considering that a construction hoist moves materials and labors at the same time, algorithm process identifies construction hoists that are operating over 80% of the load capacity.

Sixth, since this study deals with lifting operation rather than lifting planning, the optimum travel route will be to assign a construction hoist for materials and labors in the shortest time.

An analysis was made on the lifting operating simulation to examine reliability of the optimum construction hoist selection algorithm according to the six aforementioned conditions.

The study conducted simulation to measure lifting cycle time of the materials, and the result was compared to the manual construction hoist operation. Currently, an operator judges the floor calls, and sometimes multiple hoists are operated redundantly, lacking information on their status. This happens because an operator merely responds to floor calls without considering the overall management of the lifting operation.

The proposed algorithm for selection of optimum unmanned smart hoist uses a formula to decide travel route that reduces total cycle time, and it can solve the problem of redundant operation. The lifting model was based on the lifting cycle time of four construction hoists; assuming that the hoists are in operation, current position, direction, velocity, weight and call order were randomly assigned to each hoist.

Then, floor calls were generated at random floors at a regular interval to calculate lifting cycle time for both the current system and the proposed algorithm for optimum construction hoist selection.
Table 1 Simulation Conditions

<table>
<thead>
<tr>
<th>Loading/ Unloading Time</th>
<th>Door Open/Close Time</th>
<th>Total floor</th>
<th>Number of Hoist</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 min</td>
<td>0.05 min</td>
<td>30</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hoist</th>
<th>Model</th>
<th>Maximum Capacity</th>
<th>Weight</th>
<th>Velocity</th>
<th>Direction</th>
<th>Floor</th>
<th>Call Floor</th>
<th>Target Floor</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>High</td>
<td>3000kg</td>
<td>600kg</td>
<td>100 m/min</td>
<td>↑</td>
<td>9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>Medium</td>
<td>2000kg</td>
<td>1000kg</td>
<td>70 m/min</td>
<td>↓</td>
<td>27</td>
<td>15</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>Medium</td>
<td>2000kg</td>
<td>800kg</td>
<td>70 m/min</td>
<td>↓</td>
<td>23</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>D</td>
<td>High</td>
<td>3000kg</td>
<td>0kg</td>
<td>0 m/min</td>
<td>↓</td>
<td>5</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Maximum Capacity: 3000kg, Weight: 600kg, Velocity: 100 m/min.
Direction: ↑ for up, ↓ for down.
Floor: 9, 27, 23, 5.
Call Floor: 15, 21, 9, -.
Target Floor: 1, 1, 25, -.

Lifting Priority:
- Call Floor: 13, 7, 17, 29
- Floor Call Time: 0:00, 0:30, 1:00, 1:30
- Target Floor: 1, 1, 18, 13

Cycle time was calculated based on the cycle time calculation formula (Cho et al, 2010)[7].

6.2 Simulation Result and Analysis

As a result of simulation via simulation model table 1, cycle time data of two masts is table 2 and table 3. When 4 hoists in 2 masts are operated under twin or multi-cage algorithm, in following table 2 and 3, every possible duplicated operation can be eliminated. Black marked part means selection of hoist at floor call is occurred.

When 4th floor calls are occurred through simulation condition of table 1. The time spending for total lifting is 154.25sec considering material loading time and door-open time.

Table 2 Operation Cycle Time <Mast 1>

<table>
<thead>
<tr>
<th>EVENT</th>
<th>FLOOR CALL A</th>
<th>FLOOR CALL B</th>
<th>FLOOR CALL C</th>
<th>FLOOR CALL D</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td>13</td>
<td>9</td>
<td>-</td>
<td>15</td>
</tr>
<tr>
<td>↓</td>
<td>27</td>
<td>3</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>↓</td>
<td>11</td>
<td>11</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>↓</td>
<td>11</td>
<td>11</td>
<td>2</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3 Operation Cycle Time <Mast 2>

<table>
<thead>
<tr>
<th>EVENT</th>
<th>FLOOR CALL A</th>
<th>FLOOR CALL B</th>
<th>FLOOR CALL C</th>
<th>FLOOR CALL D</th>
</tr>
</thead>
<tbody>
<tr>
<td>↑</td>
<td>23</td>
<td>12</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>↓</td>
<td>16</td>
<td>12</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>↑</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>↓</td>
<td>1</td>
<td>7</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4 Simulation Conditions

<table>
<thead>
<tr>
<th>Loading/ Unloading Time</th>
<th>Door Open/Close Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 min</td>
<td>0.05 min</td>
</tr>
</tbody>
</table>

7 Conclusion and Further study

In this study, we proposed a unmanned smart lifting system and devised an optimum construction hoist selection algorithm on twin or multi-cage. And estimate the productivity of proposed system through simulation method. Proposed algorithm is considering hoist velocity, direction, position and weight capacity. Using optimum construction hoist selection algorithm process, eliminate all the duplicated call operation and minimize the queue time of materials and labor.

But, the detailed mechanic design of this system should be subjoined. Through searching the limitation factors of wireless network in construction site, we can consider optimization of information transmission between each construction hoist and central control server.
Acknowledgement

This research was supported by a grant (Code#’09 R&D A01) from the Cutting-edge Urban Development Program funded by the Ministry of Land, Infrastructure, Transport affairs of the Korean government.
This work is supported by Korea Minister of Ministry of Land, Infrastructure, Transport affairs as Convergence Engineering of Future City Master and Doctoral Grant Program

References


Introduction of Human-Robot Cooperation Technology at Construction Sites

Seungyeol Lee and Jeon Il Moon

Robotics System Research Division, Daegu Gyeongbuk Institute of Science & Technology, Republic of Korea
E-mail: syl@dgist.ac.kr, jimoon@dgist.ac.kr

Abstract -
Currently, a new construction method using a robotic system is widely spreading in construction sites. This study is related to introduction of human-robot cooperation technology which can improve convenience and productivity through the efficient interaction between a worker and a robotic system while doing glass panel installation works. Based on the analysis on glass panel installation with a glazing robot, functional requirements and approaches to address these requirements that can implement human-robot cooperative manipulation at construction sites. A practical example, which is applied to a specific target construction site, is also described in this paper. After field test at a real construction site, productivity and safety of the proposed system are compared with the existing glass panel installation system.

Keywords -
Human-robot cooperation, Construction robot, Glazing robot, Glass panel, Curtain-wall

1 Introduction

Since the late 1980s, construction robots have helped operators perform hazardous, tedious, and health-endangering tasks in heavy material handling [1~10]. Iwamoto et al. stated a similar problem that reduces the need for a labor force and provides improved productivity and safety [11]. Isao et al. discussed the appropriateness of the automation technology for installation of a curtain wall [12]. Masatoshi et al. proposed the automated building interior finishing system, and a suitable structural work method is described [13].

Generally, almost half of construction work is said to be building materials handling. Building materials and components are much larger and heavier than many other industrial materials. Buildings are made of many kinds of materials and each material may be a different shape. Glass panel is one type of building materials for interior/exterior finishing. The demand for larger glass panels has been increasing along with the number of high-rise buildings and the increased interest in building design. A glass panel has been designed to pursue beautifulness and satisfy the requirements from customers. Nevertheless, the size and weight of glass panel have to be limited in consideration of feasibility on entire process from transportation to installation. Because of the lack of suitable handling/installation equipment for glass panels, the construction process is always complicated and hazardous, relying on a number of construction workers. As shown in figure 1, handling heavy construction materials (e.g. curtain-walls etc.) has been, for the most part, eliminated for outside work by cranes and other various lifting equipment. Such equipment, however, is not available for precise work. To address curtain-walls handling needs for precise work, especially, 'TRCI (a Teleoperated Robot for Curtain-wall Installation)' has been developed and applied to the real construction site as shown in figure 2 [14]. This system comprised of two types of a robotic manipulator. One is a hydraulic actuated manipulator to implement curtain-wall handling (e.g. lifting, moving etc.) motion, the other is an electric actuated manipulator to implement precise curtain-wall installing motion with teleoperation.

A robotic system can be classified into two groups: those that can carry out work and coexist with humans in atypical environments, and those that do repeated work according to a standard program such as part assembly or welding and coating in the automobile and electronic industries. Thus, manufacturing robots are stationary and the product moves along an assembly line. In contrast, construction projects require a stationary product, that is the building, and the robots change location. Moreover, in manufacturing, robotic repetition provides identical products, whereas, in construction, the product is custom-made and robots must be reprogrammed to operate in each given condition [15]. Consequently, construction robots are defined as field robots that execute orders while operating in a dynamic environment where structures, operators, and equipment are constantly changing. Therefore, a guidance or teleoperated (or
remote-controlled) system is the natural way to implement construction robot manipulators. However, during operation of a teleoperated construction robot, problems arise due to operators receiving limited working information; the contact force when it carries out press pits between materials, thus reducing the ability to respond to the constantly changing operational environments. A human-robot cooperative method [16–18], in which an operator can handle/install glass panels intuitively, is described as improvements in this paper.

![Figure 1. Glass curtain wall installation by cranes (or winches) and workers in high-rise building](image1)

![Figure 2. TRCI (a Teleoperated Robot for Curtain-wall Installation)](image2)

2 Functional Requirements and Approaches

After applying TRCI to a real construction site, we analyzed that the construction methods with/without TRCI had advantages and disadvantages in glass panels installing. The construction methods with TRCI is capable of motions needed high speed or power, whereas the methods without TRCI (i.e. construction method depends on workers) is sluggish, releasing only small amounts of energy, and commits errors frequently. On the other hand, the methods without TRCI is much more flexible and adaptable in thinking, motion, and behavior. Based on the upper analysis on construction methods with/without TRCI, we deduced functional requirements to be improved that integration of advantages of both construction methods with/without TRCI, and incorporation of them into the Human-Robot Cooperative (HRC) manipulation, would improve the efficiency or quality of glass panel installation as below. First, we considered a robotic system that can correctly follow a worker (or a robot operator)’s motion intention for glass panel installation at construction sites. Second, we also examined a robotic system that can share workspace with a construction worker in safety guaranteed. Third, we designed a coordination program for efficient cooperation between a worker and a robotic system in each unit work of glass panel installation. Finally, we worked out a detailed plan for a dexterous robot control that can reflect worker’s technological know-how.

Studies on the human-robot cooperation have been ceaselessly performed so far. In 1962’s the Cornell Aeronautical Laboratory researched master-slave system to amplify the human operator’s strength [19]. And then, further work was done by GE. GE designed Hardiman which is exoskeleton typed man-amplifier [20]. In 1980’s, kazerooni approached the innovative man-amplifier, extender, which is different from a master-slave system [21–23]. Power and operational signals are directly translated from human to robot. Kosuge presented a control algorithm for the human-robot cooperation using maneuverability and amplification factor [24].

To address upper functional requirements, we deduced an approach to human-robot cooperative manipulation in process of glass panel installation at construction sites. This approach is related to design of a robotic controller that can amplify force of operator with a certain force augmentation ratio so that operator can manage a heavy glass panel with relatively scaled-down force. And, to feel reaction forces helps intuitive operation by reflection of force from environments. Figure 3 describes a schematic of conceptual design based on the approach for introducing human-robot cooperation technology at glass panel installation sites. This concept design, especially, is considered interactions among human, robot and environment, and generation of target dynamics. To implement the human-robot cooperation in constrained condition (i.e. when a glass panel is under installing to a panel frame by the robot), the impedance control method, which was proposed by Hogan [25], is applied as a basic robot force control method.

The development methods for construction robots can be classified into two categories. The first category involves developing entirely new robots that can achieve requested work. The second category involves new robots implementing with existing similar construction equipment. The first method is beneficial in optimizing
specifically requested work. However, the cost and time required by this method are the major drawbacks to developing new robots. The second method is difficult to optimize for target projects, but it can achieve efficiency with limited cost and time requirements. In this study, the second method is introduced to implement the suggested robotic system.

3 A Practical Example

The existing glass panel installation process, which is complicated and hazardous, relies on scaffolding (or aerial lift) and construction workers. This process exposes workers to falling accidents or vehicle rollovers. In addition, inappropriate working posture is a major element that increases the frequency of accidents by causing various musculo-skeletal disorders and decreasing concentration. That is to say, it becomes a direct cause of decreasing productivity and safety in building construction.

Figure 4 shows the target construction site and glass panel installation position related to the first application of human-robot cooperation at construction sites. The building size is 32m × 22m and the installation position of glass panels is 7.9m above the ground. The glass panel used for installation can be classified into two categories. The first category is a glass panel that has 3000mm × 1500mm dimensions and is 120kg. The second category is a glass panel that has 1500mm × 1500mm dimensions and is 60kg. This paper introduces the ‘Module T&H-bar’ installation method, which represents the ‘Lay-in’ to place the glass panel on frames. According to analysis of the target work, it is deduced the functional requirements for implementing a glass panel installation system based on human-robot cooperation as below. First, this system must be able to lift heavy glass panels, a worker, and the installation equipment. It requires engines, batteries, or motors to lift them. Second, this system must be able to handle heavy and fragile glass panels. It requires sophisticated force and position control based on human-robot cooperative manipulation. Third, this system must be devised to help construction workers, not to replace them. This requires a smart human-robot interfacing device to interact with a robot operator and a robotic system. The system must share the workspace with a robot operator. Fifth, this system must be able to reflect the technical operator’s skills that are required to obtain homogeneous construction quality. Thus, the system must follow the operator’s motion intentions in various working processes and environments at unstructured construction sites. Last, this system must belong in the task planning. This is required to prevent worker’s accidents and help workers increase productivity, by reducing the recovery time from system malfunctions and decreasing the worker’s duty time in dangerous works.
selected model (KUKA Industrial Robots). In order to control the motion of the manipulator, kinematic and dynamic analysis is required. As operator’s safety is influenced by these types of motion, while any singularities in the hardware and software should be considered carefully. The human-robot interfacing device is involved with installing glass panels by cooperating a robot operator (In here, a robot operator is one of normal construction workers, not a robotic engineer.) with the basic system. This device plays a role in delivering the robot operator’s motion intentions to the basic system’s motion controller. It is positioned between the flange of the multi-DOF manipulator and the vacuum suction device, while it is composed of two types F/T sensors. If the robotic operator puts external force containing a motion intentions (i.e. operational commands) on a handle of the HRI, it is converted into a control signal to operate the manipulator from operational sensor (6 DOF force/torque sensor; ATI Industrial Automation, Inc.) and a manipulator’s motion controller. Here, if the manipulator comes in contact with an external object (e.g. glass panel’s frame, obstacles, etc.), information on the contact force is transmitted to the manipulator’s motion controller through environmental sensor (6 DOF force/torque sensor). It is important to note that external force transmitted through environmental sensor and that transmitted to operational sensor should operate separately from each other. End-effector types of a robotic manipulator varies according to the properties of the construction materials. Since this paper aims at installing construction materials with relatively smooth surfaces, such as glass panels, a vacuum suction pad is used as the end-effector. Lastly, an outrigger to prevent a robot from tumbling, additional safety devices for a robotic operator, and an alarm device to alert neighboring workers of robot operation are necessary, with consideration for the operational environments and characteristics of construction sites. Figure 6 shows the field test with the proposed robotic system at the target construction site.

4 Results of Field Test and Conclusions

Table 2 shows the results of the field test. Comparison with manual installation process is not executed because the glass panel is too heavy to handle by construction workers. However, to prove advances in handling heavy construction materials, the existing (manual) installation process [14] of curtain-wall construction is introduced. Working time means the whole time consumed in loading the glass panel from the ground and installing (including finishing) it in the panel frame. Labor intensity means the degree of manpower strength required of workers during the glass panel installation process. Convenience indicates the degree of difficulty of the installation work,
and safety shows derived degree of safety.

The resulting comparison and analysis in Table 2 can be changed according to the working environment of the target construction site. In the case of installing a glass panel on smaller buildings, the work may depend on manpower. But according to the tendency of current construction trends towards larger and more sophisticated buildings, we are looking forward to highlighting the glass panel installation method based on human-robot cooperation in the near future.

Table 2. 1 Results of field test

<table>
<thead>
<tr>
<th>Working time</th>
<th>Curtain-wall installation with manpower and winch [14]</th>
<th>Glass panel installation based on human-robot cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor intensity</td>
<td>18 min. High momentary labor intensity</td>
<td>Avg. 26 min/piece (including finishing) Generally low labor intensity</td>
</tr>
<tr>
<td>Convenience</td>
<td>Profoundly dangerous work under obstacle interference</td>
<td>Generally convenient work Reducing in danger; fewer accidents</td>
</tr>
<tr>
<td>Safety</td>
<td>Generally dangerous; scattered accidents</td>
<td>Reduction in danger; fewer accidents</td>
</tr>
<tr>
<td>Number of workers</td>
<td>3</td>
<td>2 (deck:1, aerial lift:1)</td>
</tr>
</tbody>
</table>

Acknowledgment

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT and Future Planning of Korea (14-RS-02).

References

[18] Miller, J.S. The Myotron – A Servo-Controlled


Study on the Optimal Digging Range for Intelligent Excavation

Jeonghwan Kim, Joonhyun Jang, Leeho Lee, and Jongwon Seo

Dept. of Civil Engineering, Hanyang University, Korea
Korea Construction Engineering Development Collaboratory Program, Korea
Dept. of Civil Engineering, Hanyang University, Korea
Dept. of Civil Environmental Engineering, Hanyang University, Korea (Corresponding Author)
E-mail: kimjh0418@gmail.com, jang1013@koced.net, leedon62@naver.com, jseo@hanyang.ac.kr,

Abstract - Excavators play an important role at construction sites owing to high application and economic advantages, but the number of skilled operators of the excavator decreased because young operators do not like to enter to construction business. Research on automation of excavator and earthwork equipment was actively done such as IES (Intelligent Excavating System) in Korea. Much of research relied upon skilled worker's heuristics at making decision of minimum working unit without considering the conditions of the ground and excavator load, and it exposed to fall down of excavator at collapse of the ground and difficult to improve working efficiency. This study suggested optimal digging range considering heuristics of skilled workers, ground of various kinds of working areas and excavator specification to supplement existing methods and to improve safety and performance.

Keywords - Automated Excavation; Geotechnical Conditions; Culmann’s method; Optimal Digging Range

1 Introduction

The earthwork at construction sites produces basic ground to build up facilities at natural land. Earthwork project occupy about 20% of construction project cost and it has great influence upon construction project and to rely upon construction machinery more than other civil engineering businesses do [1]. The productivity, quality and safety, etc. may vary depending upon use and operation of construction machinery [2].

Nowadays, on-the-spot construction work has become particularly difficult at large-scaled and complicated construction projects, and these include a variety of construction businesses. Consequently, risks in handling heavy weight machines and a much higher rate of negligence-related accidents have created a situation where many young labourers do not feel motivated to work at construction sites. As a result, the number of skilled operators of construction machinery is likely to decrease day by day: Nonetheless, the construction business has relied upon machinery and equipment. Recently, Studies on construction machines and equipment can prevent negligent accidents from happening in the absence of skilled operators, as well as prevent productivity and quality from being lowered. In other words, it creates a decision-making system of not only operators but also managers of construction machines to make operation plans of construction machines and to put them into practice using a full automation system. In light of this, the Ministry of Land, Infrastructure and Transportation supported R&D project of Intelligent Excavating System (IES) to improve conditions and environment at construction sites. The development of IES expanded the applications and use of excavation machines, providing an automation system of simple and repeated excavation.

The system adopted the digging range of excavators as basic operation unit of excavated area. Size of digging area or range can be varied based on the depth and length of excavation work. Increasing the excavation depth and length seem to good for excavation in terms of productivity, it may also cause an critical incident, such as overturn, by making an unstable ground condition due to excessive excavation. Therefore, we need to find a balance point between two things; maximizing the volume of excavation area and guaranteeing a safety work condition (e.g. soil collapse).

At the initial stage of IES development, digging range relied upon information of skilled excavator operator's heuristics. In other words, excavator operators' experience and knowledge were generalized to apply them into the system. However, that approach had always been applied in the same way, neglecting current ground conditions and weight of excavators. In fact, this implied exposing the machines to collapse or accidental falls due to possibilities of unexpected ground settlement, thus leading to lower levels of operation efficiency. In this study, the optimal digging range of IES that reflects ground condition and
excavator specifications (e.g. weight) at various kinds of operation areas was suggested to assure of safety and performance for intelligent excavation.

The goal of the IES was not to produce excavating robots with new types of specifications but to perform excavation effectively by using current excavators. That did not make any great changes on the existing specifications, its focus was to remodel some organs and add communications equipment, such as posture and positioning sensors, as well as laser scanners and cameras. 13-ton caterpillar type excavator (Model No. DX140, supplied by Doosan Infracore Co., Ltd.) were investigated in IES development, which is often used at local construction sites to develop prototype. In this study, modules for optimal digging range of each excavator were developed to expand use. This study suggested Vertical Excavation Depth ($H_v$), Safety Length ($L_s$) and Horizontal Excavation Length ($L_{opt}$) of digging range as shown in figure 1, minimum working unit of excavation, to let IES excavate effectively in several steps; First, the study investigated IES and task planning system development to find out problems of IES digging range generation as well as needs of new optimal digging range generation.

Second, as mentioned before, the study investigated specifications of excavators of Doosan Infracore to examine excavator movement depending upon operations and specifications that made change in accordance with environment of earthwork, and to discover parameters of modules of optimal digging range generation.

Third, the study investigated alternatives for the production of optimal digging range generation module to suggest optimal alternative and to give development method of modules and to set up optimal digging range generation algorithm.

Figure 1. Configuration of digging range

2 Optimal Digging Range

2.1 Definition of digging range

The digging range that an excavator excavates without rotation is the minimum task unit of the task planning system as figure 2. In other words, the size of the digging range has great influence upon efficiency and safety of IES task plan.

![Figure 2. Digging range](image)

Excavated amount of dirt and cycle time determines the productivity of excavation task. Wide digging range for quick operation can increase excavating task excessively, and deep digging depth without considering physical properties of the soil and safety (buffer) area around the excavator could destroy slopes to overturn excavator. Smaller size of digging range that stabilizes ground can lower excavation efficiency so that optimal digging range size is needed to assure of safety and efficiency of excavation.

The digging range cross section consists of $H_v$ (vertical excavation depth), $L_s$ (safety length) and $L_{opt}$ (horizontal excavation length in optimal) as shown in figure 1. $L_s$ includes distance from center of excavator to start point of the slope to apply bucket width of excavator specification to the digging range. And, $H_v$ is 2 to 2.5 m at earthwork site: In this study, heuristics of skilled excavation operator who did not excavate more than 2.5 meters was reflected.

The failure surface can be made by using critical break angle based on angle of internal friction. At this time, the study has inspected slope stability considering downward moving force of earth along the failure surface, resistance and weight of the excavator.

2.2 Slope Stability Analysis Theory

The methods of slope stability analysis for development of digging range generation algorithm of soil properties includes not only straight line but also curve of failure surface (Table 1).
The 31st International Symposium on Automation and Robotics in Construction and Mining (ISARC 2014)

<table>
<thead>
<tr>
<th>Table 1 Analysis of Slope Safety</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Failure Method</td>
</tr>
</tbody>
</table>

These days, the limit equilibrium method (Bishop’s section method) that is similar to failure surface has been often used, and curve of failure needs to reflect variables that can be produced at many complicated environments of calculation. Also, break without enough depth can be close to plane break, and Culmann method that is subject to plane of failure surface can produce satisfactory results [3]. That is, the digging range generation algorithm was designed based on Culmann method subject to plane of failure. The steep slope with $H_s$ of excavating depth has failure surface ($\overline{AB}$) with critical break angle ($\theta$) as shown in Fig. 3. Equation 1 is used to estimate weight of earth on the failure surface ($\Delta ABC$).

$$W_s = \frac{1}{2} (\overline{AC})(H_r)(\gamma) = \frac{1}{2} (H_r \tan \theta)(H_r)(\gamma)$$

(1)

Where, $W_s$: Soil weight on the failure surface $H_r$: Excavation depth (vertical) $\gamma$: Unit weight of soil

Weight of earth ($W_s$) of failure surface ($\overline{BC}$) has normal and tangent Equation (2):

$$N_s = W_s \cos \theta = \frac{1}{2} \gamma H^2 \sin \theta \cos \theta$$

(2)

$$T_s = W_s \sin \theta = \frac{1}{2} \gamma H^2 \sin \theta \sin \theta$$

(3)

Equation 3 makes change with tangent of earth weight ($W_s$) and excavator’s partial weight ($W_{EB}$):

$$N_s = (W_s + W_{EB}) \cos \theta = \left(\frac{1}{2} \gamma H^2 \sin \theta \cos \theta + W_{EB}\right) \cos \theta$$

(4)

$$T_s = (W_s + W_{EB}) \sin \theta = \left(\frac{1}{2} \gamma H^2 \sin \theta \sin \theta + W_{EB}\right) \sin \theta$$

(5)

Not only normal stress on average but also shear stress on average ($\tau_s$) is applied to failure surface ($\overline{BC}$):

$$\frac{\sigma}{\alpha} = \frac{N_s}{\overline{BC}} = \frac{N_s}{\overline{BC}} \left(\frac{\overline{AC}}{\overline{BC}}\right) = \left(\frac{A \tan \theta \cos \theta}{\alpha}\right) \left(\frac{1}{2} \gamma H^2 \sin \theta \cos \theta + W_{EB}\right)$$

(6)

$$\tau_s = \frac{\tau_s}{\alpha} = \frac{\tau_s}{\alpha} \left(\frac{\overline{AC}}{\overline{BC}}\right) = \left(\frac{A \tan \theta \cos \theta}{\alpha}\right) \left(\frac{1}{2} \gamma H^2 \sin \theta \cos \theta + W_{EB}\right)$$

(7)

The shear stress on average of failure surface ($\overline{BC}$) can be obtained by using normal stress on average Equation (6) Equation (8):

$$\tau_s = c + \sigma \tan \theta$$

$$\tau_s = \left(\frac{\overline{AC}}{\overline{BC}}\right) = \left(\frac{A \tan \theta \cos \theta}{\alpha}\right) \left(\frac{1}{2} \gamma H^2 \sin \theta \cos \theta + W_{EB}\right) \tan \theta$$

(8)

The study obtains not only shear stress on average ($\tau_s$) but also shear stress on average ($\tau_s$) on failure surface ($\overline{BC}$) by using partial load ($W_{EB}$) of failure load ($\overline{AC}$) of total excavator load on ground surface as well as weight ($W_s$) of the earth ($\Delta ABC$): The shear stress on average ($\tau_s$) can be compared with shear stress strength on average ($\tau_s$) to judge break of steep slope.

Safety factor shall be applied to shear stress on average ($\tau_s$) considering risks Equation (9). The regulation of road design of Korea Expressway Corporation [4] adopted FS of more than 1.5 at cutting of soil layer as well as weathered rock that did not consider underground water level at dry season. This study supplemented special situation that slope stability analysis did not consider not only continuous change of weight center of the excavator but also vibration at
movement of the excavator, and it adopted factor of safety of 1.6:

\[ F_s \tau_s < \tau_r \]  \hspace{1cm} (9)

\[ F_s = \frac{\sin \theta}{\eta_r} \left( \frac{1}{2} \gamma H_v^2 \alpha \tan \theta + W_{BK} \right) < \]

\[ c + \left( \frac{\sin \theta \tan \theta}{\eta_r} \right) \left( \frac{1}{2} \gamma H_v^2 \alpha \tan \theta + W_{BK} \right) \tan \psi \]  \hspace{1cm} (10)

So, shear stress strength on average that resists break more than 1.6 times than shear stress on average of break of failure surface (\(\overline{AC}\)) is thought to be safe.

2.3 Horizontal Excavation Length

\(L_{opt}\) (Horizontal Excavation Length) on which size of digging range relies shall be decided before making not only \(H_v\) but also \(L_s\).

![Figure 3. Horizontal digging distance (\(L_{opt}\))](image)

The excavator can often reach a distance of maximum arm length and boom, and skilled excavator operators excavated earth surface with 130° to 135° of boom and/or arm angle of the excavator (\(\theta\) of fig 18): \(L_{opt}\) that reflected the experience and knowledge of excavator operators based on Equation (11):

\[ L_{opt} = \sqrt{a^2 + (b + c)^2} - 2a(b + c) \cos \theta \]  \hspace{1cm} (11)

2.4 Vertical Excavation Depth and Safety Length

The initial process of discovery of both \(H_v\) (Vertical Excavation Depth) and \(L_s\) (Safety Length) starts from \(H_v\) and specification of the excavator. A slope stability analysis is to be done from 0 ton having no effect of \(W_{BK}\) to 13 ton of total weight of the excavator subject to 1 meter of \(H_v\) and 13 ton of excavator total weight to find out excavator \(W_{BK}\) at critical break point that the slope broke down. \(L_s\) can be estimated with 1 meter of \(H_v\) with \(L_t\) being total excavator weight versus track length (\(L_{opt}\)) Equation (12).

\[ L_s = H_v \cot \theta - K L_t + \frac{1}{2} L_t \]  \hspace{1cm} (12)

In essence, after the first process is done, the second process is carried out by increasing \(H_v\) gradually from 1 meter, and the process is repeated for every \(H_v\). Earthwork sites often have 2.5 meters or less of \(H_v\) considering drainage at excavation, dump loading and connections with other types of constructions which have up to 2.5 meters of \(H_v\). At both the first process and the second process, each \(H_v\) is to be obtained for \(L_s\) of \(W_{BK}\) at critical break point, and the area of cross section of digging range is used to decide upon both \(H_v\) and \(L_s\). The length of excavation can be obtained by deducting \(L_s\) from \(L_{opt}\) to estimate cross section of digging range by using \(H_v\) and length of the excavation. When excavation length exceeds three times of bucket length, excavation can be done smoothly. Optimum \(H_v\) and \(L_s\) have been set at three times or more of excavation length than bucket length as well as the largest area of cross section of digging range.

2.5 Optimal Digging Range Generation Algorithm

Figure 4 shows an optimal digging range generation using the algorithm mentioned above.

The process of determination of both \(H_v\) and \(L_s\) is the most important at optimal digging range algorithm.

First, standard of \(H_v\) is to be set with IES specifications as well as soil parameters. The example has 1 meter of \(H_v\) as well as 3 meters of \(L_s\) (length from excavator center to end of the excavation) as shown in figure 5. 14.4-ton excavator on firm and stable ground may collapse slope of more than 1-meter of excavation depth according to skilled excavator operators: Consequently, 1 meter \(H_v\) has been set. \(L_s\) has been set considering length of excavator boom and front side.

Second, slope stability analysis is done with reference condition of \(H_v\) and \(L_s\). In the example, reference condition is safe. Third, \(L_s\) decreases at the reference condition. In the example, \(L_s\) decreased by 0.5 meters. Basically, with a smaller \(L_s\), \(W_{BK}\) of the excavation can be added to weight of the earth on failure surface to increase break force.

Fourth, slope stability analysis is to be done at less \(L_s\). In the example, given \(L_s\) is safe even at 2.5 meter. And the same process has been repeated until break of the slope. In the example, the slope was broken down at 1.5 meter of \(L_s\). Fifth, aforementioned process is to be repeated with deep \(H_v\).
Figure 4. Optimal digging range generation algorithm
2.6 Algorithm Test before Module Production

We conducted the test by using spreadsheet of MS Office Excel before producing modules with optimal digging range generation algorithm. The spreadsheet for testing consisted of input of both excavator specifications and soil parameters, calculation of slope stability depending upon place of the excavator, and output showing optimal digging range to check results by applying soil quality that can be found out at common earthwork sites (Fig. 7).

![Figure 7. Algorithm implementation](image)

The Road Design Manual [4] gives soil parameters of natural specimen of the design to classify SM and SC into 1.7tf/m$^3$ of unit weight, 30 degree of angle of internal friction and 3tf/m$^3$ or less of adhesiveness. The author increased adhesiveness from 1tf/m$^3$ to 3tf/m$^3$ to estimate optimal digging range by using the spreadsheet for the test. Two types of excavator (Table 2) were used and results are shown in Figure 8 and 9.

![Table 2. Specifications of excavator](image)

3 Case study

In this study, IES with task planning system having digging range generation module was used to conduct test at earthwork site by many IES technicians and professionals who are experienced in earthwork. The purpose of the test was to inspect $H_v$, $L_{opt}$ and $L_s$ and to verify exactness and stability by testing sizes of digging range of excavation of IES: An earth-work site was chosen at Hanyang University's Erica campus at Ansan, where banking with 15 meters in width, 15 meters in length and 1.5 meter in height was made to demonstrate
IES in operation and to conduct tests after skilled operators excavated several times (Fig. 10).

Figure 8. Result of optimal digging range generation algorithm test of DX 140

Figure 9. Result of optimal digging range generation algorithm test of DX 220

Figure 10. Construction Sites for IES Test

The findings are as follows: $H_i$ had an error of about $\pm 6.45$cm, and $L_{opt}$ had an error of about $\pm 7.93$cm and $L_s$ had an error of about $\pm 13.23$cm. The information of digging range generation module differed a little from the results of actual excavation. Nonetheless, we admitted of the existence of errors because effects of posture and position sensor exactness as well as mechanical errors could not be removed completely.

On top of that, some professionals said that the size of the digging range had conservative values from the point of view of ground properties at earthwork site. The digging range generation module gives enough digging range information for IES to excavate safely without collapse of excavating slope, and the digging range size is smaller than when compared to operations conducted by skilled excavator operators.

Skilled excavator operators, however, mentioned that they could not take immediate and flexible actions against working environment when they took on an excavator, and that they needed to keep stability relying upon conservative values. Therefore, the excavation amount of the digging range generation model was a little smaller than that of skilled excavator operators which increase excavation efficiency considerably, with stability of the excavator. Also, not only digging range information of the modules but also errors of actual IES task results could be used to improve sensors and mechanical parts.

4 Conclusion

This study has improved conventional method that decided upon size of digging range, minimum task unit of Intelligent Excavation System (IES) that relied upon skilled operator’s heuristics being development factor of IES and it developed auto generation module of optimal digging range size by using soil strength properties at earthwork sites and IES specifications. The slope stability analysis was done repeatedly subject to plane of hypothetical break surface and load of the excavator to give optimal digging range for better IES performance and to improve safety and productivity in
simple and prompt way. Further studies were needed to protect excavators from risks that were not acknowledged prior to application to actual earthwork sites, and to inspect various kinds of ground environments, and to examine not only static factors but also dynamic factors.

References


Information Modelling on Mechanized Earthworks

Shigeomi Nishigaki\textsuperscript{a}, Katsutoshi Saibara\textsuperscript{b}, and Shigeo Kitahara\textsuperscript{c}

\textsuperscript{a}Mazaran, Co., Ltd., and Kick, Co., Ltd., Japan
\textsuperscript{b}Kick, Co., Ltd., Japan
\textsuperscript{c}Kumagaigumi, Co., Ltd., Japan

E-mail: sleepingbear@c2m.com, saibara@c2mp.com, skitahar@ku.kumagaigumi.co.jp

Abstract - Mechanized earthworks for post-disaster reconstruction largely forks into unmanned construction at hazardous working sites and manned one at less hazardous sites. In order to avoid secondary disasters, it is significant and required to compensate operator’s realistic sensation in order to enhance spatial awareness in the mechanized earthworks. The mechanized earthworks consist of a series of discrete works, which are sometimes independently and at other time interactively performed by different workers and managed by small and medium contractors. Generally, a variety of construction machines made by different manufactures are used there. On the other hand, diverse machine guidance systems come from different vendors. Most of them are self-contained and constitute an independent unit in and of itself. Here, are demanded inexpensive and easy-to-use software systems with sensors capable of easily mounting and demounting. The software systems of those should be not embedded in specific construction machines but available to multiple ones. To overcoming problems like those, our primary efforts target on fabricating information modelling system for mechanized earthworks by utilizing open source software. The information modelling here means a process involving acquisition of field data and feedback of relevant digital representations of physical and functional characteristics of works in progress. This paper presents framework of field data acquisition and exploratory spatio-temporal visualizations, enhancement of spatial awareness, enhancement of communication, and conclusions with a discussion of the contributions included in this paper. Finally, some open problems are also pointed out.

Keywords - Information modelling; Mechanized earthworks; Unmanned construction; Realistic sensations; Potential hazards; Exploratory spatio-temporal visualization; Enhancement of Communication

1 Introduction
Motivations and objectives in this study, and structure of this paper are reported in this Section.

1.1 Compensation of Operator’s Realistic Sensations
Mechanized earthworks for post-disaster reconstruction largely forks into unmanned construction at hazardous working sites and manned one at less hazardous ones. When facing such dangerous situations, in order to avoid any secondary disasters, we demand unmanned operations.

Handling joysticks to operate construction machines, either directly or remotely, is an inherently eye-hand coordination task. The eye-hand coordination means to control eye movements with hand movements, and as processing visual representation of the situational views to handle joysticks along with the use of proprioception of the hands, or vice versa.

In case of manned operation at less hazardous working site for loading, hauling and dumping (LHD) of earth and rocks, operators would handle their machines based on their own realistic sensation given on the seat of their pants. For examples, distance and clearance perception based on front views, possible motion space, and velocity, acceleration, jerk, jolting, rattling, rolling, pitching, and yawing of the machine body.

In case of unmanned operation at hazardous working sites for removal of earth and rocks, operators can't feel their realistic sensation, and have to remotely operate their machines in a narrow field of view from camera-monitor system at the control station. Problems here include:

1. Machine operability,
2. Difficulty in a task at hand,
3. Performances limited by bearings of operators in
1.3 Objectives

To overcoming the problems like these as described above, our primary efforts target on fabricating an information modelling system on mechanized earthworks by utilizing open source software based on GNU General Public License [3]. The information modelling here means a process involving acquisition of field data and feedback of relevant digital representations of physical and functional characteristics of works in progress.

1.4 Structure of this Paper

This paper is organized as follows. First, we introduce framework of field data acquisition and exploratory spatio-temporal visualizations in Section 2. Secondly, in Section 3, we report enhancement of spatial awareness, put concretely, feedback of quantitative indexes and exploratory spatio-temporal visualizations of potential hazards in unmanned construction. Thirdly, enhancement of communication in the mechanized earthworks is given in Section 4. In Section 5, we conclude with a discussion of the contributions included in this paper. Finally, some open problems are also pointed out.

2 Framework of Field Data Acquisition and Information Feedback

Discrete spatio-temporal events are occurred in space with time by construction work practices. Data pertaining to the events are automatically collected in real time by the inexpensive devices. These devices are capable of easily mounting and demounting, which consist of six-degree accelerometers, GPS receivers, active tags, compact PCs, smart phones, etc., and sometimes manually acquired by key-in as needed.

Figure 1 shows framework to acquire filed data and feedback relevant digital representations of physical and functional characteristics of works in progress.

1.1 Objectives

To overcoming the problems like these as described above, our primary efforts target on fabricating an information modelling system on mechanized earthworks by utilizing open source software based on GNU General Public License [3]. The information modelling here means a process involving acquisition of field data and feedback of relevant digital representations of physical and functional characteristics of works in progress.

1.2 Enhancement of Communication

Mechanized earthworks consist of a series of discrete works, e.g., digging, loading, hauling, dumping, grading, embanking, compacting, etc., which are sometimes independently and at other time interactively performed by different workers and managed by small and medium contractors. Generally, a variety of new and used construction machines made by different manufactures are utilized in the mechanized earthworks.

It is said that machine guidance systems will be able to improve construction accuracy and productivity [1], [2]. At present, diverse machine guidance systems come from different vendors. Most of them are self-contained and constitute an independent unit in and of itself. Considering use of different construction machines, it is difficult for workers to coordinate data among the diverse machine guidance systems, and to generate significant information based on the data as linking them to the other systems. In addition, the up-front cost of the machine guidance systems might be a huge burden and a ball and chain to the small and medium contractors in Japan. They would demand inexpensive and easy-to-use software systems with devices, which are capable of easily mounting and demounting on different construction machines made by various manufactures. The software systems of those should be not embedded in specific construction machines but available to multiple ones.

Therefore, it might be difficult for workers to grasp current conditions of the each work related to one another, and, not to mention, almost impossible for them to know or reason the present and future situations of the whole work process on a real time basis. Here, it is very important to enhance communication as to work process and operations between the people concerned.
values of attributes: qualitative changes and changes of ordinal or numeric characteristic.

Exploratory spatio-temporal visualization of the above changes would be done by the open source software the R, which is a language and environment for statistical computing and graphics [4], [6] as focusing on:

(1) When and where, then what,
(2) When and what, then where, and
(3) Where and what, then when,

Figure 2 shows an image of the exploratory spatio-temporal visualization in this study. Numerical and graphic indicators here would be derived by:

(1) Detecting trends of consecutive increasing or decreasing values and change points latent in time series of the data, and
(2) Providing visual representation of work environments based on-demand processing of requests.

Figure 2. Image of the exploratory spatio-temporal visualization in this study

3 Enhancement of Spatial Awareness

In this Section, we present feedback of quantitative indexes and exploratory spatio-temporal visualization that will be able to enhance operator’s spatial awareness.

3.1 Feedback of Quantitative Indexes

The spatial awareness is very important for operators at their control room to infer and understand the present and future surroundings of their own construction machines or heavy-duty dump trucks and to watch themselves in the work space. In addition, it is significant and required to compensate operator’s realistic sensation, when remotely operating construction machine at dangerous spots, for example, at the steep slope, at the ridge of steep cliff, etc.

Figure 3 shows the information flow to assist operators. Construction machines have on-board devices that are composed of tools as follows:

(1) Tri-axial accelerometer,
(2) Tri-angular velocity meter,
(3) GPS receiver,
(4) Smart Phone or tablet PC, and
(5) Communication module.

These devices would be inexpensive and capable of mounting and demounting on any construction machine for earthworks.

Besides, are used wireless LAN-based devices that consist of active tags and the detection instruments with about 100 m effective range [7]. The each dump truck carries an active tag. The detection instruments are put in position on the side of the entrance to the loading and the dumping spots, respectively. The detection instruments successively receive signals from the active tags with the dump trucks that come into the effective cover range. Then the detection instruments automatically send the Web server system via the Internet the time stamps and the IDs of the active tags.

Based on field data given by the above on-board devices, are periodically and automatically generated and feedback quantitative indexes regarding operators’ bearings [8], the productivity [9], [10], and the physical cues [11].

Examples of the quantitative indexes as to operators’ bearing include:

(1) Mean cross rate,
(2) Magnitude of impact loads,
(3) Operability of orientation control,
(4) Learning capability,
(5) Difficulty in a task at a hand, and
(6) Response capability

Examples of the quantity indexes related to productivity include:

![Diagram of Information Flow](image-url)
(1) Moving distance, 
(2) Number of repositions, 
(3) Actual operation rate, and 
(4) Productivities of the LHD such as number of loading and dumping, distance hauled, work volume hauled, cycle time, etc.

Physical cues pertaining to posture of machine body and hazard factors latent in work places are fed back in real time. Examples of the physical cues include: 
(1) Posture of machine body such as inclination, rolling, pitching, 
(2) Positions of bucket pin, 
(3) Advance alert as to skidding, bumping, and rolling over, and 
(4) Advance proximity warning

To let operators recognize danger efficiently when remotely operating their machines, the numeric and graphic indicators with respect to the quantitative indexes would be intuitively context-sensitive assists to compensate operators’ realistic sensation.

Even under direct viewing, operators and the people concerned could not accurate at making inferences about the relative position of objects, the productivity of work in progress, and operators’ bearings. For securing safety, reliable and effective remote control, it becomes very important to provide operators with the information described in this Section.

3.2 Exploratory Spatio-temporal Visualization of Potential Hazards

Since operators could not easily envisage any hazard latent in unmanned earthwork process without any background information in a narrow field of view from camera-monitor system at the control station. It would be helpful and significant for operators in the control station to see severe dangerous spots with marks, and to watch the advance proximity warning on their monitors, when approaching to them.

Besides, exploratory spatio-temporal visualization of terrain and hazards are given with information resources such as digital elevation model and digital maps. Examples of the exploratory spatio-temporal visualization include:
(1) Scatter plot, 
(2) Ground surface contour, 
(3) Heat map, 
(4) Slope and aspect, 
(5) Marginal plot of terrain, and 
(6) Hazard map

Figure 4 shows the ground surface contour of construction area ant the trajectory of the heavy-duty dump truck (blue points), which hauls no-slump concrete to construct a multilayer check dam against avalanche of earth and rocks in Unzen-Fugendake. The multilayer check dam as shown in Figure 5 has been and is being constructed in severe hillside of which the gradients and aspects are shown in Table 1, and the gradients is visualized in Figure 6.
Figure 7 displays hazards regarding crosswise skiddy spots (red bubbles) latent in the no slump concrete haulage roads (blue dots). Also, we could visualize the longitudinal skiddy spots and the bumpy spots. For want of space, the charts like those are left out in this paper. The hazards latent in haulage road are detected based on existence of trends such as consecutive increasing or decreasing values in time series of tri-axial acceleration responses with longitudinal and latitudinal coordinates.

<table>
<thead>
<tr>
<th>Slope</th>
<th>Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>4.289</td>
</tr>
<tr>
<td>1st Qu</td>
<td>6.034</td>
</tr>
<tr>
<td>Median</td>
<td>7.224</td>
</tr>
<tr>
<td>Mean</td>
<td>7.896</td>
</tr>
<tr>
<td>3rd Qu.</td>
<td>9.938</td>
</tr>
<tr>
<td>Max.</td>
<td>13.705</td>
</tr>
</tbody>
</table>

Table 1. Gradient and aspect (unit: degree)

Hazardous spots like those might emerge at different places as the earthworks are progressing with time. To ensure safe and reliable haulage works, it is very significant to display the hazardous spots in advance on the monitor before starting and during work operations.

4 Enhancement of Communications

We have been and are developing the information modelling system, which enables the people concerned to share information on the no-slump concrete haulage and the compaction works in progress and to enhance communications among them. Figure 8 shows system configuration to enhance communications among the people concerned. This system enables them to peruse the analysis results, summary statistics and graphics of the works in progress anytime anywhere. Figure 9 shows a confirmation scene of compaction works in progress by the vibration roller in the outdoors other than at the control room.

4.1 Visualizations of No-slump Concrete Haulage Works

The no-slump concrete is hauled by 10 ton dump trucks from the ready-mixed concrete batching plant to the spot where it is transferred to heavy-duty dump trucks. Subsequently, the heavy-duty dump trucks
being remotely controlled haul the no-slump concrete to the designated dumping spot. In addition, a bulldozer spreads the no-slump concrete, and subsequently a vibration roller compacts [7].

Figure 10 shows the line chart of cumulative numbers both of arrivals of ton dump trucks to and the departures from the concrete batching plant for loading no-slump concretes.

![Figure 10. Line chart of arrivals and departures of 10 ton dump trucks at concrete batching plant](image1)

Figure 11 shows the progress curve of the no-slump concrete haulage.

![Figure 11. Progress curve of the no-slump concrete haulage](image2)

Besides, Figure 11 includes a base line for a comparison objective, which connects points of 0%, 33%, 66%, 100% at the time line axis and the corresponding 0%, 25%, 75%, 100% at the axis of the concrete volume hauled with a green line.

Figure 12 shows a spatio-temporal visualization of movements of a heavy-duty dump truck being remotely controlled to haul the no-slump concrete toward the designated dumping spot. Then the heavy-duty dump truck runs at an average of 18 kilometres an hour.

![Figure 12. Movements of heavy-duty dump truck being remotely controlled](image3)

The people concerned can watch these Figures anytime, anywhere. It can be easily seen from these Figures that there are no queuing lengths at both of the ready-mixed concrete batching plant and the transhipment place to the heavy-duty dump tracks. In addition, there is nothing of accumulative delays in the movements of the heavy-duty dump trucks.

4.2 Visualizations of Compaction Works

Figure 13 shows an example of display of compaction works by the machine guidance system for the vibration roller at the control room. This machine guidance system is self-contained and constitute an independent unit in and of itself only for the operator at the control room. Therefore, the people except the operator can not watch the display somewhere other than at the control room.

In contrast, the information modelling system enables them to peruse analysis results, summary statics and exploratory spatio-temporal visualization of the compaction works in progress by the vibration roller anytime, anywhere.

Figure 14 shows the number of compactions in progress. The Information modelling system also provides data regarding work starting time, current time, work time elapsed, distance covered, and work finish.
time. Then the vibration roller walks at an average of 6.7 kilometres an hour.

![Operator to handle a vibration roller at the control room](image)

Figure 13. Display of compaction works by machine guidance system at control room

![Machine guidance display for compaction works](image)

Figure 14. Number of compactions in progress

Figure 15 shows the exploratory spatio-temporal visualization of the compaction works in progress. Furthermore, Figure 16 shows the progress curve of the compaction works.

The compaction works by the vibration roller are done intermittently. As reported in Section 4.1, the concrete haulage still reserves the capacity. So, it is possible to increase more the concrete placement areas in parallel. In contrast, earthworks volume per hour burden on a backhoe would be larger.

![Exploratory spatio-temporal visualization of compaction works in progress](image)

Figure 15. Exploratory spatio-temporal visualization of compaction works in progress

![Progress curve of compaction works](image)

Figure 16. Progress curve of compaction works

5 Conclusion and Some Open Problems

The information modelling presented in this paper means a process including:
(1) Automatically gathering field data,
(2) Feedback of physical cues as to machine behaviour,
(3) Showing exploratory spatio-temporal visualization of hazards latent in in-situ earthwork process, and
Enhancement of communication regarding work process and operations. And then the information modelling system enables workers to:

1. Pick a set of building block of information at the right level of abstraction,
2. Provide operators with the opportunities to reflect on their own bearings as facing their own works at hand,
3. Reduce likely stress in remote control operations of machines, and
4. Take timely and quickly correct actions based on detailed visibility of appearances and motions of the mobile entities in a whole unmanned construction process.

Taken together, the workers concerned could get in-depth insight into the work process while showing hints to improve safety and productivity with detailed reporting. Accordingly, it would be possible to secure traceability and accountability with the mechanized earthworks.

In future, remotely dexterous operations in close proximity to other would be required, when unmanned construction system would be deployed in various kinds of post-disaster recovery and reconstruction, for example, approaching to the target (e.g., inspection, repair, monitoring another one), positioning or repositioning dangerous materials, connecting an object with other, grasping and stacking sandbags or gabions, inserting nozzle into a fuel filler port for replenishment, and so on. For securing safety, reliable and effective operations, we will be not just focusing on the best way but the variety of better ways to provide operators with physical cues, which would enhance

1. Awakening spatial awareness,
2. Perception of distance and clearance,
3. Recognition of possible motion space,
4. Grasping posture of machine body, and
5. Improvement of operator’s skill.

In addition, we will do further research and development on exploratory spatio-temporal visualization of:

1. Earthworks by backhoe,
2. Terrain being changed with earthwork operations, and
3. Civil engineering structure under construction

References


[2] MALAGA Demonstration & Learning Center


AUTHOR INDEX
AUTHOR INDEX

A
Ahmadian Fard Fini, Alireza 466
Ahmed, Essmeil 278
Ahmed, Riham 910
Ahn, Changbum R. 854
Aho, Timo 667
Akbarnezhad, Ali 385, 466, 714, 722
Akinci, Burcu 746
Al-Hussein, Mohamed 254, 262, 270, 286, 410
Alipour, Mujtaba Hossein 706
Alkass, Sabah 278
Allen, Jeremy 531
Altaf, Mohammed Sadiq 254
Ansuini, Roberta 506
Aram, Shiva 434
Aria, Sepideh S. 854
Askari, Mohsen 72
Awwad, Rita 354
Azzi, Merched 7

B
Balaguer, Carlos 19, 837
Barkokebas, Beda 270
Bazik, Peter 918
Beceorik-Gerber, Burcin 549
Bernold, Leonhard 868
Bertram, Torsten 778
Bock, Thomas 11, 142, 624, 762, 768, 778, 795
Bosche, Frederic 813
Bouferguène, Ahmed 262, 410
Brennan, Jane 499, 523
Bulgakov, Alexey 142, 762, 768, 778
Buzalo, Natalia 142

C
Caldas, Carlos 654
Capone, Pietro 313
Carbonari, Alessandro 506, 829
Carneiro, Joao 549
Castro-Lacouture, Daniel 536, 576
Chan, Greg 393
Chan, Wen-Shuo 593
Chang, Ting-Yu 616
Charlesraj, V. Paul C. 936
Chen, Chi-Jim 568
Chen, Hongsu 337
Chen, Hung-Ming 616
Chen, Pei-Yin 302, 821
Chen, Po-Han 561
Chen, Yit-Jin 795
Cheng, Jack 450
Cheng, Min-Yuan 661, 689
Cheng, Yuan-Yuan 593
Chi, Hung-Lin 738
Chien, Hsiao-Che 302
Chiu, Chien-Kuo 661
Chiu, Yung-Fang 661
Cho, Hunhee 474
Cho, Jehyun 224
Cho, Yong 134
Cho, Yun-Gu 95
Choi, Jinwon 210
Chong, Heap-Yih 4
Chuo, Baeksuk 784
Costin, Aaron 646

D
de La Casa, Santiago 837
Martinez
Ding, Grace 499, 523
Dissanayake, Gamini 43
Doloi, Hemanta 610, 904
Duc, Edward 238
Duc, Hiep Nguyen 544, 584
Dunston, Phillip 126
<table>
<thead>
<tr>
<th>Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastman, Charles</td>
<td>434</td>
</tr>
<tr>
<td>Eggers, Kai</td>
<td>514</td>
</tr>
<tr>
<td>El Souki, Omar</td>
<td>354</td>
</tr>
<tr>
<td>Emelianov, Sergey</td>
<td>762</td>
</tr>
<tr>
<td>Eom, Sunghoon</td>
<td>790</td>
</tr>
<tr>
<td>Er, Michael</td>
<td>346</td>
</tr>
<tr>
<td>Ergan, Semiha</td>
<td>746</td>
</tr>
<tr>
<td>Evgenov, Alexey</td>
<td>778</td>
</tr>
<tr>
<td>Fang, Gu</td>
<td>870</td>
</tr>
<tr>
<td>Fava, Gabriele</td>
<td>829</td>
</tr>
<tr>
<td>Felkl, Jakub</td>
<td>646</td>
</tr>
<tr>
<td>Feng, Chen</td>
<td>163</td>
</tr>
<tr>
<td>Fiander-Mccann, Dawn</td>
<td>329</td>
</tr>
<tr>
<td>Florez, Laura</td>
<td>536</td>
</tr>
<tr>
<td>Forsythe, Perry</td>
<td>238, 480</td>
</tr>
<tr>
<td>Gandhi, Shalini</td>
<td>346</td>
</tr>
<tr>
<td>Gansonre, Abdoul-Aziz</td>
<td>377, 593</td>
</tr>
<tr>
<td>Gasparik, Jozef</td>
<td>918</td>
</tr>
<tr>
<td>Georgoulas, Christos</td>
<td>624, 795</td>
</tr>
<tr>
<td>Georgy, Maged</td>
<td>398, 404</td>
</tr>
<tr>
<td>Gerber, David</td>
<td>549</td>
</tr>
<tr>
<td>Getuli, Vito</td>
<td>313</td>
</tr>
<tr>
<td>Gil, Myeongsu</td>
<td>929</td>
</tr>
<tr>
<td>Giretti, Alberto</td>
<td>506</td>
</tr>
<tr>
<td>Giusti, Tommaso</td>
<td>313</td>
</tr>
<tr>
<td>Golovina, Olga</td>
<td>646</td>
</tr>
<tr>
<td>Gucunski, Nenad</td>
<td>50</td>
</tr>
<tr>
<td>Guenet, Emeline</td>
<td>813</td>
</tr>
<tr>
<td>Güttler, Jorg</td>
<td>795</td>
</tr>
<tr>
<td>Han, Sanghyeok</td>
<td>262, 410</td>
</tr>
<tr>
<td>Hansen, Christian</td>
<td>514</td>
</tr>
<tr>
<td>Hanski, Sami</td>
<td>88</td>
</tr>
<tr>
<td>Heikkilä, Rauno</td>
<td>171, 667, 672, 697, 948, 953</td>
</tr>
<tr>
<td>Heilmann, Andrea</td>
<td>523</td>
</tr>
<tr>
<td>Hermann, Ulrich</td>
<td>262, 410</td>
</tr>
<tr>
<td>Heydarian, Arsalan</td>
<td>549</td>
</tr>
<tr>
<td>Hipel, Keith</td>
<td>329</td>
</tr>
<tr>
<td>Hong, Chang Hee</td>
<td>185</td>
</tr>
<tr>
<td>Hong, Daehie</td>
<td>157</td>
</tr>
<tr>
<td>Hong, Guang</td>
<td>584</td>
</tr>
<tr>
<td>Hossein Rashidi, Taha</td>
<td>466</td>
</tr>
<tr>
<td>Hou, Lei</td>
<td>295, 321</td>
</tr>
<tr>
<td>Hsiang, Chia-Chi</td>
<td>966</td>
</tr>
<tr>
<td>Hsieh, Shang-Hsien</td>
<td>738</td>
</tr>
<tr>
<td>Hsieh, Yo-Ming</td>
<td>860</td>
</tr>
<tr>
<td>Huang, Jingying</td>
<td>925</td>
</tr>
<tr>
<td>Huang, Ting</td>
<td>393</td>
</tr>
<tr>
<td>Huh, Jaemyung</td>
<td>157</td>
</tr>
<tr>
<td>Hwang, Nahyae</td>
<td>194</td>
</tr>
<tr>
<td>Ibrahim, Magdy</td>
<td>805</td>
</tr>
<tr>
<td>Ibrahim, Moheeb</td>
<td>398, 404</td>
</tr>
<tr>
<td>Ismail, R. M. T. Raja</td>
<td>149</td>
</tr>
<tr>
<td>Jabbour, Melanie</td>
<td>354</td>
</tr>
<tr>
<td>Jadidi, Hossein</td>
<td>706</td>
</tr>
<tr>
<td>Jang, Joohyun</td>
<td>984</td>
</tr>
<tr>
<td>Jardón, Alberto</td>
<td>19, 837</td>
</tr>
<tr>
<td>Jeon, Gyueob</td>
<td>217</td>
</tr>
<tr>
<td>Jeong, Hoyoung</td>
<td>942</td>
</tr>
<tr>
<td>Jiang, Jung-Yi</td>
<td>302</td>
</tr>
<tr>
<td>Kaaranka, Annemari</td>
<td>171, 602, 672, 697, 953</td>
</tr>
<tr>
<td>Kamat, Vineet</td>
<td>163</td>
</tr>
<tr>
<td>Kan, Meng-Shen</td>
<td>561</td>
</tr>
<tr>
<td>Kang, Julian</td>
<td>676</td>
</tr>
<tr>
<td>Kang, Kyung-In</td>
<td>474</td>
</tr>
</tbody>
</table>

1002
<table>
<thead>
<tr>
<th>Name</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGee, Wesley</td>
<td>163</td>
</tr>
<tr>
<td>Metia, Santanu</td>
<td>584</td>
</tr>
<tr>
<td>Mine, Naoto</td>
<td>426</td>
</tr>
<tr>
<td>Miro, Jaime Valls</td>
<td>43</td>
</tr>
<tr>
<td>Moghadam, Mana</td>
<td>270</td>
</tr>
<tr>
<td>Monter, R.</td>
<td>19</td>
</tr>
<tr>
<td>Moon, Jeon II</td>
<td>790, 978</td>
</tr>
<tr>
<td>Moon, Sungkon</td>
<td>868</td>
</tr>
<tr>
<td>Moon, Sung-Min</td>
<td>157</td>
</tr>
<tr>
<td>Moselhi, Osama</td>
<td>805</td>
</tr>
<tr>
<td>Moussavi Nadoushani, Zahra</td>
<td>385, 714</td>
</tr>
<tr>
<td>Sadat</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Na, Ok-Pin</td>
<td>95</td>
</tr>
<tr>
<td>Nahangi, Mohammad</td>
<td>847</td>
</tr>
<tr>
<td>Nanda, A.</td>
<td>960</td>
</tr>
<tr>
<td>Naticchia, Berardo</td>
<td>829</td>
</tr>
<tr>
<td>Negnevitsky, Michael</td>
<td>58</td>
</tr>
<tr>
<td>Nevala, Kalervo</td>
<td>697</td>
</tr>
<tr>
<td>Nguyen, Luan</td>
<td>50</td>
</tr>
<tr>
<td>Nishigaki, Shigeomi</td>
<td>458, 992</td>
</tr>
<tr>
<td>Nissilä, Jaakko</td>
<td>667</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>O</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Oduro, Seth Daniel</td>
<td>584</td>
</tr>
<tr>
<td>Olanrewaju, Abdulateef</td>
<td>426</td>
</tr>
<tr>
<td>Ooka, Ryozo</td>
<td>247</td>
</tr>
<tr>
<td>Orr, Kirsten</td>
<td>238, 346</td>
</tr>
<tr>
<td>Ortmaier, Tobias</td>
<td>514</td>
</tr>
<tr>
<td>Ou, Yu-Chen</td>
<td>661</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Park, Sung Jun</td>
<td>230</td>
</tr>
<tr>
<td>Park, Sungjae</td>
<td>309</td>
</tr>
<tr>
<td>Paukkeric, H.</td>
<td>948</td>
</tr>
<tr>
<td>Peltomaa, Irina</td>
<td>171, 602</td>
</tr>
<tr>
<td>Periaraj, Vivek</td>
<td>680</td>
</tr>
<tr>
<td>Pirinen, Tuomo</td>
<td>88</td>
</tr>
<tr>
<td>Pursimo, Juha</td>
<td>88</td>
</tr>
<tr>
<td>Q</td>
<td></td>
</tr>
<tr>
<td>Qin, Datong</td>
<td>925</td>
</tr>
<tr>
<td>Quan, Steven Jige</td>
<td>576</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rahnamayiezekavat, Payam</td>
<td>868</td>
</tr>
<tr>
<td>Raphael, Benny</td>
<td>81</td>
</tr>
<tr>
<td>Ravanshadnia, Mehdi</td>
<td>706</td>
</tr>
<tr>
<td>Raza, Ahmad</td>
<td>795</td>
</tr>
<tr>
<td>Rey, David</td>
<td>722</td>
</tr>
<tr>
<td>Romo, Ilkka</td>
<td>667</td>
</tr>
<tr>
<td>Rotimi, James</td>
<td>443</td>
</tr>
<tr>
<td>Ryu, Eunseo</td>
<td>224</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Safa, Mahdi</td>
<td>329, 531</td>
</tr>
<tr>
<td>Safa, Majeed</td>
<td>329, 531</td>
</tr>
<tr>
<td>Saibara, Katsutoshi</td>
<td>458, 992</td>
</tr>
<tr>
<td>Salter, David</td>
<td>544</td>
</tr>
<tr>
<td>Sammut, Claude</td>
<td>680</td>
</tr>
<tr>
<td>Sankaran, Shankar</td>
<td>346</td>
</tr>
<tr>
<td>Sattineni, Anoop</td>
<td>361</td>
</tr>
<tr>
<td>Sayfeddine, Daher</td>
<td>762</td>
</tr>
<tr>
<td>Seo, Jongwon</td>
<td>984</td>
</tr>
<tr>
<td>Sepasgozar, Samad M.E.</td>
<td>876</td>
</tr>
<tr>
<td>Shahi, Arash</td>
<td>329</td>
</tr>
<tr>
<td>Shan, Tzu-Chi</td>
<td>966</td>
</tr>
<tr>
<td>Shen, Xuesong</td>
<td>103</td>
</tr>
<tr>
<td>Shen, Yi-Jiun</td>
<td>568</td>
</tr>
<tr>
<td>Shiau, Yeu-Horng</td>
<td>821</td>
</tr>
<tr>
<td>Shih, Shih-Yu</td>
<td>593</td>
</tr>
<tr>
<td>Shin, Chang-Yeob</td>
<td>157</td>
</tr>
<tr>
<td>Shin, Dongbin</td>
<td>790</td>
</tr>
<tr>
<td>Shin, Joong-Hwan</td>
<td>972</td>
</tr>
<tr>
<td>Shin, Minkyu</td>
<td>202, 210, 217</td>
</tr>
<tr>
<td>Shirowzhan, Sara</td>
<td>876, 897</td>
</tr>
<tr>
<td>Shou, Wenchi</td>
<td>321</td>
</tr>
<tr>
<td>Skibniewski, Miroslaw J.</td>
<td>110, 337</td>
</tr>
<tr>
<td>Son, Hyojoo</td>
<td>194, 640</td>
</tr>
<tr>
<td>Su, Feng</td>
<td>890</td>
</tr>
<tr>
<td>Subhy, Ehab</td>
<td>404</td>
</tr>
<tr>
<td>Sun, Wei</td>
<td>925</td>
</tr>
<tr>
<td>Sun, Yong</td>
<td>754</td>
</tr>
</tbody>
</table>
Szczesny, Kamil 632 Watt, Sean 544
Wei, Chia-Chen 966
Weng, Shao-Wei 377
Willette, Aaron 163
Won, Jongsung 224
Wong, Johnny Kwok Wai 393
Wonschik, Robert 499, 523
Wu, Xianguo 110, 337
Wu, Xiaodong 103
Wu, Yu-Wei 661

T
Takeuchi, Masaki 426
Teizer, Jochen 646
That, Nguyen Dinh 120, 149
Thanh, Pham Tam 120
Ting, Pi-Hui 302
Tokmakov, Georgy 768
Tran, Tri 35
Trieu, Toan 544
Truyen, Nguyen The 731

U
Usmani, Altaf 960

V
Vakiloroaya, Vahid 489
Varghese, Koshy 2, 418
Vessalas, Kirk 499, 523
Victores, Juan Gonzalez 19, 837
Viljamaa, Esa 171, 602
Vinh, Nguyen The 731
Virtanen, H. 948
Vu, The Nam 557
Vu, Tuan Anh 557
Vuran, Mehmet C. 854

W
Walbridge, Scott 847
Waldron, Kenneth 43
Waller, Steven Travis 466, 680
Wang, Chao 134
Wang, Han-Hsiang 377
Wang, Haoran 393
Wang, Jun 321
Wang, Kun-Chi 593
Wang, Shih-Hsu 593
Wang, Wei-Chih 377, 593
Wang, Xiangyu 4, 295, 321, 738
Wang, Limao 110, 337
Zhong, Jingbing 110, 337
Zabel, Nael 398
Zaeri, Fahimeh 443
Zankoul, Emile 368
Zayed, Tarek 278