

Lecture Notes in Electrical Engineering 352

James J. (Jong Hyuk) Park
Han-Chieh Chao
Hamid Arabnia
Neil Y. Yen *Editors*

Advanced Multimedia and Ubiquitous Engineering

Future Information Technology



Springer

Lecture Notes in Electrical Engineering

Volume 352

Board of Series editors

Leopoldo Angrisani, Napoli, Italy
Marco Arteaga, Coyoacán, México
Samarjit Chakraborty, München, Germany
Jiming Chen, Hangzhou, P.R. China
Tan Kay Chen, Singapore, Singapore
Rüdiger Dillmann, Karlsruhe, Germany
Haibin Duan, Beijing, China
Gianluigi Ferrari, Parma, Italy
Manuel Ferre, Madrid, Spain
Sandra Hirche, München, Germany
Faryar Jabbari, Irvine, USA
Janusz Kacprzyk, Warsaw, Poland
Alaa Khamis, New Cairo City, Egypt
Torsten Kroeger, Stanford, USA
Tan Cher Ming, Singapore, Singapore
Wolfgang Minker, Ulm, Germany
Pradeep Misra, Dayton, USA
Sebastian Möller, Berlin, Germany
Subhas Mukhopadhyay, Palmerston, New Zealand
Cun-Zheng Ning, Tempe, USA
Toyoaki Nishida, Sakyo-ku, Japan
Bijaya Ketan Panigrahi, New Delhi, India
Federica Pascucci, Roma, Italy
Tariq Samad, Minneapolis, USA
Gan Woon Seng, Nanyang Avenue, Singapore
Germano Veiga, Porto, Portugal
Haitao Wu, Beijing, China
Junjie James Zhang, Charlotte, USA

About this Series

“Lecture Notes in Electrical Engineering (LNEE)” is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer’s high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer’s other Lecture Notes series, LNEE will be distributed through Springer’s print and electronic publishing channels.

More information about this series at <http://www.springer.com/series/7818>

James J. (Jong Hyuk) Park · Han-Chieh Chao
Hamid Arabnia · Neil Y. Yen
Editors

Advanced Multimedia and Ubiquitous Engineering

Future Information Technology

Editors

James J. (Jong Hyuk) Park
Department of Computer Science
and Engineering
Seoul University of Science & Technology
Seoul
Korea

Han-Chieh Chao
Institute of Computer Science
and Information Engineering
National Ilan University
I-Lan
Taiwan

Hamid Arabnia
Department of Computer Science
University of Georgia
Georgia
USA

Neil Y. Yen
School of Computer Science
and Engineering
The University of Aizu
Fukushima
Japan

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-3-662-47486-0 ISBN 978-3-662-47487-7 (eBook)
DOI 978-3-662-47487-7

Library of Congress Control Number: 2015940892

Springer Heidelberg New York Dordrecht London

© Springer-Verlag Berlin Heidelberg 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer-Verlag GmbH Berlin Heidelberg is part of Springer Science+Business Media
(www.springer.com)

Message from the FutureTech 2015 General Chairs

FutureTech 2015 is the 10th event of the series of international scientific conference. This conference takes place on May 18-20, 2015 in Hanoi, Vietnam. The aim of the FutureTech 2015 is to provide an international forum for scientific research in the technologies and application of information technology. FutureTech 2015 is the next edition of FutureTech 2014 (Zhangjiajie, China), FutureTech 2013 (Gwangju, Korea), FutureTech 2012 (Vancouver, Canada), FutureTech 2011 (Loutraki, Greece), FutureTech 2010 (Busan, Korea, May 2010) which was the next event in a series of highly successful the International Symposium on Ubiquitous Applications & Security Services (UASS-09, USA, Jan. 2009), previously held as UASS-08 (Okinawa, Japan, Mar. 2008), UASS-07 (Kuala Lumpur, Malaysia, August, 2007), and UASS-06 (Glasgow, Scotland, UK, May, 2006).

The conference papers included in the proceedings cover the following topics: Hybrid Information Technology High Performance Computing, Cloud and Cluster Computing, Ubiquitous Networks and Wireless Communications Digital Convergence, Multimedia Convergence, Intelligent and Pervasive Applications, Security and Trust Computing, IT Management and Service Bioinformatics and Bio-Inspired Computing, Database and Data Mining, Knowledge System and Intelligent Agent, Game and Graphics Human-centric Computing and Social Networks, Advanced Mechanical Engineering, Computer Aided Machine Design, Control and Automations & Simulation. Accepted and presented papers highlight new trends and challenges of future information technologies. We hope readers will find these results useful and inspiring for their future research.

We would like to express our sincere thanks to Steering Chair: James J. Park (SeoulTech, Korea) and Hamid R. Arabnia (The University of Georgia, USA). Our special thanks go to the Program Chairs: Joon-Min Gil (Catholic University of Daegu, Korea), Neil Y. Yen (The University of Aizu, Japan), Muhammad Khurram Khan (King Saud University, Saudi Arabia), all Program Committee members and all reviewers for their valuable efforts in the review process that helped us to guarantee the highest quality of the selected papers for the conference.

We cordially thank all the authors for their valuable contributions and the other participants of this conference. The conference would not have been possible without their support. Thanks are also due to the many experts who contributed to making the event a success.

C.S. Raghavendra, University of Southern California, USA
Jason C. Hung, Oversea Chinese University, Taiwan
Doo-soon Park, SoonChunHyang University, Korea
Jianhua Ma, Hosei University, Japan

FutureTech 2015 General Chairs

Message from the FutureTech 2015 Program Chairs

Welcome to the 10th International Conference on Future Information Technology (FutureTech 2015), which will be held in Hanoi, Vietnam on May 18-20, 2015. FutureTech 2015 will be the most comprehensive conference focused on the various aspects of information technologies. It will provide an opportunity for academic and industry professionals to discuss recent progress in the area of future information technologies. In addition, the conference will publish high quality papers which are closely related to the various theories and practical applications in multimedia and ubiquitous engineering. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in these important subjects.

For FutureTech 2015, we received many paper submissions, after a rigorous peer review process, we accepted only articles with high quality for the FutureTech 2015 proceedings, published by the Springer. All submitted papers have undergone blind reviews by at least two reviewers from the technical program committee, which consists of leading researchers around the globe. Without their hard work, achieving such a high-quality proceeding would not have been possible. We take this opportunity to thank them for their great support and cooperation. We would like to sincerely thank the following invited speakers who kindly accepted our invitations, and, in this way, helped to meet the objectives of the conference: Prof. Han-Chieh Chao, National Ilan University, Taiwan and Prof. Timothy K. Shih, National Central University, Taiwan. Finally, we would like to thank all of you for your participation in our conference, and also thank all the authors, reviewers, and organizing committee members. Thank you and enjoy the conference!

Joon-Min Gil, Catholic University of Daegu, Korea
Neil Y. Yen, The University of Aizu, Japan
Muhammad Khurram Khan, King Saud University, Saudi Arabia

FutureTech 2015 Program Chairs

Organization

Steering Chair

James J. Park
Hamid R. Arabnia

SeoulTech, Korea
The University of Georgia, USA

General Chairs

C.S. Raghavendra
Jason C. Hung
Doo-soon Park
Jianhua Ma

University of Southern California, USA
Oversea Chinese University, Taiwan
SoonChunHyang University, Korea
Hosei University, Japan

General Vice-Chairs

Hwa-Young Jeong
Cho-Li Wang

Kyung Hee University, Korea
University of Hong Kong, Hong Kong

Program Chairs

Joon-Min Gil
Neil Y. Yen
Muhammad Khurram Khan

Catholic University of Daegu, Korea
The University of Aizu, Japan
King Saud University, Saudi Arabia

Workshop Chairs

Deok-Gyu Lee
Vincent Huang

Ka Lok Man

Seowon University, Korea
National Taichung University of Science
and Technology, Taiwan
Xi'an Jiaotong-Liverpool University, China

International Advisory Committee

Yi Pan	Georgia State University, USA
Qun Jin	Waseda, Japan
Fatos Xhafa	Technical University of Catalonia, Spain
Hsiao-Hwa Chen	National Cheng Kung University, Taiwan
Ivan Stojmenovic	University of Ottawa, Canada
Laurence T. Yang	St. Francis Xavier University, Canada
Young-Sik Jeong	Dongguk University, Korea

Publicity Chairs

Eunyoung Lee	Dongduk Women's University, Korea
Byung-Gyu Kim	Sun Moon University, Korea
Sung-Ki Kim	Sun Moon University, Korea

Program Committee

Alfredo Cuzzocrea	University of Calabria, Italy
Amagasa Toshiyuki	University of Tsukuba, Japan
Byna Suren	Lawrence Berkeley National Laboratory, USA
Caldelli Roberto	Universita degli Studi di Firenze
Cerquitelli Tania	Politecnico di Torino, Canada
Cha Yue-Shan	National Taipei University, Taiwan
Chen Bing	Memorial University of Newfoundland, Canada
Chen WeiFeng	California University of Pennsylvania, USA
Chi-Fu Huang	National Chung Cheng University, Taiwan
Ching-Hsien Hsu	Chung Hua University, China
Davidovic Tatjana	Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia
Edward Hua	QED Systems, USA
Gonzalez JoseAntonio	Universidad de Malaga, Spain
Homenda Wadysaw	Instytut Badan Systemowych Polskiej Akademii Nauk, Poland
Hu Yu-Chen	Providence University, Taiwan
Ivetic Dragan	University of Novi Sad, Serbia
Jiqiang Lu	Institute for Infocomm Research, Singapore
Kapetanios Epaminondas	University of Westminster, UK
Klyuev Vitaly	University of Aizu, Japan
Kyungbaek Kim	Chonnam National University, Korea
Lu Leng	Southwest Jiaotong University Emei Campus, China
Maumita Bhattacharya	Charles Sturt University, Australia
Pai-Ling Chang	Shih-Hsin University, Taiwan
Picard Willy	Poznan University of Economics, Poland

Qiang He	Swinburne University of Technology, Australia
Raylin Tso	National Chengchi University, China
Ren-Song Ko	National Chung Cheng University, Korea
Rios Ruben	Universidad de Malaga, Spain
Rudolf Tsoy	Far Eastern State Academy, Russia
Salem Abdelbadeeh	Ain Shams University, Egypt
Song Fu	University of North Texas, USA
Wei-Chuen Yau	Multimedia University, Malaysia
Wookey Lee	Inha University, Korea
Wyne Mudasser	National University, USA
Xiao Liu	East China Normal University, China
Yeun Chan	Yeob Khalifa University of Science, Technology and Research, UAE
Yo-Ping Huang	National Taipei University of Technology, Taiwan
Zhang Yunquan	State Key Lab of Computer Science, China
Zhang Zhiqiang	Harbin Engineering University, China

Message from the MUE 2015 General Chairs

MUE 2015 is the 9th event of the series of international scientific conference. This conference takes place on May 18-20, 2015 in Hanoi, Vietnam. The aim of the MUE 2015 is to provide an international forum for scientific research in the technologies and application of Multimedia and Ubiquitous Engineering. Ever since its inception, International Conference on Multimedia and Ubiquitous Engineering has been successfully held as MUE-14 (Zhangjiajie, China, May 2014), MUE-13 (Seoul, Korea, May 2013), MUE-12 (Madrid, Spain, July 2012), MUE-11 (Loutraki, Greece, June 2011), MUE-10 (Cebu, Philippines, August 2010), MUE-09 (Qingdao, China, June 2009), MUE-08 (Busan, Korea, April 2008), and MUE-07 (Seoul, Korea, April 2007).

The conference papers included in the proceedings cover the following topics: Multimedia Modeling and Processing, Ubiquitous and Pervasive Computing, Ubiquitous Networks and Mobile Communications, Intelligent Computing, Multimedia and Ubiquitous Computing Security, Multimedia and Ubiquitous Services, Multimedia Entertainment, IT and Multimedia Applications. Accepted and presented papers highlight new trends and challenges of Multimedia and Ubiquitous Engineering. We hope readers will find these results useful and inspiring for their future research.

We would like to express our sincere thanks to Steering Chair: James J. (Jong Hyuk) Park (SeoulTech, Korea). Our special thanks go to the Program Chairs: Gangman Yi (Gangneung-Wonju National University, Korea) and Chengcui Zhang (The University of Alabama at Birmingham, USA), all Program Committee members and all reviewers for their valuable efforts in the review process that helped us to guarantee the highest quality of the selected papers for the conference.

Shu-Ching Chen, Florida International University, USA
Young-Sik Jeong, Dongguk University, Korea
Han-Chieh, Chao National Ilan University, Taiwan

MUE 2015 General Chairs

Message from the MUE 2015 Program Chairs

Welcome to the 9th International Conference on Multimedia and Ubiquitous Engineering (MUE 2015), which will be held in Hanoi, Vietnam on May 18-20, 2015. MUE 2015 will be the most comprehensive conference focused on the various aspects of multimedia and ubiquitous engineering. It will provide an opportunity for academic and industry professionals to discuss recent progress in the area of multimedia and ubiquitous environment. In addition, the conference will publish high quality papers which are closely related to the various theories and practical applications in multimedia and ubiquitous engineering. Furthermore, we expect that the conference and its publications will be a trigger for further related research and technology improvements in these important subjects.

For MUE 2015, we received many paper submissions, after a rigorous peer review process, we accepted only articles with high quality for the MUE 2015 proceedings, published by the Springer. All submitted papers have undergone blind reviews by at least two reviewers from the technical program committee, which consists of leading researchers around the globe. Without their hard work, achieving such a high-quality proceeding would not have been possible. We take this opportunity to thank them for their great support and cooperation. Finally, we would like to thank all of you for your participation in our conference, and also thank all the authors, reviewers, and organizing committee members. Thank you and enjoy the conference!

Gangman Yi, Gangneung-Wonju National University, Korea
Chengcui Zhang, University of Alabama at Birmingham, USA

MUE 2015 Program Chairs

Organization

Steering Chair

James J. Park

SeoulTech, Korea

General Chairs

Shu-Ching Chen

Young-Sik Jeong

Han-Chieh

Florida International University, USA

Dongguk University, Korea

Chao National Ilan University, Taiwan

General Vice-Chairs

Weijia Jia

Hwa-Young Jeong

City U. of Hong Kong, Hong Kong

Kyung Hee University, Korea

Program Chairs

Gangman Yi

Chengcui Zhang

Gangneung-Wonju National University, Korea

University of Alabama at Birmingham, USA

Workshop Chairs

Namje Park

Xu Shuo

Jeju National University, Korea

ISTIC, China

International Advisory Committee

Borko Furht

Thomas Plagemann

Roger Zimmermann

Florida Atlantic University, USA

University of Oslo, Norway

National University of Singapore, Singapore

XVIII Organization

Hamid R. Arabnia	The University of Georgia, USA
Stephan Olariu	Old Dominion University, USA
Albert Zomaya	University of Sydney, Australia
Yi Pan	Georgia State University USA
Koji Nakano	University of Hiroshima, Japan

Publicity Chairs

Cheonshik Kim	Anyang University, Korea (Leading Chair)
Mohamed Gaber	University of Portsmouth, UK
Ryan Leong Hou U	Universidade De Macau, China
Chengjiu Yin	Kyushu University, Japan
Jian-Lian Chen	Aletheia University, Taiwan
Junbo Wang	University of Aizu, Japan
Nan-Chen Hsieh	National Taipei University and Health Sciences, Taiwan

Program Committee

Afrand Agah	West Chester University of Pennsylvania, USA
Akihiro Sugimoto	National Institute of Informatics, Japan
Angel D. Sappa	Universitat Autònoma de Barcelona, Spain
Bin Lu	West Chester University, USA
Ch. Z. Patrikakis	Technological Education Institute of Pir, Greece
Chao-Tung Yang	Tunghai University, Taiwan
Dakshina Ranjan Kisku	Asansol Engineering College, India
Dalton Lin	National Taipei University, Taiwan
Debzani Deb	Winston-Salem State University, USA
Dongkyun Kim	KISTI, Korea
Ezendu Ariwa	University of Bedfordshire, UK
Guillermo Camara Chavez	Universidade Federal de Minas Gerais, Brasil
Hai Jin	Huazhong University of Sci & Tech, China
HaRim Jung	Sungkyunkwan University, Korea
HeonChang Yu	Korea University, Korea
Hermann Hellwagner	Klagenfurt University, Austria
Jin Kwak	Ajou University, Korea
Jungong Han	Civolution Technology, Netherlands
Jun-Won Ho	Seoul Women's University, Korea
Kilhung Lee	Seoul National University of Science, Korea
Kwang Sik Chung	Korea National Open University, Korea
Marco Cremonini	University of Milan, Italy
Mario Doeller	University of Applied Science, Germany
Paisarn Muneesawang	Naresuan University, Thailand

Pascal Lorenz	University of Haute Alsace, France
Rainer Unland	University of Duisburg-Essen, Germany
Reinhard Klette	The University of Auckland, New Zealand
Se-Hak Chun	Seoul National University of Science, Korea
Seunghae Kim	KISTI, Korea
Sokratis Katsikas	University of Piraeus, Greece
Teng Li	Baidu Inc., China
Wesley De Neve	Ghent University, Belgium
Yan Liu	Hong Kong Polytechnic University, China
Young-Gab Kim	Sejong University, Korea
Zheng-Jun Zha	National University of Singapore, Singapore

Contents

Development of a Bidirectional Transformation Supporting Tool for Formalization with Logical Formulas and Its Application	1
<i>Shunsuke Nanaumi, Kazunori Wagatsuma, Hongbiao Gao, Yuichi Goto, Jingde Cheng</i>	
An Extension of QSL for E-testing and Its Application in an Offline E-testing Environment	7
<i>Zhe Wang, Yuan Zhou, Bo Wang, Yuichi Goto, Jingde Cheng</i>	
A Tasking Deadlock Detector for Ada 2012 Programs	15
<i>Bo Wang, Takeo Ekiba, Yuichi Goto, Jingde Cheng</i>	
Automated Theorem Finding by Forward Reasoning Based on Strong Relevant Logic: A Case Study in Graph Theory	23
<i>Hongbiao Gao, Yuichi Goto, Jingde Cheng</i>	
A Middleware Supporting Query Processing on Distributed CUBRID	31
<i>Hyeong-Il Kim, Min Yoon, YoungSung Shin, Jae-Woo Chang</i>	
A Discovery Support Scheme for Inter-domain DDS Gateways in Cyber-Physical Systems	39
<i>Wooyeob Lee, Sungmoon Chung, Sungryung Cho, Inwhee Joe, Jeman Park</i>	
Robust Facial States Estimation against Occlusion and Inference of Driver's Drowsiness Using Hidden Markov Model	45
<i>In-Ho Choi, Chan-Hee Jeong, Yong-Guk Kim</i>	
The Disaster Rescue Robot Design and Implementation Using Open Source	53
<i>Yung-Hui Chen, Jyu-Wei Wang</i>	

The Direct and Indirect Effects of Perceived Usability of Smartphone Applications on User Satisfaction	61
<i>Wonjin Jung, Taehwan Kim</i>	
A New Query Integrity Verification Method for Encrypted Data in Database Outsourcing	67
<i>Miyoung Jang, Min Yoon, Youngho Song, Jae-Woo Chang</i>	
Preventing Private Information in Secure Dissemination	73
<i>Hye-Kyeong Ko</i>	
Possibility of Using Embedded Sensors of Smart Devices for Augmented Reality Application	79
<i>Ondrej Bilek, Ondrej Krejcar</i>	
Economic and Technological Aspects of Social Networks in European Business Sector	85
<i>Petra Maresova</i>	
Decision Making Criteria for Cloud Computing Deployment	93
<i>Petra Maresova</i>	
Periocular Recognition Based on LBP Method and Matching by Bit-Shifting	99
<i>So Ra Cho, Gi Pyo Nam, Kwang Yong Shin, Dat Tien Nguyen, Kang Ryoung Park</i>	
Residual Histogram Shifting Technique Based on Cascading Prediction for Reversible Data Hiding	105
<i>Yu-Chen Hu, PiYu Tsai, Jieh-Shan Yeh, Wu-Lin Chen</i>	
Color and Depth Image Correspondence for Kinect v2	111
<i>Changhee Kim, Seokmin Yun, Seung-Won Jung, Chee Sun Won</i>	
Reduced Reference Quality Metric for Depth Images	117
<i>Thanh-Ha Le, Seongjo Lee, Seung-Won Jung, Chee Sun Won</i>	
Sustainable Operation Algorithm for High Availability with Integrated Desktop Storage Based on Virtualization	123
<i>Hyun-Woo Kim, HwiRim Byun, Eun-Ha Song, Young-Sik Jeong</i>	
Analysis on Temporal Sparsity of Human Face for Various Emotion Stimuli	129
<i>Unsoo Jang, Byoung Cheul Kim, Yoonkyoung Kim, Min Woo Park, Eui Chul Lee</i>	
Monocular Eye Tracking System Using Webcam and Zoom Lens	135
<i>Kun Ha Suh, Yun-Jung Kim, Yoonkyoung Kim, Daejune Ko, Eui Chul Lee</i>	

Many-to-Many Data Collection for Mobile Users in Wireless Sensor Networks	143
<i>Chi-Fu Huang, Wei-Chen Lin</i>	
Method of Path Computation Using National Research Network Federation	149
<i>Jinhyung Park, Joon-Min Gil, Woohyun Kim, Jeongwook Park, Hyunhun Cho</i>	
Robot Reinforcement Learning for Automatically Avoiding a Dynamic Obstacle in a Virtual Environment	157
<i>Phuong Chu, Hoang Vu, Donghyeon Yeo, Byeonggwon Lee, Kyhyun Um, Kyungeun Cho</i>	
Smart Virtual Lab Using Hand Gestures	165
<i>Warda Ikram, Yoonji Jeong, Byeonggwon Lee, Kyhyun Um, Kyungeun Cho</i>	
Internal Topology Based Flexible Shortest Path Planning Method for Indoor Navigation	171
<i>Yan Li, Byeong-Seok Shin</i>	
Result Verification Scheme with Resource Clustering in Desktop Grids	177
<i>Joon-Min Gil, Yong-Hyun Cho, Sung-Hwa Hong</i>	
RDEA: A Novel Video Encryption Algorithm	183
<i>Zhiyong Li, Xingjun Wang, Yuxi Lin, Chao Cheng</i>	
Embedding Data into Audio Signal by Combining Sliding Window Technique with a Novel Marking Bit Method	191
<i>Vu Van Tam, Tran Duc-Tan, Nguyen Thanh Thuy, Phan Trong Hanh</i>	
A Study on the Hierarchical Structure of Conditional Replenishment Algorithm for SC-MMH 3DTV System	199
<i>Kyeong-Hoon Jung, Sung-Hoon Kim, Dong-Wook Kang</i>	
Performance Comparison for CFDP NAK Modes in Deep Space Communications	205
<i>Yong Li, Kyungrak Lee, Yantao Kan, Inwhee Joe, Yeonyi Choi</i>	
The Design and Implementation of Mesh-Based DTN for Disaster Situations	211
<i>Jaewon Lee, Abdul Aziz, Suoning Ma, Inwhee Joe, Yeonyi Choi</i>	
A Fire Evacuation Guidance System Based on Ubiquitous Sensor Networks	217
<i>Yeonyi Choi, Seokjoon Hong, Inwhee Joe</i>	
Stage Image Control System Using Visual Tracking	223
<i>Sooyeon Lim, Jaeha Lyu</i>	

Real-Time Virtual Lego Brick Manipulation Based on Hand Gesture Recognition	231
<i>Tran Van Thanh, Dongho Kim, Young-Sik Jeong</i>	
Object Detection Based on Exemplar Object Expression	239
<i>Yuanyou Wang, Xiaoru Wang, Junping Du, Tianming Du</i>	
An Algorithm for Image Classification Based on Semantic Transfer Learning	249
<i>Tianming Du, Xiaoru Wang, Junping Du, Yuanyou Wang</i>	
Automatic Dehumidifier Control System for Greenhouse Using Smart Phone	257
<i>Young-Jae Lee, Kyung-Wook Park, Eung-Kon Kim</i>	
Thermal Image-Based Disc Pads Diagnosis System in Grab Crane	265
<i>Yeon-Jae Oh, Kyoung-Wook Park, Eung-Kon Kim</i>	
Image Acquisition System for Construction Inspection Based on Small Unmanned Aerial Vehicle	273
<i>Sungsuk Choi, Eungkon Kim</i>	
Portable Video System for Farm Growth Monitoring	281
<i>Mi-jeong Park, Eung-gon Kim</i>	
Design of Flight Stabilization System for Acquisition of UAV Based Monitoring Image	289
<i>Oh-hoon Cho, Eung-Kon Kim</i>	
The Smart and Secure Protocol for Mobile Office Environments	297
<i>Hyun-A. Park, JongSung Park, Jee-In Kim, JaeHyun Park</i>	
Research for Personalized Recommendation of Learning Resource in Mobile Computing Context	309
<i>Lina Yang, Yi Yang</i>	
Terrain Classification Using Adaboost Algorithm Based on Co-occurrence and Haar-like Features	317
<i>Ngoc-Hoa Nguyen, Dong-Min Woo</i>	
Data Recording System Based on FPGA and DSP	323
<i>Jun Wang, Xiaoliang Wang, Zhipeng Zhao, Hong Xiang, Yuxi Zhang</i>	
Implementation of High-Speed Serial Interconnects for Multi-Processor Parallel System	331
<i>Jun Wang, Zhipeng Zhao, Bin Yang, Wengui Fan, Yuxi Zhang</i>	

A Twin-Screw Rotor Profile Design Method Based on Computational Fluid Dynamics	337
<i>He Xueming, Pan Chenglong, Wu Meiping, Ji Xiaogang</i>	
Design of Nanograting Structures for Optoelectronic Devices Based on Rigorous Coupled-Wave Analysis	343
<i>Nghia Nguyen-Huu, Jaromir Pistora, Michael Cada</i>	
An Improved Algorithm of Network Maximum Flow Based on Network Flow Matrix	351
<i>ZhenChao Wang, LiPing Zhang, WeiDong Hao</i>	
Author Index	357

Development of a Bidirectional Transformation Supporting Tool for Formalization with Logical Formulas and Its Application

Shunsuke Nanaumi, Kazunori Wagatsuma, Hongbiao Gao,
Yuichi Goto, and Jingde Cheng

Department of Information and Computer Sciences,
Saitama University, Saitama 338-8570, Japan
cheng@aise.ics.saitama-u.ac.jp

Abstract. In many applications in computer science and artificial intelligence, logical formulas are used as a formal representation to represent and/or specify various objects and relationships among them. Transforming logical formulas into informal propositional statements, e.g., declarative sentences and mathematical formulas, is important as well as transforming informal propositional statements into logical formulas. When people obtain new logical formulas as results of deduction/reasoning based on logic, investigating the obtained formulas is also not an easy task for them. Although information systems with proving, e.g., automated theorem proving systems, formal verification systems, etc., are used in various field, in the future, information systems with reasoning, e.g., automated theorem finding systems, will also be developed and used in various fields. Thus, a tool to support bidirectional transformation between informal propositional statements and logical formulas will be demanded at that time. This paper presents an implementation of a bidirectional transformation supporting tool for formalization with logical formulas. The paper also shows application of the tool in a case study of automated theorem finding with forward reasoning.

Keywords: Bidirectional transformation, Logical formula, Formalization, Informal propositional statement.

1 Introduction

In many applications in computer science and artificial intelligence, logical formulas are used as a formal representation to represent and/or specify various objects and relationships among them [3]. Transforming logical formulas into informal propositional statements e.g., declarative sentences and mathematical formulas, is important as well as transforming the informal propositional statements into logical formulas. When people obtain new logical formulas as results of deduction/reasoning based on logic, investigating the obtained formulas is also not an easy task for them. Although information systems with proving, e.g., automated theorem proving systems, formal verification systems, etc., are used in various field, in the future, information systems

with reasoning, e.g., automated theorem finding systems [2], will also be developed and used in various fields.

Thus, a tool to support bidirectional transformation between informal propositional statements and logical formulas will be demanded at that time. Therefore, we proposed a bidirectional transformation method for formalization with logical formulas and designed a supporting tool for transformation between informal propositional statements and logical formulas [7], but have not implemented it yet.

This paper presents an implementation of a bidirectional transformation supporting tool for formalization with logical formulas. The paper also shows an application of the tool in a case study of automated theorem finding with forward reasoning.

The rest of this paper is organized as follows. Section 2 presents a bidirectional transformation supporting tool. Section 3 explains implementation of a bidirectional transformation supporting tool. Section 4 shows application of the tool in a case of automated theorem finding with forward reasoning, and concluding remarks is given in section 5.

2 Bidirectional Transformation Supporting Tool

2.1 Transformation Method

Any bidirectional transformation between informal propositional statements and logical formulas consists of six activities [7]. Four of six activities are the activities of transforming informal propositional statements into logical formulas, i.e., unifying, complementing, assigning, and generating. “Unifying” is to unify expressions which have same meaning. “Complementing” is to complement omitted expressions. “Assigning” is to assign words/phrases of the informal propositional statements to symbols of logical formulas. “Generating” is to generate the logical formulas from the informal propositional statements. Two of six activities are the activities of transforming logical formulas into informal propositional statements, i.e., replacing and changing. “Replacing” is to replace symbols of logical formulas with words/phrases of target informal propositional statements according to the assignment of symbols to words/phrases used in transforming the target informal propositional statements into logical formulas. “Changing” is to change word order of the replaced formula according to a rule of the informal propositional statements.

2.2 Design of a Bidirectional Transformation Supporting Tool

We designed a bidirectional transformation supporting tool for formalization with logical formulas [7]. A bidirectional transformation supporting tool gives its users instructions and supports to do transformation between informal propositional statements and logical formulas easily. When the users do bidirectional transformation tasks, the supporting tool instructs the users to do transformation tasks every process in the six activities.

As shown in Fig. 1, the supporting tool consists of ten components: unifying component to support “unifying”, complementing component to support “complement”, assigning component to support “assigning”, generating component to support “generating”,

word order changing component to support “changing”, symbol replacing component to support “replacing”, a morphological analyzer, and two databases [7]. A morphological analyzer is an external tool to do morphological analysis of declarative sentences. A synonym DB is a database that synonyms in declarative sentences are saved and to find synonyms in sentences. A user’s task DB is a database in which transforming data are stored.

The supporting tool can deal with different kinds of informal propositional statements from the view point of data structure [7]. It is not necessary to change the data structure if a function to transform a new kind of informal propositional statements is added into the supporting tool, because each data table in the supporting tool is independent with different kinds of informal propositional statements.

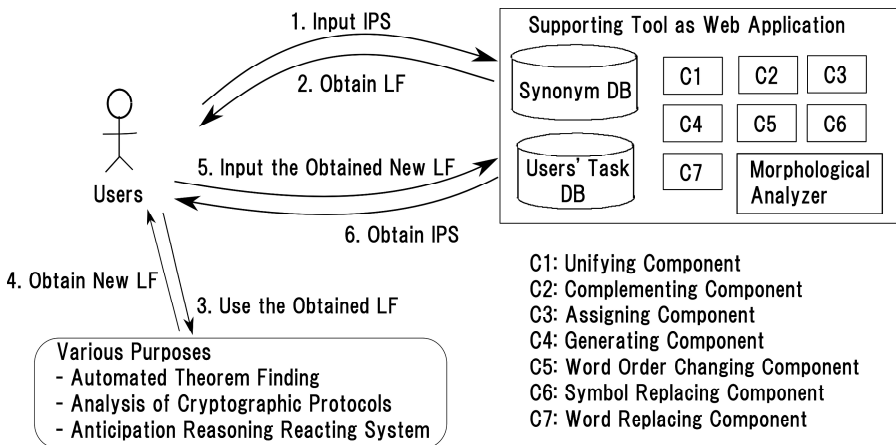


Fig. 1. Overview of a supporting tool

3 Implementation of a Bidirectional Transformation Supporting Tool

We implemented a bidirectional transformation supporting tool for formalization with logical formulas as web application. We implemented it by Java and Play Framework [4] which is a web application development framework. Moreover, we used WordNet [1, 6] as a lexical database, Enju [8] and Cabocha [5].

When a user uses the tool, the user inputs a set of informal propositional statements or logical formulas via own web browser. After that, the user does transformation tasks every one statement according to instructions given by the tool. Moreover, the tool can lighten the burden imposed on users in transforming a lot of statements or formulas because the users can do transformation tasks at same time in several people via the Internet.

4 Application of the Tool in a Case Study of Automated Theorem Finding with Forward Reasoning

The problem of automated theorem finding (ATF for short): “What properties can be identified to permit an automated reasoning program to find new and interesting theorems, as opposed to proving conjectured theorems?” [2]. In mathematical fields, to find new theorems, mathematicians continue to define more complex concepts by using previously given definitions and axioms, and already defined concepts. Then, mathematicians think, assume, and prove propositions by using the defined complex concepts. After that, they obtain new theorems. A systematic methodology of ATF by forward reasoning based on the strong relevant logic [2] has been proposed. To use the methodology, it is necessary to transform theorems expressed by mathematical formulas into theorems expressed by logical formulas. Moreover, it is necessary to transform new theorems expressed by logical formulas obtained by the methodology into theorems expressed by mathematical formulas. Our tool is needed to do such as bidirectional transformation such that it can help mathematicians to do ATF. To show the application of the tool, we used our tool for transformation between logical formulas and mathematical formulas in a case study of ATF in von Neumann-Bernays-Gödel (NBG) set theory [9].

Table 1. State of formulas in each process of bidirectional transformation

	(1)	(2)
Inputted	$(x \in V) \Rightarrow (U(x) \in V)$	$\forall x (O(x) \rightarrow C(x, F(V, V)))$
Step1	$\forall x ((x \in V) \Rightarrow (U(x) \in V))$	$O(x) \rightarrow C(x, F(V, V))$
Step2	$\forall x (\in(x V) \Rightarrow \in(U(x) V))$	$ONEONE(x) \rightarrow \subseteq(x, \times(V, V))$
Step3	$\forall x (B(x V) \Rightarrow B(U(x) V))$	If ONEONE(x), then $\subseteq(x, \times(V, V))$
Outputted	$\forall x (B(x V) \rightarrow B(U(x) V))$	If ONEONE(x), then $(x \subseteq (V \times V))$

In a following formula, you should add quantifiers **[ALL.x]** or **[EXIST.x]** and its scope to following formula in variable **x**.

Target statement:

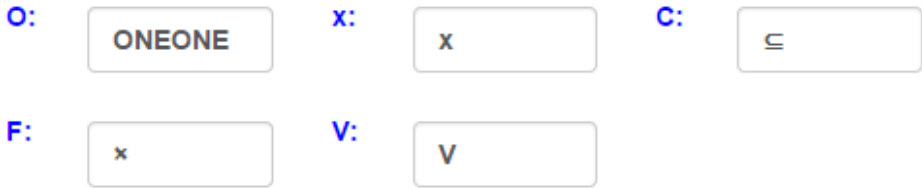
$(x \in V) \Rightarrow (U(x) \in V)$

NEXT

Fig. 2. Adding quantifiers in “Complementing” process in the tool

Table 2. Assignment of symbols of mathematical formulas to symbols of logical formulas

Math.	ONEONE	x	\subseteq	\times	\forall	=	\in	U	u
Logic.	O	x	C	F	V	E	B	U	u



Target formula is:

$$\forall x(O(x) \Rightarrow (C(x, F(V, V))))$$

Replacing

Fig. 3. “Replacing” process in the tool

Table 1 shows progress of transformation between mathematical formulas and logical formulas. Transformation of a mathematical formula into a logical formula is shown in (1) of Table 1. First, users add quantifiers to an inputted mathematical formula. Fig. 2 shows its process in the tool. Secondly, the users replace the target formula with one expressed in polish notation. Thirdly, the users assign symbols of the mathematical formula with predicates and terms of logical formulas. Fourthly, the users replace connectives of mathematical formulas with those of logical formulas. Finally, the users generate a logical formula from the inputted mathematical formula. Moreover, transforming a logical formula into a mathematical formula is shown in (2) of Table 1. First, users remove quantifiers from the inputted formula. Secondly, the users replace predicates and terms of logical formulas with symbols of mathematical formulas according to the assignment of symbols of logical formulas to symbols of mathematical formulas used in transforming mathematical formulas into logical formulas. Table 2 shows the assignment relation and Fig. 3 shows its process. Thirdly, the users replace connectives of logical formulas with those of mathematical formulas. Fourthly, the users change word order of logical formulas to that of mathematical formulas. Finally, the users generate a mathematical formula. As results of such as bidirectional transformation, we showed an application of the tool in a case of automated theorem finding with forward reasoning by showing that the tool supports to do bidirectional transformation.

5 Concluding Remarks

We presented an implementation of a bidirectional transformation supporting tool for formalization with logical formulas and showed an application of the tool in a case of automated theorem finding with forward reasoning. By using the bidirectional transformation supporting tool for bidirectional transformation between informal propositional statements and logical formulas, we can use logical formulas generated by the tool for a formal representation in many application in computer science and artificial intelligence, and we can use informal propositional statements generated by the tool for investigating new logical formulas obtained as results of deduction/reasoning based on logic. Our supporting tool can support many information systems with forward reasoning in various field in the future.

As future works, we will adapt the bidirectional transformation supporting tool to the other kinds of informal propositional statements, e.g., Chinese declarative sentences, expand the field to use our tool, and make the tool used easily.

References

1. Bond, F., Isahara, H., Fujita, S., Uchimoto, K., Kuribayashi, T.: Kanzaki. K.: Enhancing the Japanese WordNet, In: Proc. 7th Workshop on Asian Language Resources, pp. 1–8 (2009)
2. Gao, H., Goto, Y., Cheng, J.: A Systematic Methodology for Automated Theorem Finding. *Theoretical Computer Science* 554, 2–21 (2014)
3. Hurley, P.: *A Concise Introduction to Logic*. Wadsworth (1996)
4. Karunakaran, K.: *Play Framework 2 – For Java Developers*. CreateSpace Independent Publishing Platform (2013)
5. Kudo, T., Matsumoto, Y.: Japanese Dependency Analysis Using Cascaded Chunking. In: Proc. 6th Workshop on Computational Language Learning, Taipei, pp. 63–69 (2002)
6. Miller, A.: WordNet: A Lexical Database for English. *Communications of the ACM* 38(11), 39–41 (1995)
7. Nanaumi, S., Wagatsuma, K., Gao, H., Goto, Y., Cheng, J.: A Bidirectional Transformation Supporting Tool for Formalization with Logical Formulas. In: Nguyen, N.T., Trawiński, B., Kosala, R. (eds.) *ACIIDS 2015*. LNCS, vol. 9011, pp. 634–643. Springer, Heidelberg (2015)
8. Ninomiya, T., Matsuzaki, T., Miyao, T., Tsuji, J.: A Log-linear Model with an N-gram Reference Distribution for Accurate HPSG Parsing. In: Proc. 10th Conference on Parsing Technologies 2007, Prague, pp. 60–68 (2007)
9. Quaife, A.: *Automated Development of Fundamental Mathematical Theories*. Kluwer Academic (1992)

An Extension of QSL for E-testing and Its Application in an Offline E-testing Environment

Zhe Wang, Yuan Zhou, Bo Wang, Yuichi Goto, and Jingde Cheng

Department of Information and Computer Sciences,
Saitama University, Saitama, Japan
{wangzhe, shuugen, wangbo,
gotoh, cheng}@aise.ics.saitama-u.ac.jp

Abstract. E-testing is to perform all processes from preparing questions to marking collected answer of the questions in a completely electronic way. Now, various e-testing are done, and there are many kinds of e-testing systems. When people want to do an e-testing or order a new e-testing system, they should specify the e-testing or e-testing system. A Specification language helps to create precise and adequate specifications of e-testing and e-testing systems is demanded. QSL is a specification language for specifying various e-questionnaire and e-questionnaire systems. QSL is a hopeful candidate of the required specification language because both testing and questionnaire have similar processes. However, the current version of QSL does not take e-testing and e-testing systems into account. This paper presents an extension of QSL to deal with e-testing and e-testing systems, and shows a real application of extended QSL in case of an offline e-testing environment.

Keywords: QSL, Specification language, E-testing, Offline e-testing environment.

1 Introduction

Test is a general and indispensable method, and widely used to assess people's achievement, ability, and characteristics in education, enterprise, medicine, and government [16]. E-testing is to perform all processes from preparing questions to marking collected answer sheets in a completely electronic way. An e-testing system is a system that provides an environment with its users to do e-testing. In recent years, there are various e-testing systems have been developed ad hoc because purposes and procedures of tests are different from each other. When examiners want to use an e-testing system to execute various e-testing, they should specify the e-testing. If an e-testing system cannot satisfy examiners' needs, they may order a new e-testing system. They also should specify the e-testing system. Therefore, people need a specification language which can help them to create precise and adequate specifications for various e-testing and e-testing systems.

QSL [23] is a specification language for specifying various e-questionnaire and e-questionnaire systems, and it is a hopeful candidate of the required specification language owing to significant similarities of contents and structures between e-questionnaire

and e-testing. However, the current version of QSL does not take e-testing and e-testing systems into account. This paper proposes the extension of QSL for e-testing so that users can use QSL to specify various e-testing as well as e-testing systems. The paper also shows a real application of extended QSL in case of an offline e-testing environment. The rest of the paper is organized as followed: Section 2 gives introduces in QSL. Section 3 presents an extension of QSL for e-testing. Section 4 shows an offline e-testing environment as a use case of extended QSL. Finally, some concluding remarks are given in Section 5.

2 QSL: A Specification Language for E-questionnaire Systems

QSL serves as a communication tool among specifying, developing, and using various e-questionnaires and e-questionnaire systems [23]. QSL addresses the needs by providing various desirable functions of e-questionnaire systems, pointing out necessary items to make specification clear and precise, and being used as a format for data exchange.

QSL is based on XML [22]. QSL provides primitive elements that are used to specify various e-questionnaires and e-questionnaire systems. A primitive element consists of entity and representation. Entity is uniquely identified regardless of changing representations. Representation is to express the corresponding entity. We can describe specifications of various e-questionnaire systems by combining the primitive elements.

QSL can be used in three ways [22, 23]. At first, QSL can be used to specify e-questionnaire systems. In other words, QSL can be used to specify the functions the system provides. For example, QSL provides tags for four kinds of participants: sponsors who organize an e-questionnaire, questioners who design and ask questions, respondents who take part in answering, and analysts who analyze the result and give the results. Secondly, QSL can be used to specify e-questionnaires on a system, i.e., QSL can be used to describe restrictions in each process of e-questionnaire. For example, for an e-questionnaire, QSL is used to describe the distribution method such as e-mail, web-link, or other offline methods. At last, QSL can be used for a format of data exchange to describe questionnaire data and response data. For instance, QSL is used to describe various questions and answers.

3 An Extension of QSL for E-testing

3.1 E-questionnaire and E-testing

To clarify the differences between e-questionnaire and e-testing, we investigated 20 e-testing systems [1-15, 17-21] on the Internet. The differences between e-questionnaire and e-testing are in five aspects: participants, process, data, logic, and authorization. Firstly, participants are the people who take part in an e-questionnaire or e-testing. An e-testing system contains not only those four kinds of participants mentioned in Section 2, but also two new kinds of participants who are monitor and marker. Monitor is a person who monitors the whole e-testing for illegal behavior,

and marker is a person who marks the responses of respondents and gives the results. Secondly, both e-questionnaire and e-testing have similar procedure except marking. After collecting the response data of all respondents, marker should mark the response data and give scores of each question and total scores of each respondent. Thirdly, data of e-questionnaire and e-testing are used to describe many research papers and record many response data by different respondents. Unlike e-questionnaire, sample answer of each question is demanded to mark answers in e-testing. Similarly, score of each answer and total score of answers are also demanded. Besides, questions in e-testing involves wider and more professional field, such as mathematical formula, periodic table of chemical elements, etc. Moreover, the logic in an e-questionnaire is a facility to control which question is showed in which order. In e-testing, the logic is used to prevent cheating activities. The new logic type for e-testing is randomization which is showing the questions and options to respondents in a random order so that they cannot peep other respondents for answers. At last, the authority is used to describe the authority of all the participants both e-questionnaire and e-testing, such as the writability and readability for a research paper, response, sample answer (for e-testing) and result. However, e-testing has more strict management of authority in order to stop and prevent illegal behaviors.

To specify various e-testing and e-testing systems, these differences should be covered by QSL. However, the current version of QSL does not deal with those differences.

3.2 QSL Extension for E-testing

According to the differences between e-questionnaire systems and e-testing system, some new elements have been extended to specify various e-testing and e-testing systems.

To deal with the difference of participants, we add two new elements to describe new participants. They are *monitor* and *marker*. An example is shown below. The participant element is *monitor* which has an attribute named "id", the value of the element is used to describe the name of the monitor.

```
<monitor id="m001">John Smith</monitor>
```

In order to cover the difference of data and marking procedure, we add three new elements which are *sampleanswer*, *score*, and *result*, and we add an attribute *type* for *score* which contains "question", "section", "total", "respquestion", "respsection", and "resptotal." The element *sampleanswer* is defined to describe the sample answer of a question. The types "question", "section", and "total" are used to describe the score of question, section and the total score that defined by questioner, and the types "respquestion", "respsection", and "resptotal" are used to describe the score of question, section and the total score getting by respondents. An example is shown below. "Blue" is the sample answer of the question "qu1." The score of the question is "2". The score of the question that the respondent gets is "0", and the respondent also get "0" of total score.

```

<result>
  <marker>Bryan</marker>
  <question>
    <sampleanswer>Blue</sampleanswer>
    <score type="question">2</score>
    <score type="respquestion">0</score>
  </question>
  <score type="resptotal">0</score>
</result>

```

For the difference of question fields, we add a new element which is *formula*. We also add an attribute *lang* for the *formula* element. The element *formula* and its attribute *lang* are defined to describe formulas written by XML-based formula languages. An example is shown below. It is a question which contains a formula written by “MathML”.

```

<question>
  <text>
    <formula lang="MathML">
      ...
    </formula>
  </text>
</question>

```

At last, to deal with difference of authority, we add two attributes which are *writable* and *readable*. We use four numbers to describe the authority for each part of test. The first number stands for information of test, the second number stands for sample answer, the third number stands for response, the last number stands for the result. 0 stands for no, and 1 stands for yes. An example is shown below. It means that the respondent is only writable for response part, and the respondent cannot read sample answer of the test.

```

<respondent writable="0010" readable="1011" />

```

4 Use Case: An Offline E-testing Environment

4.1 Offline E-testing Environment

The offline e-testing environment is an e-testing environment to provide users with e-testing service which can execute various offline e-testing. E-testing executed offline means that all the devices in the test progress are in a closed network which cannot connect to the Internet. We developed the offline e-testing environment as a web application shown in Fig. 1.

In the offline e-testing environment, there are four kinds of devices. They are the offline e-testing server terminal, the wireless router, the client terminals, and the USB flash memory. All test data should be specified by QSL and imported into the offline e-testing server. The e-testing will be distributed through LAN and wireless router to

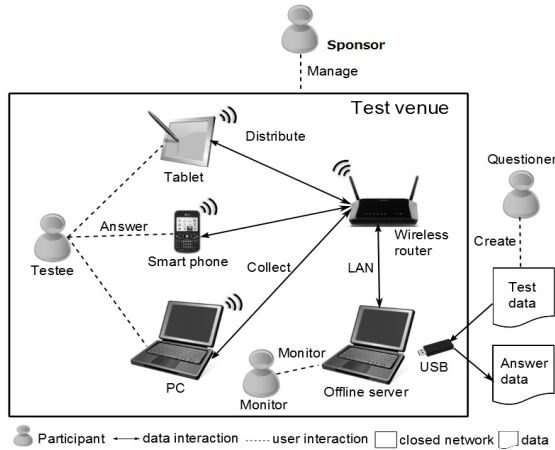


Fig. 1. Offline E-testing Environment

respondents without the Internet. The respondents will answer questions and submit their answers to offline e-testing server also as QSL format with multiple clients, such as smart phone, tablets, PC and so on. Before executing offline e-testing, an operating system and the necessary execution environment for the offline server are set into the USB flash memory as a live boot. Sticking the USB flash memory into a PC and starting the PC, the offline e-testing server can be started.

There are four kinds of participants taking part in the offline e-testing environment. They are sponsor, questioner, monitor, and respondent. The sponsor takes the responsibilities of initiating and organizing an e-testing which contains to specify the participants of the e-testing. The responsibilities of questioner are to designs questions of the e-testing and specify all the questions with QSL. Moreover, the responsibilities of monitor contain that importing test files to the offline e-testing server, managing all test that have been imported, monitoring the state of the e-testing and respondents, collecting and exporting response data of respondents. At last, the respondent takes the responsibilities of answering questions and submitting responses of the e-testing.

At present, the procedure of offline e-testing with the offline e-testing environment is as follows. A monitor imports the test file which has been described by QSL and manages all tests, the monitor monitors the connection state and test state of respondents, and respondents answer questions via web browser. The monitor orders offline e-testing server to integrate answer data automatically and the monitor exports answer files also as QSL format. Our general-purpose offline E-testing environment provides multi-language of user interface. Current environment supports English, Japanese, and Chinese. Our environment supports multi-device, smart phone, tablet, and PC all can be the client devices, and web page will adapt to the screen of device according to resolution of the device. Fig. 2 shows screenshots of the monitoring facility which uses PC as the client device (left), and the answering facility which uses smart phone as the client device (right).

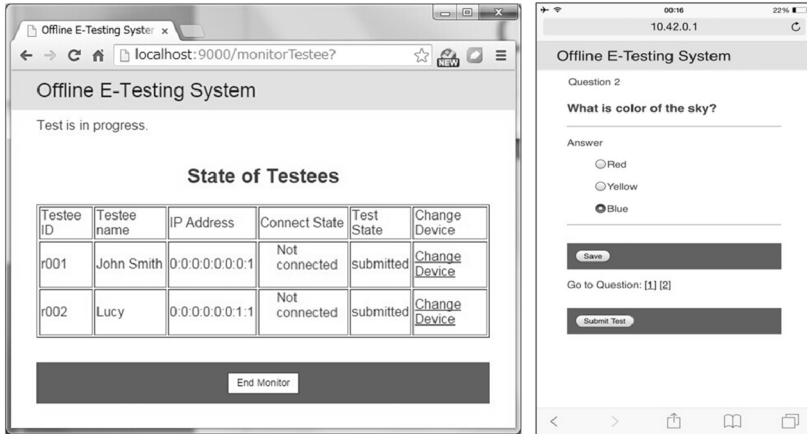


Fig. 2. Monitor Function (left) and Answer Function (right)

4.2 Specification of E-testing and E-testing System

We specified the offline e-testing environment by QSL. An e-testing which executed on the offline e-testing environment has 5 processes. They are creation, distribution, answering, monitor, and collection. All the functions are specified by QSL shown below. Offline e-testing environment distributes test data and collects answer data through wireless router. Before storing answer data into USB flash memory, all the answer data will be integrated by using zip method. For the answering process, respondents can use PC, smart phone, or tablet to answer e-testing. In the process of monitor, e-testing environment will “ping” the IP address of each respondents in order to monitor the connection state of respondents. The environment also monitors the test state of respondents. At last, the authorization of all participants is set in the security part of offline e-testing environment.

```

<qenvironment>
  <qdistribution>
    <func-distribute>
      <method type="offline">USB flash memory</method>
    </func-distribute>
  </qdistribution>
  <qcollection>
    <func-collect>
      <method type="offline">USB flash memory</method>
    </func-collect>
    <func-integrate>
      <method>zip</method>
    </func-integrate>
  </qcollection>
</qanswer>

```

```

<func-answer>
  <device>PC</device>
  <device>smart phone</device>
  <device>tablet</device>
</func-answer>
</qanswer>
<qmonitor>
  <func-monitor>
    <method type="ping">IP address</method>
    <method type="common">Test state</method>
  </func-monitor>
</qmonitor>
</qenvironment>
<qsecurity>
  <authorization/>
  <participants>
    <sponsor writable="1100" readable="1111"/>
    <questioner writable="1100" readable="1100"/>
    <monitor writable="0000" readable="1000"/>
    <respondent writable="0010" readable="1011"/>
  </participants>
</qsecurity>

```

We also specified an e-testing on the offline e-testing environment. An example is shown below. Creating an e-testing named “An E-Testing Example”. It defines the language of the e-testing which is English, it also defines the start and end time of this e-testing. The e-testing contains one section. In the section, it represents a multiple-choice question. The logic is randomization which is showing options in a random order. The score of the question is 2.

```

<qdata>
  <questionnaire name="q1">
    <text type="title">An E-Testing Example</text>
    <language>en</language>
    <time type="start">2014-12-08T09:30:00Z</time>
    <time type="end">2014-12-08T11:30:00Z</time>
    <section name="section1">
      <question name="qul" type="multiple-choice">
        ...
        <logic type="randomization" />
        <sampleanswer>Blue</sampleanswer>
        <score type="question">2</score>
      </question>
    </section>
  </questionnaire>
</qdata>

```

5 Concluding Remarks

We have proposed an extension of QSL for various e-testing and e-testing systems. We also shown an application which is an offline e-testing environment as use case of extended QSL, and shown specifications of the environment and e-testing on the environment. Moreover, we have shown that QSL is also useful format for data exchange in offline e-testing.

In the future, we will continue working on improving QSL for various e-testing and e-testing systems, and consider extending it to contain various e-voting systems owing to significant similarities of contents and structures with e-questionnaire and e-testing.

References

1. AddPoll, <http://www.addpoll.com/>
2. Constant Contact, <http://www.constantcontact.com/>
3. Datagle, <http://www.datagle.com/>
4. Diao Cha Quan (in Chinese), <http://www.diaochaquan.cn/>
5. eSurv, <http://eSurv.org/>
6. ExamSoft, <http://learn.examsoft.com/>
7. Examcoo (in Chinese), <http://www.examcoo.com/>
8. Mobo Survey, <http://www.mobosurvey.com/>
9. MySurveyLab, <http://enq-maker.com/>
10. OCR, <http://www.ocr.org.uk/>
11. oExam (in Chinese), <http://www.orivon.com/>
12. OQSS, <http://www.oqss.com/>
13. PHPEMS Online Testing System (ver. 2.0.1), <http://www.phpems.net/>
14. ProProfs, <http://www.proprofs.com/>
15. QuestionPro, <http://www.questionpro.com/>
16. Suzuki, J., Goto, Y., Cheng, J.: Development of a General-Purpose E-testing Server for Ubiquitous Test. In: Zhang, W. (ed.) Software Engineering and Knowledge Engineering. AISC, vol. 162, pp. 797–803. Springer, Heidelberg (2012)
17. So Jump (in Chinese), <http://www.sojump.com/>
18. Sogo Survey, <http://www.sogosurvey.com/>
19. SurveyMonkey, <http://www.surveymonkey.com/>
20. Tomexam Online Testing System (ver. 2.7), <http://www.tomexam.com/>
21. Yong Dao Offline Testing (in Chinese), <http://www.onlinedown.net/soft/49716.htm>
22. Zhou, Y.: QSL Manual, ver. 1.5 (2014), <http://www.aise.ics.saitama-u.ac.jp/QSL/>
23. Zhou, Y., Goto, Y., Cheng, J.: QSL: A Specification Language for E-questionnaire Systems. In: Proc. Of the 5th IEEE International Conference on Software Engineering and Service Science (ICSESS 2014), Beijing, China, pp. 224–230. IEEE (2014)

A Tasking Deadlock Detector for Ada 2012 Programs

Bo Wang, Takeo Ekiba, Yuichi Goto, and Jingde Cheng

Department of Information and Computer Sciences,
Saitama University, Saitama 338-8570, Japan
{wangbo, ekiba, gotoh, cheng}@aise.ics.saitama-u.ac.jp

Abstract. To avoid and resolve tasking deadlocks in Ada programs, it is indispensable to identify and detect all types of tasking deadlocks. Various combinations of synchronization waiting relations concerning synchronization waiting tasks may lead to various types of tasking deadlocks. As a substantial expansion of Ada 2005, Ada 2012 has many new facilities, and therefore, these changes have a great impact on Ada tasking deadlocks and their detection. Though a tasking deadlock detector for Ada 2005 and Ada 95 was developed, it cannot detect new types of tasking deadlocks in Ada 2012 programs. This paper presents a new tasking deadlock detector for Ada 2012 programs that we are developing. At first, we analyzed various types of tasking deadlocks concerning new synchronization waiting relations defined in Ada 2012. After that we designed the tasking deadlock detector, and implemented it.

Keywords: Ada 2012, tasking deadlock, synchronization waiting relation, task waiting for graph, run-time detection.

1 Introduction

Tasking deadlocks is a serious and complex issue in concurrent Ada programs [2-7, 13]. A task is said to be blocked in an execution state of a concurrent Ada program, if it is waiting at some synchronization points in its thread of control for synchronization with one or other tasks or even itself [2-7, 13]. This waiting state will be kept until that either the synchronization has occurred or the task is aborted. A tasking deadlock in a concurrent Ada program is a situation where some tasks form a circle of synchronization waiting relations at some synchronization points that cannot be resolved by the program itself (including the behaviors of other tasks), and hence can never proceed with their computation by themselves [2, 3].

To avoid and resolve Ada tasking deadlocks, it is indispensable to identify all types of tasking deadlocks. A task synchronization waiting relation between tasks is a relation such that to synchronize with the other task or tasks, a task is blocked until the synchronization takes place, unless the synchronization waiting has a deadline [2-7, 13].

Cheng proposed a way to completely classify all types of tasking deadlocks by different combinations of various synchronization waiting relations between tasking objects [2, 3]. According to this way of classification, various combinations of synchronization waiting relations concerning synchronization waiting tasks may lead to various types of

tasking deadlocks. Moreover, a method and a tool to detect tasking deadlocks at run-time are also proposed [5, 7].

Now there is the revised standard of Ada, called Ada 2012 [11], which is updated from the Ada 2005 standard [10]. As a substantial expansion of Ada 2005, Ada 2012 has many new facilities [1, 11], and therefore, these changes have a great impact on Ada tasking deadlocks and their detection. Moreover, the paper [13] has proposed some queue operation related tasking deadlocks.

This paper presents a new tasking deadlock detector for Ada 2012 programs that we are developing. At first, we analyzed various types of tasking deadlocks concerning new synchronization waiting relations defined in Ada 2012. After that we designed the tasking deadlock detector, and implemented it.

The paper consists of the following contents: Section 2 presents some various synchronization waiting relations in Ada 2012 programs; Section 3 gives an example of tasking deadlocks in an Ada 2012 program; Section 4 declares the principle of detecting tasking deadlocks; Section 5 exposes run-time detection of tasking deadlocks in Ada 2012 programs; Section 6 shows concluding remarks.

2 Synchronization Waiting Relations in Ada 2012 Programs

Ada 83, 95 defined eight types of synchronization waiting relations, i.e., activation waiting relation, finalization waiting relation, completion waiting relation, acceptance waiting relation, entry-calling waiting relation, protection waiting relation, protected-entry-calling waiting relation, and suspension waiting [1, 4-6, 8, 9, 13].

In Ada 2005, there is no new synchronization waiting relation, although there are some changes for the tasking and real-time facilities in Ada 2005. For instance, one major extension is the ability to combine the interface feature with the tasking model. There are also many additional predefined packages in the Real-Time Systems annex concerning matters such as scheduling and timing [10]. The changes have no effect on synchronization waiting relations which have been defined. They also do not cause new synchronization waiting relations.

Ada 2012 defined four new operations that cause a situation where a task is blocked. The three of the four operations, i.e., the enqueue waiting relation and the dequeue waiting relation defined in Annex A.18.27, which are about queues, and the barrier-release waiting relation defined in Annex D.10.1, which is about barriers, may cause the new synchronization waiting relations. The other, i.e., procedure **Suspend_Until_True_And_Set_Deadlines** defined in Annex D.10, does not cause new synchronization waiting relation. Enqueue waiting [6] is a state that a task calling Enqueue is blocked when a queue in Synchronized Queue Interfaces implemented is full until another task calls Dequeue for this queue. If a queue is empty, then a procedure call for Dequeue will result in that the task calling Dequeue is blocked until an item becomes available. Dequeue waiting [6] is a state that a task called Dequeue is blocked, when and only when a queue is empty until another task calls Enqueue for this queue in implemented Synchronized Queue Interfaces. Enqueue waiting may be resolved by other task calling Dequeue. If no task has the possibility to resolve enqueue waiting, a task in enqueue waiting state will be blocked forever and may be

involved with a tasking deadlock. Dequeue waiting may be resolved by other task calling Enqueue. If no task has the possibility to resolve dequeue waiting, a task in dequeue waiting state will be blocked forever and may be involved with a tasking deadlock. Barrier-release waiting [6] is a state that tasks calling `Wait_For_Release` are blocked by other possible tasks calling it, until the number of blocked tasks associated with the Synchronous Barrier object is equal to threshold. If the needed number of all tasks those plan to call `Wait_For_Release` reaches to threshold, they can resolve barrier-release waiting. If any tasks cannot have possibility to call `Wait_For_Release`, a task in barrier-release waiting state will be blocked forever and may be involved with a tasking deadlock.

Also, there are some queue operation related tasking deadlocks in [13] including a requeue operation. Though, the Requeue is not synchronization waiting relation, it can cause tasking deadlocks indirectly since the days of Ada 95.

3 An Example of Tasking Deadlocks in an Ada 2012 Program

3.1 Task-Wait-For Graph

A Task-Wait-For Graph (TWFG for short) is a kind of arc-classified digraph to represent task waiting state in an execution of an Ada program [3]. The TWFG explicitly represents various types of synchronization waiting relations at the time of an Ada program execution. In a TWFG, vertices indicate tasking objects and arcs indicate synchronization waiting relations which are binary relations between tasking objects [3].

In a TWFG of an Ada 2012 program, they are 11 types of arcs [2-7]: activation waiting arc, finalization waiting arc, completion waiting arc, acceptance waiting arc, entry calling waiting arc, protection waiting arc, protected-entry calling waiting arc, suspension waiting arc, enqueue waiting arc, dequeue waiting arc, and barrier-release waiting arc, respectively, corresponding to an activation waiting relation, finalization waiting relation, completion waiting relation, acceptance waiting relation, entry-calling waiting relation, protection waiting relation, protected-entry-calling waiting relation, suspension waiting relation, enqueue waiting relation, dequeue waiting relation, and barrier-release waiting relation. Here the affixing characters show types of arcs. \rightarrow Act, \rightarrow Fin, \rightarrow Com, , \rightarrow EC, \rightarrow Pro, \rightarrow Pec, \rightarrow Sus, \rightarrow Enq, \rightarrow Deq, and \rightarrow BR denote above, respectively.

3.2 Example Program: Tasking Deadlocks with All Synchronization Waiting Relations

This program has 11 types of synchronization waiting relations, which form five cycles [6], which may cause five tasking deadlocks in this program. The five cycles of synchronization waiting relations in this program are below:

$$\begin{aligned}
& T_1 \rightarrow Acc T_4 \rightarrow P E C V \rightarrow F i n G E T 1 \rightarrow E C T_2 \rightarrow C o m T \\
& T_1 \rightarrow Acc T_5 \rightarrow F i n B \rightarrow F i n T_7 \rightarrow S u s T_2 \rightarrow C o m T_1 \\
T_1 \rightarrow Acc T_5 \rightarrow F i n B \rightarrow F i n T_8 \rightarrow D e q T_6 \rightarrow F i n T_9 \rightarrow F i n G E T 2 \rightarrow B R T_2 \rightarrow C o m T_1 \\
& T_1 \rightarrow Acc T_5 \rightarrow F i n B \rightarrow F i n W \rightarrow P r o V \rightarrow F i n G E T 1 \rightarrow E C T_2 \rightarrow C o m T \\
& T_4 \rightarrow P E C V \rightarrow F i n G E T 1 \rightarrow E C T_2 \rightarrow C o m T_3 \rightarrow E n q T_4
\end{aligned}$$

```

pragma Task_Dispatching_Policy
(Fifo_Within_Priorities);
with System;
with Ada.Containers.Synchronized_Queue_Interfaces;
with Ada.Containers.Bounded_Synchronized_Queues;
with Ada.Synchronous_Task_Control;
with Ada.Synchronous_Barriers;
with Ada.Text_IO;
use Ada.Synchronous_Task_Control;
use Ada.Synchronous_Barriers;
use Ada.Containers;
use Ada.Text_IO;
procedure d14 is
  type Queue_Element is new String(1..0);
  package String_Queues
    is new Synchronized_Queue_Interfaces
      (Element_Type => Queue_Element);
  package String_Priority_Queues
    is new Bounded_Synchronized_Queues
      (Queue_Interfaces
       => String_Queues.Default_Capacity =>1);
  Q1, Q2: String_Priority_Queues.Queue;
  E1 : Queue_Element;
  Number_Of_Tasks : constant := 2;
  BO : Ada.Synchronous_Barrier;
  Synchronous_Barrier(Number_Of_Tasks);
  Notif : Boolean := False;
  type ITEM is new Integer;
  task T1 is entry E1; end T1;
  task T2 is entry E2; end T2;
  task T3; task T4;
  S : Suspension_Object;
  function GET1 return ITEM is
  begin
    T2.E2;
    return 0;
  end GET1;
  function GET2 return ITEM is
  begin
    Wait_For_Release(BO, Notif);
    return 0;
  end GET2;
  protected V is
    procedure W (X : in ITEM);
    entry R (X : out ITEM);
  private
    Var : ITEM := 0;
  end V;
  protected body V is
    procedure W (X : in ITEM) is
    begin
      Var := X;
    end W;
    entry R (X : out ITEM) when True is
    begin
      X := GET1;
    end R;
  end V;

  task body T1 is
    task T5; task T6;
    task body T5 is
    begin
      B : declare
        task T7; task T8;
        task body T7 is
        begin
          Suspend_Until_True(S);
        end T7;
        task body T8 is
        begin
          Q2.Dequeue(E1);
        end T8;
        Y : ITEM;
      begin
        V.W (Y);
      end B;
    end T5;
    task body T6 is
      task T9;
      task body T9 is
        I : ITEM := GET2;
      begin
        null;
      end T9;
    begin
      Q2.Enqueue("");
    end T6;
  begin
    accept E1;
  end T1;
  task body T2 is
  begin
    select when False =>
      accept E2;
    or
      terminate;
    end select;
    Set_True (S);
    Wait_For_Release(BO, Notif);
  end T2;
  task body T3 is
  begin
    Q1.Enqueue("");
    Q1.Enqueue("");
  end T3;
  task body T4 is
    Z : ITEM;
  begin
    V.R (Z);
    Q1.Dequeue(E1);
    T1.E1;
  end T4;
  begin
    null;

```

Fig. 1. Example of Ada 2012 program

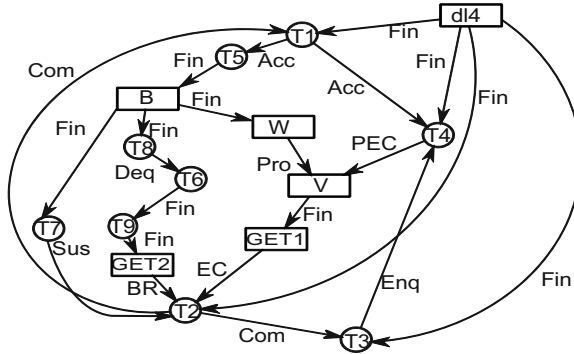


Fig. 2. TWFG of example program

4 Principle of Detecting Tasking Deadlocks

In Ada 83 there is synchronization waiting relation concern tasking deadlocks: Activation waiting, Finalization waiting, Completion waiting, Acceptance waiting, Entry-calling waiting; Ada 95 has added Protection waiting, Protected-entry-calling waiting, Suspension waiting; Ada 2012 has further extended by Enqueue waiting, which is that a task calling Enqueue is blocked when a queue in implemented Synchronized Queue Interfaces is full until another task calls Dequeue for this queue, Dequeue waiting, which is that a task calling Dequeue is blocked when a queue in implemented Synchronized Queue Interfaces is empty until another task calls Enqueue for this queue, and Barrier-release waiting, which is that tasks calling Wait For Release are blocked by other possible tasks calling it until the number of blocked tasks associated with the Synchronous Barrier object is equal to threshold. In [13], some tasking deadlocks are proposed, when some queue operations are blocked.

A task that issued a simple entry call on a protected entry is blocked until the execution of the corresponding entry body has finished. Similarly, a task that issued a timed entry call, a conditional entry call, and the corresponding protected action has started, and the execution of the corresponding entry body has started is blocked until the execution of the corresponding entry body has finished.

Note that in the above synchronization waiting relations we do not consider those selective accept statements with open delay alternatives or else part and those statements of asynchronous selects which have not yet been accepted because a task reaching any such selective accept or entry call can change its own waiting state by itself. As a result, all of above synchronization waiting relations have a common property. Therefore, a circular synchronization waiting relation formed among some tasks implies that a tasking deadlock might have occurred there.

As we described the TWFG in previous section, taking TWFGs as a formal representation for the waiting state of task synchronization, a run-time detector for tasking deadlocks can work by monitoring the tasking behavior of the program, managing a TWFG for the program, detecting cycles in the TWFG, and reporting detected tasking deadlocks.

5 Run-Time Detection of Tasking Deadlocks in Ada 2012 Programs

There are alternative approaches to monitor the tasking behavior of Ada programs. One is the source program transformation approach. In the approach, a target concurrent Ada program P is transformed by a preprocessor into another Ada program P' , called subject program of P , such that P' preserves the tasking behavior of P and during its execution P' will communicate with a run-time monitor when each tasking event of P occurs in P' and pass information about the tasking event to the run-time monitor. The other is the run-time environment support approach. In the Approach, a run-time monitor is implemented as a component of the underlying run-time support environment of Ada language and information concerning tasking events in a target concurrent Ada program is provided directly by the run-time monitor.

We implemented the Ada 2012 tasking deadlock detector in the source program transformation approach to monitor tasking behavior of Ada 2012 programs. Our tool is in two separate components, i.e., a preprocessor and a run-time monitor. The preprocessor transforms a target Ada program P into its subject program P' such that some Ada codes are inserted at each point where a tasking event occurs for passing information about the tasking event to the runtime monitor. To monitor tasking behavior of P , P' is compiled, linked with the separately compiled run-time monitor, and then executed. During its execution, P' communicates with the run-time monitor when a tasking event of P occurs in P' and pass information about the tasking event to the runtime monitor. The run-time monitor records the collected information, manages the TWFG of the target program, detects cycles in the TWFG when it is updated, and reports detected tasking deadlocks, if any.

In monitoring a concurrent program, since its behavior is generally run-time dependent, there is an “uncertainty principle,” i.e., any monitoring mechanism will interfere with the performance of the program and even may alter its behavior. Therefore, an important issue in implementing an execution monitor for concurrent programs is how to devise mechanisms to reduce the interference and alteration imposed on the behavior of target programs by monitoring actions. The authors have proposed a notion, named “partial order transparency,” for monitoring concurrent systems. It is a minimum requirement for an execution monitor in order to reduce its interference and alternation to the behavior of a target system being monitored. An execution monitor satisfying the requirement is transparent to the partial order with respect to occurrences of concurrent events during an execution of the target system. Our monitor of Ada 2012 tasking deadlock detector is partial order transparent.

We have used the ASIS [12] to develop a run-time detector for tasking deadlocks in Ada 95 programs. The ASIS is short for the Ada Semantic Interface Specification, which is an interface between an Ada environment and any tool requiring information from it, is supposed to be effective in implementing the preprocessor. An Ada compiler is regarded as a part of the system of our tasking deadlock detector and an ASIS-based tool makes use of an Ada environment which is managed by the compiler. As a future work, we should translate the current preprocessor into an ASIS-based tool in terms of the cooperation between the compiler and the tool.

We have added some codes for transforming a target Ada 2012 program P into another Ada 2012 program P' . Another Ada 2012 program P' transformed is inserted some procedure calls and function calls to the tasking information collector at each tasking information extraction point in target programs to observe tasking events.

There is the execution result of Figure 1 corresponding to the program above, such that they represent that some tasking deadlocks will occur at some run-time. Namely, all 11 types of tasking deadlocks in Ada 2012 programs can be run-time detected. We just show a snippet of results.

```
Tasking deadlock will occur in this program!
* At 0.005565337 sec after execution start,
  there is such a situation as follows :
* task Main.T2 is waiting for terminating together with
  task Main.T3
* TASKS Main.T3 is calling entry R of protected object Main.V
* PROTECTED_OBJECTS Main.V is calling subprogram Main.GET
* FUNCTIONS Main.GET is calling entry E2 of task Main.T2
* Status of all tasks are follows:
* task name      => MAIN_TASK
* state of the task => WORKING_FOR_INTERNAL_AFFAIRS
* task name      => Main.T1
* parent of the task => Main
* state of the task => ACCEPTING
* communication entry => Main.T1.E1
* task name      => Main.T1.T4
* parent of the task => Main.T1
* state of the task => WORKING_FOR_INTERNAL_AFFAIRS
* task name      => Main.T1.T5
* parent of the task => Main.T1
* state of the task => ACTIVATING
* task name      => Main.T1.T5.T8
* parent of the task => Main.T1.T5
* state of the task => SIMPLE_ENTRY_CALLING
* communication entry => Main.T2.E2
* task name      => Main.T2
* parent of the task => Main
* state of the task => TERMINATION_SELECTING
* task name      => Main.T3
* parent of the task => Main
* state of the task => SIMPLE_ENTRY_CALLING
```

6 Concluding Remarks

We present a new tasking deadlock detector for Ada 2012 programs that we are developing. We analyzed various types of tasking deadlocks concerning new synchronization waiting relations defined in Ada 2012. After that we designed the tasking deadlock detector, and implemented it.

As a future work, a method for detection of possibility of occurring cancellation at a detected deadlock should be investigated. If a local deadlock occurs in a program, it

may be cancelled by another deadlock-free task, which executes an abort statement or an Abort Task procedure in Task Identification package of the Systems Programming Annex, or by a completion of the triggering part in the asynchronous select statement. To record the dependence relations of the cancellations into the TWFG, possibility of causing a cancellation of some deadlocks can be detected. It is related with another problem, a run-time resolution mechanism for tasking deadlocks. It also should be investigated. As a detailed problem, composite types which are arrays and records, including task types are not coped with. A destination of an entry calling is changeable if the task designation is a variable, so a communication-dependent task set is not manageable for the program including such an entry calling by the present tool.

Another open problem about tasking deadlock detection in Ada programs is how to detect tasking deadlocks concerning those queue operations [13].

References

1. Barnes, J.: Programming in Ada 2012. Cambridge University Press (2014)
2. Cheng, J.: A Classification of Tasking Deadlocks. *ACM Ada Letters* 10(5), 110–127 (1990)
3. Cheng, J.: Task-Wait-For Graphs and Their Application to Handling Tasking Deadlocks. In: Proc. 3rd ACM Annual TRIAda Conference, pp. 376–390. ACM, New York (1990)
4. Cheng, J., Ushijima, K.: Tasking Deadlocks in Ada 95 Programs and Their Detection. In: Strohmeier, A. (ed.) *Ada-Europe 1996*. LNCS, vol. 1088, pp. 135–146. Springer, Heidelberg (1996)
5. Cheng, J.: Run-Time Detection of Tasking Deadlocks in Real-Time Systems with the Ada 95 Annex of Real-Time Systems. In: Pinho, L.M., González Harbour, M. (eds.) *Ada-Europe 2006*. LNCS, vol. 4006, pp. 167–178. Springer, Heidelberg (2006)
6. Ekiba, T., Goto, Y., Cheng, J.: New Types of Tasking Deadlocks in Ada 2012 Programs. *ACM Ada Letters* 33(1), 169–179 (2013)
7. Nonaka, Y., Cheng, J., Ushijima, K.: A Tasking Deadlock Detector for Ada 95 Programs. *Ada User Journal* 20(1), 79–92 (1999)
8. ISO/IEC: ISO/IEC 8652:1987 (E): Information Technology - Programming Language - Ada (1987)
9. ISO/IEC: ISO/IEC 8652:1995 (E): Information Technology - Programming Language - Ada (1995)
10. ISO/IEC: ISO/IEC 8652:2007 (E), Ed. 3: Information Technology - Programming Language - Ada (2006)
11. ISO/IEC: ISO/IEC 8652:2012 (E): Information Technology - Programming Language - Ada (2012)
12. ISO/IEC: ISO/IEC 15291:1999 (E): Information Technology - Programming Language - Ada Semantic Interface Specification (ASIS) (2013)
13. Wang, B., Goto, Y., Cheng, J.: Queue Operation Related Tasking Deadlocks in Ada 2012 Programs. *ACM Ada Letters* 34(2), 9–25 (2014)

Automated Theorem Finding by Forward Reasoning Based on Strong Relevant Logic: A Case Study in Graph Theory

Hongbiao Gao, Yuichi Goto, and Jingde Cheng

Department of Information and Computer Sciences,
Saitama University, Saitama 338-8570, Japan
{gaohongbiao,gotoh,cheng}@aise.ics.saitama-u.ac.jp

Abstract. The problem of automated theorem finding is one of 33 basic research problems in automated reasoning which was originally proposed by Wos. The problem is still an open problem until now. To solve the problem, a systematic methodology with forward reasoning based on strong relevant logic has been proposed. This paper presents a case study of automated theorem finding in graph theory to show the generality of the methodology, and presents a future direction for automated theorem finding based on the methodology.

Keywords: Automated theorem finding, forward reasoning, strong relevant logic, graph theory.

1 Introduction

The problem of automated theorem finding (ATF for short) is one of 33 basic research problems in automated reasoning which was originally proposed by Wos in 1988 [14,15], and it is still an open problem [6, 9, 10, 12-15]. The ATF problem [14, 15]: “What properties can be identified to permit an automated reasoning program to find new and interesting theorems, as opposed to proving conjectured theorems?”

The most important and difficult requirement of the problem is that, in contrast to prove conjectured theorems supplied by the user, it asks for criteria that an automated reasoning program can use to find some theorems in a field that must be evaluated by theorists of the field as new and interesting theorems. The significance of solving the problem is obvious because an automated reasoning program satisfying the requirement can provide great assistance for scientists in various fields [2].

To solve the ATF problem, a forward reasoning approach based on strong relevant logic [1, 2] and its systematic methodology [8] have been proposed. To verify the effectiveness of the approach, we used the methodology to perform a case study of ATF in axiomatic set theory, and the result of that case study shows that the forward reasoning approach based on strong relevant logic is hopeful to solve the ATF problem [8].

This paper presents a case study of automated theorem finding in graph theory [7] to show the generality of the proposed methodology, and presents a future direction for ATF based on the methodology.

2 Basic Notions and Notations

A formal logic system L is an ordered pair $(F(L), \vdash_L)$ where $F(L)$ is the set of well formed formulas of L , and \vdash_L is the consequence relation of L such that for a set P of formulas and a formula C , $P \vdash_L C$ means that within the framework of L taking P as premises we can obtain C as a valid conclusion. $Th(L)$ is the set of logical theorems of L such that $\phi \vdash_L T$ holds for any $T \in Th(L)$. According to the representation of the consequence relation of a logic, the logic can be represented as a Hilbert style system, Gentzen sequent calculus system, Gentzen natural deduction system, and so on [3].

Let $(F(L), \vdash_L)$ be a formal logic system and $P \subseteq F(L)$ be a non-empty set of sentences. A formal theory with premises P based on L , called an L -theory with premises P and denoted by $T_L(P)$, is defined as $T_L(P) =_{df} Th(L) \cup Th^e_{\vdash_L}(P)$ where $Th^e_{\vdash_L}(P) =_{df} \{A | P \vdash_L A \text{ and } A \notin Th(L)\}$, $Th(L)$ and $Th^e_{\vdash_L}(P)$ are called the logical part and the empirical part of the formal theory, respectively, and any element of $Th^e_{\vdash_L}(P)$ is called an empirical theorem of the formal theory [3].

Based on the definition above, the problem of ATF can be said as “for any given premises P , how to construct a meaningful formal theory $T_L(P)$ and then find new and interesting theorems in $Th^e_{\vdash_L}(P)$ automatically” [3].

The notion of predicate abstract level [8] is defined as follows: (1) Let $pal(X) = k$ denote that an abstract level of a predicate X is k where k is a natural number, (2) $pal(X) = 1$ if X is the most primitive predicate in a target field, (3) $pal(X) = \max(pal(Y_1), pal(Y_2), \dots, pal(Y_n)) + 1$ if a predicate X is defined by other predicates Y_1, Y_2, \dots, Y_n in the target field where n is a natural number. A predicate X is called k -level predicate, if $pal(X) = k$. If $pal(X) < pal(Y)$, then the abstract level of predicate X is lower than Y , and Y is higher than X .

The notion of function abstract level [8] is defined as follows: (1) Let $fal(f) = k$ denote that an abstract level of a function f is k where k is a natural number, (2) $fal(f) = 1$ if f is the most primitive function in the target field, (3) $fal(f) = \max(fal(g_1), fal(g_2), \dots, fal(g_n)) + 1$ if a function f is defined by other functions g_1, g_2, \dots, g_n in the target field where n is a natural number. A function f is k -level function, if $fal(f) = k$. If $fal(f) < fal(g)$, we call the abstract level of function f is lower than g , and g is higher than f .

The notion of abstract level of a formula [8] is defined as follows: (1) $lfal(A) = (k, m)$ denotes that an abstract level of a formula A where $k = pal(A)$ and $m = fal(A)$, (2) $pal(A) = \max(pal(Q_1), pal(Q_2), \dots, pal(Q_n))$ where Q_i is a predicate and occurs in A ($1 \leq i \leq n$), or $pal(A) = 0$, if there is not any predicate in A , (3) $fal(A) = \max(fal(g_1), fal(g_2), \dots, fal(g_n))$ where g_i is a function and occurs in A ($1 \leq i \leq n$), or $fal(A) = 0$, if there is not any function in A . A formula A is (k, m) -level formula, if $lfal(A) = (k, m)$.

(k, m) -fragment of premises P , denoted by $P(k, m)$, is a set of all formulas in P that consists of only (j, n) -level formulas where m, n, j and k are natural number ($0 \leq j \leq k$ and $0 \leq n \leq m$) [8].

3 A Systematic Methodology for ATF with Forward Reasoning

The systematic methodology for ATF consists of five phases [8]. Phase 1 is to prepare logical fragments [5] of strong relevant logic for various empirical theories. The prepared logic fragments are independent from any target field, therefore they can be reused for ATF in different fields. Phase 2 is to prepare empirical premises of the target theory and draw up a plan to use collected empirical theorems. In detail, we prepare (k, m) -fragment of collected empirical premises in the target field to define a semi-lattice. A set of the prepared fragments and inclusion relation on the set is a partial order set, and is a finite semi-lattice. Moreover, a set of formal theories with the fragments and inclusion relation on the set is also a partial order set, and is also a finite semi-lattice. Partial order of the set of the prepared fragments can be used for a plan to reason out fragments of formal theories with collected empirical premises. According to the partial order, we can systematically do ATF from simple theorems to complex theorems.

Phase 3 to phase 5 are performed repeatedly until deducing all fragments of formal theory that have been planned in phase 2. In this methodology, *loop* means doing phase 3 to phase 5 at once. We use one of (k, m) -fragment of collected empirical premises to perform phase 3 to phase 5 in one loop. In detail, we firstly use lowest level fragment of premises to reason out empirical theorems. Then, we enter into phase 4 to abstract deduced empirical theorems, and then enter into phase 5 to find new and interesting theorems from empirical theorems. After that, we go back to the phase 3, and use the next level fragment of premises and theorems obtained in last loop as premises of this loop. Then, we enter into phase 4 and phase 5 to abstract theorems and find interesting theorems again. We repeat the loops until all of the (k, m) -fragment of collected empirical premises have been used.

4 Case Study of ATF in Graph Theory

The purpose of the case study is to show the generality of our methodology. We chose graph theory as the field of ATF in the case study, because graph theory can be established above axiomatic set theory. We have performed a case study of ATF [8] in axiomatic set theory by using the proposed methodology. If we can also perform ATF in graph theory by using the methodology, it means that we can also do ATF in other mathematical fields by using our methodology, because almost all of mathematical fields can be established above axiomatic set theory.

We performed the case study according to our methodology. Phase 1 is to prepare logical fragments of strong relevant logic. We prepared logical fragments in the case study of axiomatic set theory [8] and those logic fragments can be reused in the case study. Phase 2 is to prepare the empirical premises of graph theory. Diestel [7] recorded the definitions of graph theory in his book. In the case study, we chose 21 definitions in Diestel's book as empirical premises of the case study and formalized them based on the predicates and functions of NBG set theory [11].

The definitions of graph theory formalized by us are shown as follows.

1. Definition of *graph*

$$G(V, E) = \langle V, E \rangle$$

2. Definition of *empty graph*

$$0 = G(0, 0)$$

3. Definition of *incident edge*

$$\forall x \forall y \forall v ((\{x, y\} \in E) \wedge (v \in \{x, y\}) \Leftrightarrow (\{x, y\} = \text{incidentedge}(v)))$$

4. Definition of *loop*

$$\forall x (\text{loop}(x) = \{x\})$$

5. Definition of *isomorphism*

$$\forall w \forall v \forall v' \forall e \forall e' ((G(v, e) = G(v', e')) \Leftrightarrow (w \in e \Rightarrow \text{isomorphism}(w) \in e'))$$

6. Definition of \cap_g

$$\forall x \forall x' \forall y \forall y' (G(x, y) \cap_g G(x', y') = G(x \cap x', y \cap y'))$$

7. Definition of \cup_g

$$\forall x \forall x' \forall y \forall y' (G(x, y) \cup_g G(x', y') = G(x \cup x', y \cup y'))$$

8. Definition of *subgraph*

$$\forall x \forall x' \forall y \forall y' (\text{Sub}(G(x, y), G'(x', y')) \Leftrightarrow ((x \subseteq x') \wedge (y \subseteq y') \wedge (x' \subseteq V) \wedge (y' \subseteq E)))$$

9. Definition of *induced subgraph*

$$\forall v \forall x \forall x' \forall y \forall y' (\text{InducedSub}(G(x, y), G'(x', y')) \Leftrightarrow \text{Sub}(G(x, y), G'(x', y')) \wedge ((v \in (x \cap x')) \Rightarrow (\text{incidentedge}(v) \in y)))$$

10. Definition of *super graph*

$$\forall x \forall x' \forall y \forall y' (\text{Sup}(G'(x', y'), G(x, y)) \Leftrightarrow \text{Sub}(G(x, y), G'(x', y'))$$

11. Definition of *simple graph*

$$\forall x \forall m \forall n \forall v \forall e (\text{SimpleGraph}(G(v, e)) \Leftrightarrow (v \subseteq V) \wedge (e \subseteq E) \wedge ((x \in v) \Rightarrow \neg(\text{loop}(x) \in e)) \wedge ((m \in e) \wedge (n \in e) \Rightarrow \neg(m = n)))$$

12. Definition of *adjacent*

$$\forall x \forall y (\text{Adjacent}(x, y) \Leftrightarrow (\{x, y\} \in E))$$

13. Definition of *complete graph*

$$\forall x \forall y \forall v \forall e (\text{CompleteGraph}(G(v, e)) \Leftrightarrow \text{SimpleGraph}(G(v, e)) \wedge ((x \in v) \wedge (y \in v) \Rightarrow \text{Adjacent}(x, y)))$$

14. Definition of *disjoint*

$$\forall x \forall x' \forall y \forall y' (\text{Disjoint}(G(x, y), G(x', y')) \Leftrightarrow (G(x, y) \cap_g G(x', y') = 0))$$

15. Definition of $-$

$$\forall x \forall y \forall u (G(x, y) - u = G(\sim(x \cap u), \sim(y \cap \text{incidentedge}(u))))$$

16. Definition of *connected graph*

$$\forall u \forall x \forall y \forall z \forall v \forall e (\text{ConnectedGraph}(G(v, e)) \Leftrightarrow ((G(u, x) \cup_g G(y, z) = G(v, e)) \Rightarrow \neg \text{Disjoint}(G(u, x), G(y, z))))$$

17. Definition of *path*

$$\forall e \forall v \forall x \forall m \forall n \forall p ((\text{path}(v, e) = G(v, e)) \Leftrightarrow (\text{ConnectedGraph}(G(v, e)) \wedge ((m \in e) \wedge (n \in e) \wedge (p \in e) \wedge \neg(m = n) \wedge \neg(n = p) \wedge \neg(m = p) \wedge (x \in m) \wedge (x \in n)) \Rightarrow \neg(x \in p))))$$

18. Definition of *cycle*

$$\forall v \forall e \forall x \forall m \exists n ((\text{cycle}(v, e) = G(v, e)) \Leftrightarrow ((\text{path}(v, e) = G(v, e)) \wedge ((m \in e) \wedge (x \in m) \Rightarrow (x \in n) \wedge (n \in e) \wedge \neg(m=n))))$$

19. Definition of *connectivity*

$$\forall x \forall y \forall e \forall v (\text{Connect}(x, y) \Leftrightarrow ((x \in v) \wedge (y \in v) \Rightarrow (\{v\} \in \text{path}(v, e))))$$

20. Definition of *forest*

$$\forall v \forall v' \forall e \forall e' (\text{Forest}(G(v, e)) \Leftrightarrow \neg \text{Sub}(\text{cycle}(v', e'), G(v, e)))$$

21. Definition of *tree*

$$\forall v \forall e (\text{Tree}(G(v, e)) \Leftrightarrow \text{ConnectedGraph}(G(v, e)) \wedge \text{Forest}(G(v, e)))$$

Then, we defined a semi-lattice of abstract level fragments of formalized empirical premises according to the proposed methodology. In detail, first we summarized all of the predicate abstract levels in the formalized definitions: 2-level predicate: *Adjacent* (abstract from \in), *Connect* (from \in); 3-level predicate: *Sub* (from \subseteq); 4-level predicate: *Forest* (from *Sub*), *InducedSub* (from *Sub*, \in), *Sup* (from *Sub*), *SimpleGraph* (from \in , \subseteq , $=$), *Disjoint* (from $=$); 5-level predicate: *CompleteGraph* (from *SimpleGraph*, \in , *Adjacent*), *ConnectedGraph* (from $=$, *Disjoint*); 6-level predicate: *Tree* (from *ConnectedGraph*, *Forest*). Second, we summarized all of the function abstract levels in the formalized definition: 2-level function: *incidentedge* (abstract from *unordered pair*); 3-level function: *loop* (from *singleton*); 4-level function: *G* (from *ordered pair*); 5-level function: *path* (from *G*), \cap_g (from \cap , *G*), \cup_g (from \cup , *G*), *isomorphism* (from *G*), $-$ (from *G*, \sim , \cap , *incidentedge*); 6-level function: *cycle* (from *G*, *path*). Third, we summarized all of the abstract levels of formalized definitions as shown in Table 1. Finally, we defined the semi-lattice of abstract level fragments of empirical premises in graph theory as shown in Fig. 1 by using the methodology. In this case study NBG set theory is seen as the minimum element of the semi-lattice, which is different from the last case study of axiomatic set theory [8].

Table 1. The abstract level of definitions in graph theory

Abstract Level	Definition
(2, 1)	Definition of <i>adjacent</i>
(2, 5)	Definition of <i>connectivity</i>
(3, 2)	Definition of <i>incident edge</i>
(3, 3)	Definition of <i>loop</i>
(3, 4)	Definition of <i>graph, empty graph, subgraph</i>
(3, 5)	Definition of <i>isomorphism</i> , \cap_g , \cup_g , $-$
(3, 6)	Definition of <i>cycle</i>
(4, 4)	Definition of <i>super graph, simple graph, induced subgraph</i>
(4, 5)	Definition of <i>disjoint</i>
(4, 6)	Definition of <i>forest</i>
(5, 4)	Definition of <i>complete graph</i>
(5, 5)	Definition of <i>path, connected graph</i>
(6, 4)	Definition of <i>tree</i>

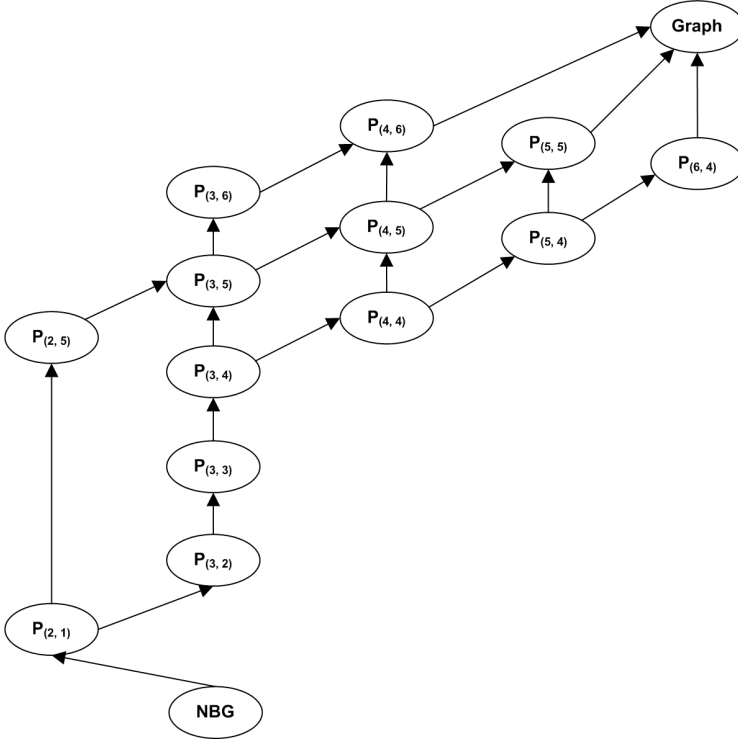


Fig. 1. The defined semi-lattice of graph theory based on NBG set theory

From Phase 3 to Phase 5, we performed deduction, abstraction, and finding of theorems. We used the general reasoning engine FreeEnCal [5] as the tool. In detail, we used the prepared logic fragments [8] from small to large according to the defined semi-lattice of strong relevant logic [8] as logic premises, and used (k, m) -fragments of premises of graph theory from lower abstract level to higher abstract level as empirical premises, we also used obtained empirical theorems of NBG set theory in last case study [8] as empirical premises to perform ATF. Then, we used filtering method to remove uninteresting theorems. We showed the results of the case study in Table 2.

The case study shows that our methodology holds generality. By using our methodology, we defined the (k, m) -fragments of empirical premises of axiomatic set theory and extended them to graph theory well. Besides, it is sure that the case study of ATF in graph theory were performed systematically in each phase. All of those empirical theorems are obtained by using forward reasoning method and our filtering method can filter most of uninteresting theorems based on syntax, more than 90% empirical theorems reasoned out by all of the five prepared logical fragments can be removed as uninteresting theorems automatically such that the scientists can find interesting theorems from the filtered results based on semantics by acceptable time.

Table 2. The results of the case study

Used logic fragments	Obtained empirical theorems	Core empirical theorems	Filtered results
$Th^{(\Rightarrow, 2)}(EeQ)$	1138	102	83
$Th^{(\Rightarrow, 3)}(EeQ)$	1662	133	93
$Th^{(\Rightarrow, 2, \neg, 1)}(EenQ)$	1139	103	84
$Th^{(\Rightarrow, 3, \neg, 1)}(EenQ)$	2885	288	216
$Th^{(\Rightarrow, 2, \neg, 1, \wedge, 1)}(EcQ)$	1216	122	97

5 A Future Direction for ATF

Cheng has proposed a semi-lattice model of formal theories [4], which can support forward reasoning approach based on strong relevant logic to do ATF in multi-fields. The core of the model is strong relevant logic (SRL) and axiomatic set theory is seen as the minimum element in the semi-lattice, and other formal theories can be established above the axiomatic set theory as shown in Fig. 2.

To do ATF based on Cheng’s semi-lattice model [4] of formal theories by using our methodology is a future direction for ATF, because it holds generality. Our methodology can support Cheng’s semi-lattice model and provide a method to establish the semi-lattice of formal theories, that is we define the (k, m) -fragment of empirical premises of axiomatic set theory and we extend them to other mathematical fields. The case study of graph theory which we have presented in Section 4 shows our methodology is effective to support Cheng’s semi-lattice model of formal theories. We consider finding process of theorems must include some concept/notion abstraction processes in other mathematical fields, so we can conclude that our methodology is suitable for ATF in other mathematical fields, such as group theory, lattice theory and number theory.

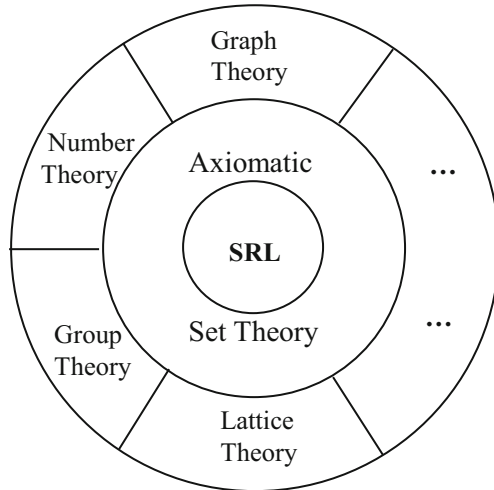


Fig. 2. The semi-lattice model of formal theories

6 Concluding Remarks

We have presented a case study of ATF in graph theory, and showed the generality of our proposed methodology through the case study. We have shown a future direction for ATF through the case study. We will do ATF in other fields like lattice theory, group theory, and number theory in future according to Cheng's semi-lattice model of formal theories.

References

1. Cheng, J.: A Relevant Logic Approach to Automated Theorem Finding. In: The Workshop on Automated Theorem Proving attached to International Symposium on Fifth Generation Computer Systems, pp. 8–15 (1994)
2. Cheng, J.: Entailment Calculus as the Logical Basis of Automated Theorem Finding in Scientific Discovery. In: Systematic Methods of Scientific Discovery: Papers from the 1995 Spring Symposium, pp. 105–110. AAAI Press - American Association for Artificial Intelligence (1995)
3. Cheng, J.: A Strong Relevant Logic Model of Epistemic Processes in Scientific Discovery. In: Information Modelling and Knowledge Bases XI. Frontiers in Artificial Intelligence and Applications, pp. 136–159. IOS Press (2000)
4. Cheng, J.: A Semilattice Model for the Theory Grid. In: Proc. 3rd International Conference on Semantics, Knowledge and Grid, pp. 152–157. IEEE Computer Society (2007)
5. Cheng, J., Nara, S., Goto, Y.: FreeEnCal: A Forward Reasoning Engine with General-Purpose. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part II. LNCS (LNAI), vol. 4693, pp. 444–452. Springer, Heidelberg (2007)
6. Colton, S., Meier, A., Sorge, V., McCasland, R.: Automatic Generation of Classification Theorems for Finite Algebras. In: Basin, D., Rusinowitch, M. (eds.) IJCAR 2004. LNCS (LNAI), vol. 3097, pp. 400–414. Springer, Heidelberg (2004)
7. Diestel, R.: Graph Theory. Springer, Heidelberg (2000)
8. Gao, H., Goto, Y., Cheng, J.: A Systematic Methodology for Automated Theorem Finding. Theoretical Computer Science 554, 2–21 (2014)
9. Gao, H., Goto, Y., Cheng, J.: Research on Automated Theorem Finding: Current State and Future Directions. In: Park, J.J.(J.H.), Pan, Y., Kim, C.-S., Yang, Y. (eds.) Future Information Technology. LNEE, vol. 309, pp. 105–110. Springer, Heidelberg (2014)
10. McCasland, R., Bundy, A., Autexier, S.: Automated Discovery of Inductive Theorems. Journal of Studies in Logic, Grammar and Rhetoric 10(23), 135–149 (2007)
11. Quaife, A.: Automated Development of Fundamental Mathematical Theories. Kluwer Academic (1992)
12. Recio, T., Velez, M.Z.: Automatic Discovery of Theorems in Elementary Geometry. Journal of Automated Reasoning 23(1), 63–82 (1999)
13. Tang, P., Lin, F.: Discovering Theorems in Game Theory: Two-Person Games with Unique Pure Nash Equilibrium Payoffs. Artificial Intelligence 175(14), 2010–2020 (2011)
14. Wos, L.: Automated Reasoning: 33 Basic Research Problem. Prentic-Hall (1988)
15. Wos, L.: The Problem of Automated Theorem Finding. Journal of Automated Reasoning 10(1), 137–138 (1993)

A Middleware Supporting Query Processing on Distributed CUBRID

Hyeong-Il Kim, Min Yoon, YoungSung Shin, and Jae-Woo Chang*

Chonbuk National University, Korea
{melipion, myoon, twotoma, jwchang}@jbnu.ac.kr

Abstract. Due to the shortages of NoSQL, studies on RDBMS based bigdata processing have been actively performed. Although they can store data in the distributed servers by dividing the database, they cannot process a query when data of a user is distributed on the multiple servers. Therefore, in this paper we propose a CUBRID based middleware supporting distributed parallel query processing. Through the performance evaluations, we show that our proposed scheme outperforms the existing work in terms of query processing time.

Keywords: Middleware, distributed parallel query processing, CUBRID.

1 Introduction

Recently, studies on the bigdata processing have been actively performed [1], [2]. With the existing IT technologies, it is very hard to efficiently store, process and analyze the bigdata. The bigdata itself is hard to be used as valuable information because of the immense volume of the bigdata. Therefore, it is necessary to analyze the bigdata to extract the meaningful information. To analyze the bigdata, a large scale of computing resources and efficient bigdata management system are required. For this, studies on NoSQL have been done [3-7]. However, NoSQL cannot satisfy the ACID properties of database transactions. Therefore, bigdata processing based on RDBMS (Relational DataBase Management System) has been spotlighted.

CUBRID Shard [8] is a RDBMS that is designed to process bigdata. To support parallel query processing, CUBRID Shard stores data in the distributed CUBRID servers by dividing the database. However, if data of a user is distributed on the multiple CUBRID servers, CUBRID Shard cannot process the query. Moreover, CUBRID Shard has a low usability because a user should specify a '*shard_hint*' in the SQL when requesting the query.

To solve these problems, in this paper we propose a CUBRID based middleware which supports distributed parallel query processing. Through our proposed middleware, users who are familiar with SQL can conveniently process the bigdata by using SQL statements. In addition, the middleware can support the aggregation queries that have not been handled on the distributed parallel computing environment.

* Corresponding author.

The rest of this paper is organized as follows. In section 2, we briefly review related work. Section 3 explains the propose middleware in detail. An empirical evaluation is presented in Section 4. Finally, we conclude this paper in section 5.

2 Related Work

NoSQL systems are increasingly used in bigdata and real-time web applications. NoSQL such as Hadoop [3], MongoDB [4], and Cassandra [5] provides a mechanism for storage and retrieval of unstructured data. The data structures used by NoSQL differ from those used in relational databases, making some operations faster in NoSQL. However, most NoSQL cannot satisfy the ACID properties of the database transactions. Especially, the major shortcoming of NoSQL is that it cannot guarantee data consistency when NoSQL supports the partition tolerance and availability.

Therefore, RDBMS have been spotlighted in the field of bigdata processing. CUBRID [9] is an object-oriented RDBMS developed by NHN (Next Human Networks). CUBRID provides predictable automatic fail-over and fail-back features based on a native CUBRID heartbeat technology. However, CUBRID cannot run on the distributed system environments because CUBRID is optimized on single machine. So, it is not efficient for dealing with bigdata. To solve the problems of CUBRID, CUBRID Shard [8] is developed. CUBRID Shard can partition the data based on the horizontal partitioning technique. CUBRID Shard allows storing a number of database shards and distributing data. With CUBRID Shard, application developers do not need to modify the application logic to divide a database into CUBRID Shards because the database system automatically handles it. CUBRID Shard provides built-in distributed load balancing, connection, and statement pooling. However, CUBRID Shard cannot process a query when data of a user is distributed on the multiple CUBRID servers. It can be a big problem when dealing with the bigdata. Moreover, CUBRID Shard has a low usability because a user should specify a *'shard_hint'* in the SQL when requesting the query.

3 Middleware Based on the Distributed CUBRID

Fig. 1 shows the overall system architecture of our proposed middleware supporting parallel query processing on the distributed CUBRID. The middleware consists of 4 components.

First, a communication component is in charge of data transmission with a user or CUBRID servers. SQL query and database connection information are transmitted through the communication component. Second, a query analysis component performs an SQL parsing to extract table names in from phrase that are used for retrieving meta tables. In addition, the component distinguishes the query types. Third, metadata retrieval component retrieves meta tables. There are 3 meta tables. i) *MinMaxTable* stores information for inserting data on the distributed CUBRID servers. The schema of the table is $\{dbName, partition, tableName, column, min, max\}$. The *column* means the name of the column that is used to partition the

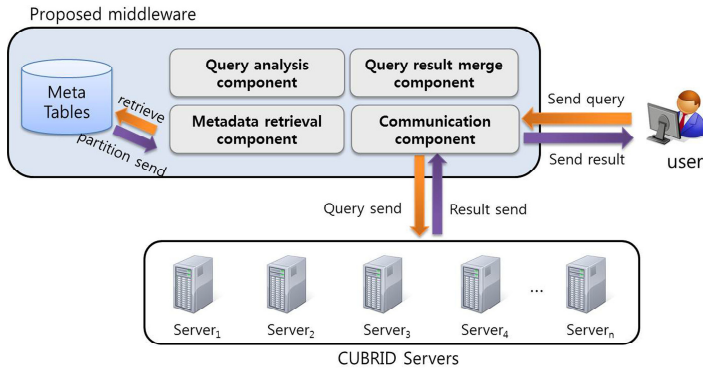


Fig. 1. The overall system architecture

tableName table. The *partition* means a CUBRID server which stores records whose values of the *column* are between *min* and *max*. ii) *SearchTable* stores information required for retrieving data that are stored on the distributed CUBRID servers. The schema of the table is $\{userID, dbName, tableName, partition\}$. By using the table, we can determine the partitions storing the *tableName* table that are necessary to process the query of the *userID*. iii) *IpPortTable* stores connection information of each CURED server. The schema of the table is $\{partition, ip, port\}$. Finally, a query result merge component merges results sent from CUBRID servers. The middleware prepares a buffer for each CUBRID server to receive each query result in parallel without any collisions. In addition, the query result merge component eliminates duplicated results and aggregates query results if needed. Finally, the query result merge component sends the final query result to the query issuer.

The overall query processing procedure with the proposed middleware is as follows. i) A user sends an SQL query to the middleware. ii) By using the query analysis component, the middleware distinguishes a type of the query. iii) The middleware reconstructs the SQL query to be processed on the distributed CUBRID servers. iv) By using the query analysis component, the middleware extracts table names in from phrase. v) By using metadata retrieval component, the middleware finds a list of CUBRID servers holding the required data to process the query. vi) The middleware generates a packet for each CUBRID server. vii) By using the communication component, the middleware sends packets to the CUBRID servers. In addition, the middleware prepares a buffer for each CUBRID server to receive query results in parallel. viii) The middleware receives a query result from each CUBRID server that processes the query. ix) By using the query result merge component, the middleware draws the final query result. x) The middleware finishes the query request by sending the final query result to the client.

Meanwhile, the middleware plays a different role according to the query type. Following describes how our proposed middleware processes each query type. First, in case of Insert phrase, the middleware stores data into the distributed CUBRID servers. To handle data insertion, data partitioning strategy of the designated table should be stored in *MinMaxTable*. By referring the table, the system can

automatically store the data into the appropriate partition. For example, for a given SQL query “Insert into *Student*(*ID*, *name*) values(20, ‘KIM’)”, the middleware can notice that the data should be inserted into the *Student* table. By referencing the *MinMaxTable*, the middleware confirms that the *Student* table is partitioned based on the *ID* column and the record with the *ID* value of 20 is related to the *partition 1*. Then, the middleware retrieves the *IpPortTable* to find the connection information of the *partition 1*. Table 2 shows an example of the *IpPortTable*. By retrieving the *IpPortTable*, the middleware finds that the *ip* and *port* of the *partition 1* are “123.456.789.001” and “9001” respectively. So, the middleware performs the data insertion by sending the SQL query to the CUBRID server (*partition 1*). Through the mechanism, the middleware achieves the distributed data insertion.

Table 1. MinMaxTable

dbName	partition	TableName	column	min	Max
db01	1	Student	ID	0	50
db01	2	Student	ID	50	100
db01	1	Graduate	ID	0	50
db01	2	Graduate	ID	50	100

Table 2. IpPortTable

partition	ip	port
1	123.456.789.001	9001
2	123.456.789.002	9002
10	123.456.789.010	9010

Table 3. SearchTable

id	dbName	TableName	Partition
user01	db01	Student	1, 2
user02	db09	Professor	1

Second, in case of Select phrase, the middleware retrieves databases in distributed manner. For this, the middleware determines which tables should be retrieved by analyzing the SQL query and retrieves *SearchTable* to find partitioning information of the tables. For example, assume that *user01* sends a query like “Select * from *Student* where *age*=21”. By analyzing the query, the middleware can notice that the *Student* table is required to process the query. When we consider the *SearchTable* shown in Table 3, the middleware can find that *Student* table of the *user01* is distributed in *partition 1* and *partition 2*. Then, the middleware accesses the *IpPortTable* to retrieve the connection information of the CUBRID servers. By sending the query to these CUBRID servers, data retrieval can be performed in parallel. Meanwhile, when processing the select query type, the middleware should consider following. The query result of each CUBRID server is sorted based on the order by conditions. If there is no order by phrase in the query, the query result of each CUBRID server is sorted based on the key value by default. So, the middleware should re-sort the query result sent from each CUBRID server based on the order by conditions to make the final query result. In addition, the middleware eliminates a duplicated record during re-sorting the query results. If there is a limit phrase in the query, the middleware terminates the query processing when the middleware writes the required number of records to the final result. Finally, the middleware completes the select query processing by sending the final query result to the client.

Third, in case of Join phrase, the middleware can process the query when the following criteria are satisfied. i) *MinMaxTable* should store the partitioning strategies of the designated tables. ii) The tables should be partitioned based on the same column and should follow the same partitioning strategy. For example, assume that the middleware receives a query like “Select * from *Student*, *Graduate* where *age=21*” and the *MinMaxTable* is given as like Table 1. *Student* and *Graduate* tables use *ID* column for partitioning and their partitioning strategy is identical (e.g., *partition 1* is in charge of storing records whose *ID* values are between 0 and 50 for both tables). In this case, the middleware can perform the join operation on the tables.

Finally, in case of Aggregation phrase, the middleware finds what kinds of aggregation operations are included in the query. According to the type, the middleware operates in different way. i) When the type is *min* or *max*, the middleware receives the minimum or maximum value from each CUBRID server and sets the smallest or largest value as the final result. ii) When the type is *count* or *sum*, the middleware receives the number of records of sum from each CUBRID server and calculates the final result by adding result values. iii) When the type is *average*, it is impossible to draw the final result by using average results sent from CUBRID servers. Therefore, the middleware reconstructs the query by using *sum* and *count* instead of *average*. Then, the middleware receives the query result (*count* and *sum*) from each CUBRID server. The middleware calculates the sum of these values respectively and calculates the average value ($total\ sum / total\ count$).

4 Performance Evaluation

We compare our middleware with the existing CUBRID in terms of query processing time varying the number of data. Because CUBRID does not support parallel query processing in distributed environments, we perform query processing of CUBRID in a sequential way. We use one master node and 3 slave nodes for the performance evaluation. We use CUBRID version 2.2.0 and compile the middleware using g++ 4.6.3 running on the Linux 3.5.0-23 with Intel® Core™ i3-3240 3.40Ghz CPU and 8 GB memory. According to Wisconsin Benchmark [10], we generate a million data for select and average operations, and 10,000 data for join operation.

Fig. 2 describes query processing time for select operation. The query processing time is increased as the number of data increases. When the number of data is 60% of all data, the query processing time of our scheme and CUBRID are 7.76 and 14.49 seconds, respectively. The middleware shows about 47% better performance than CUBRID. Fig. 3 shows the query processing time for join operation. When the number of data is 60% of all data, the query processing time of our scheme and CUBRID are 0.14 and 0.35 seconds, respectively. Our proposed middleware shows about 60% better performance than CUBRID. Fig. 4 shows query processing time for average operation. When the number of data is 60% of all data, the query processing time of our scheme and CUBRID are 0.12 and 0.31 seconds. Overall, our proposed middleware outperforms the existing CUBRID because our middleware supports parallel query processing in a distributed environment. Especially, in case of join operation, the middleware shows much performance improvement because join operation requires more computations than the select operation.

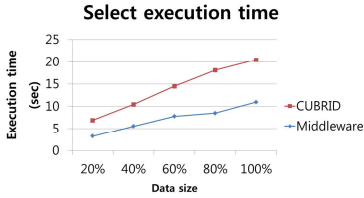


Fig. 2. Select operation performance

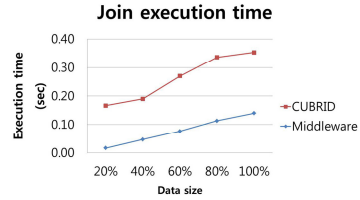


Fig. 3. Join operation performance

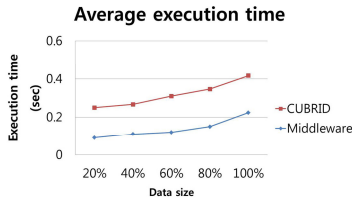


Fig. 4. Average operation performance

5 Conclusion

Existing distributed systems have problems when processing bigdata. Therefore, in this paper we propose a CUBRID based middleware which supports distributed parallel query processing. The middleware can support users who are familiar with SQL to conveniently process the bigdata by using SQL statements. In addition, the middleware can support various aggregation operators. Through the performance evaluations, we show that our proposed scheme outperforms the existing work in terms of query processing time. As a future work, we plan to expand our middleware to support various types of join with reasonable efficiency.

Acknowledgements. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2014065816).

References

1. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. *Communications of the ACM* 51(1), 107–113 (2008)
2. Rabl, T., Sadoghi, M., Jacobsen, H.: Solving Big Data Challenges for Enterprise Application Performance Management. *Vldb Endowment* 5(12), 1724–1735 (2012)
3. Apache Software Foundation, Apache Hadoop, <http://hadoop.apache.org>
4. Chodorow, K.: MongoDB: the definitive guide. O'Reilly Media Inc. (2013)

5. Dietrich, A., Mohammad, S., Zug, S., Kaiser, J.: ROS meets Cassandra: Data Management in Smart Environments with NoSQL. In: 11th International Baltic Conference (2014)
6. Shvachko, K., Kuang, H., Radia, S., Chansler, R.: The Hadoop Distributed File System. In: 26th IEEE Symposium on Mass Storage Systems and Technologies (MSST), NV (2010)
7. Han, J., Haihong, E., Guan, L.: Survey on NoSQL Database. In: 6th IEEE International Conference on Pervasive Computing and Applications, Port Elizabeth, pp. 363–366 (2011)
8. CUBRID Shard, <http://www.cubrid.org/manual/91/en/shard.html>
9. CUBRID, <http://www.cubrid.com>
10. DeWitt, D.J.: The Wisconsin Benchmark: Past, Present, and Future. In: Database and Transaction Processing System Performance Handbook (1993)

A Discovery Support Scheme for Inter-domain DDS Gateways in Cyber-Physical Systems

Wooyeob Lee¹, Sungmoon Chung¹, Sungryung Cho¹,
Inwheel Joe^{1,*}, and Jeman Park²

¹ Department of Electronics and Computer Engineering, Hanyang University, Korea

² ETRI (Electronics and Telecommunications Research Institute), Korea

{matias12, dear1115, kyougt, iwjoe}@hanyang.ac.kr,
jeman@etri.re.kr

Abstract. We propose a discovery support scheme for inter-domain DDS gateways to reduce the completion time of the discovery process. Since the conventional inter-domain services are regardless of the unbalanced EDPs load among the gateways, several problems occurs such as the delaying of completion time. Therefore, in this paper, we address the problem of completion time.

Keywords: Cyber-Physical System, CPS Middleware, Inter-domain, DDS Gateway, Fast Discovery.

1 Introduction

The CPS (Cyber-Physical Systems) is a highly reliable system which provides the real-time collaboration among the distributed heterogeneous computing devices [1]. In the cps system, all members, such as human or things, exchanges huge amount of messages. The DDS (Data Distribution Service) is a pub/sub-based real-time middleware which has advantage of exchanging messages frequently [2]. In many researches, the DDS is selected as the suited middleware for CPS because of its capability to support large system [3]. In many versions of DDS products, the inter-domain services are included and they use gateways to represent each domain [4-5] and the multicast support is not usable in the practical network [6-7]. Thus, the gateway may become a bottleneck and it may delays the completion time of the discovery process.

2 Related Works

There are some existing researches on the inter-domain service for DDS system [4-5]. The EDDS (ETRI DDS) [4] provides an inter-domain gateway based on the tunneling. All gateways are connected to each other via TCP connection and although the DDS supports a multicast, all messages are sent by unicast manner because of some

* Corresponding author.

practical reasons [6-7]. The RTI DDS provides two types of inter-domain service: bridging for local and tunneling for remote. However, as both services use the gateways, they have the same unfair bottleneck problems. Therefore we propose a discovery support scheme for EDDS inter-domain gateway to reduce completion time.

3 Discovery Support Scheme

Overall Process. Figure 1 is an example of Discovery Supporting phase. Basically, the proposed scheme is a centralized approach which uses a single server to manage the transmission of discovery messages of each gateway. The gateways of each domain send the report messages containing the status information to the central server. Then, by considering the current status of all gateways, the server makes a decision on whether to start the discovery support scheme or not. If all states are stable and it is unnecessary to control the transmission, all gateways exchange discovery messages without any control algorithm.

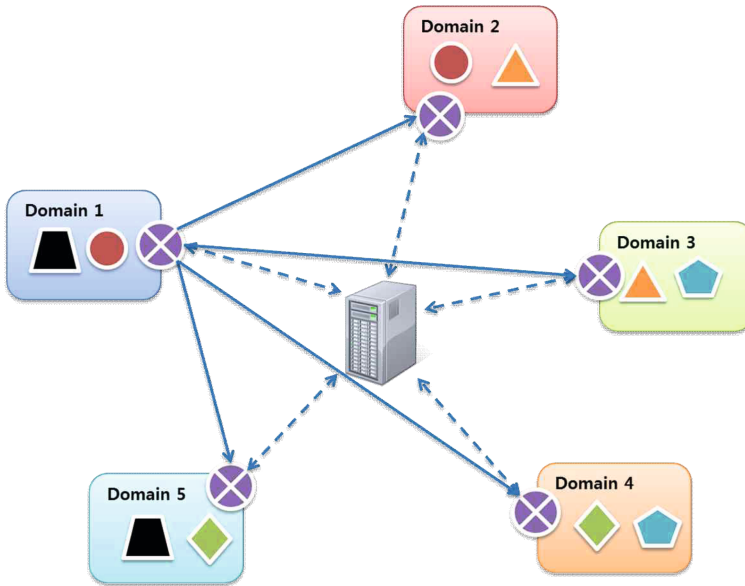


Fig. 1. An example of Discovery Supporting phase

Status Reporting. In this phase, the server monitors all gateways by receiving the report messages periodically from each gateway. The report message contains the data rate, network utilization and the number of EDPs of each gateway. While the size of total EDPs of gateways is not exceeded the data rate, there is no need for any controlling of transmission. Thus, all EDPs of each gateway will be sent directly from the source gateways. Then if the status of any gateway is changed, in other words, if the

size of total EDPs is exceeded the capacity of the data rate, the server determines whether this change is permanent or not. If the change is not temporal, the server enters the discovery supporting phase.

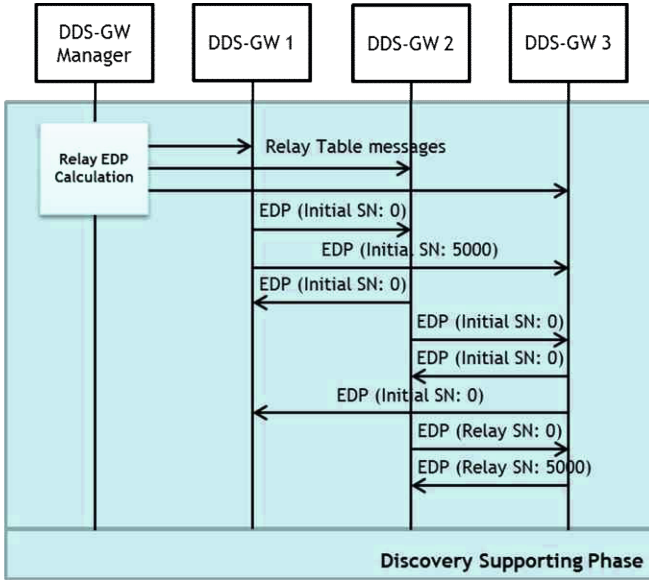


Fig. 2. An example of Discovery Supporting phase

Discovery Supporting. Fig. 2 is an example of Discovery Supporting phase. When entered this phase, the server requests the current status of all gateways, then it calculates the relative average number of EDPs to send for each gateway. Then the server reallocates the excess EDPs of overflowed gateways and creates a discovery supporting table which includes a fully ordered sequence of transmissions of all gateways. The *non-overflowed* gateway sends their EDPs firstly, and then relays the EDPs received from the *overflowed* gateways in order of discovery supporting table. Therefore, to maintain the balance of network load, the server must calculate the amount of transmission for each gateway by considering the status of each of them.

Since the completion time of the endpoint discovery process equals to the total transmission time of the EDPs of the overflowed gateways, the completion time is reduced as the gap between the amount of exceeded EDPs and the relative average number of EDPs grows more distinctive.

4 Performance Evaluation

In this chapter, we evaluate and compare our proposed discovery support scheme to the conventional method of the EDSS inter-domain gateway. Since the main idea of our proposing scheme is to share the load of the overflowed gateway among the non-overflowed gateways, we calculated the average amount of transmission and the

amount of relay transmission related to the average. The variables used for the calculation is described in the Table 1 and the equation is as follows

Table 1. Description of variables

Variables	Description
m	The number of gateways
n_T	The total number of EDPs
n_i	The number of EDPs in i th gateway
n_S	The average number of EDPs per gateway
n_R	The number of relay EDPs per gateway
DR_T	The sum of data rates of all gateways
DR_i	The data rate of i th gateway

$$n_{S_i} = \frac{n_T \cdot (m - 1) \cdot DR_i}{DR_T} \tag{1}$$

$$\begin{aligned} & \text{if } n_{S_i} \geq n_i \cdot (m - 1) && \text{if } n_{S_i} < n_i \cdot (m - 1) \\ n_{R_i} &= \frac{n_{S_i} - (n_i \cdot (m - 1))}{m - 2} && n_{R_i} = \frac{n_{S_i} - n_i}{m - 2} \end{aligned} \tag{2}$$

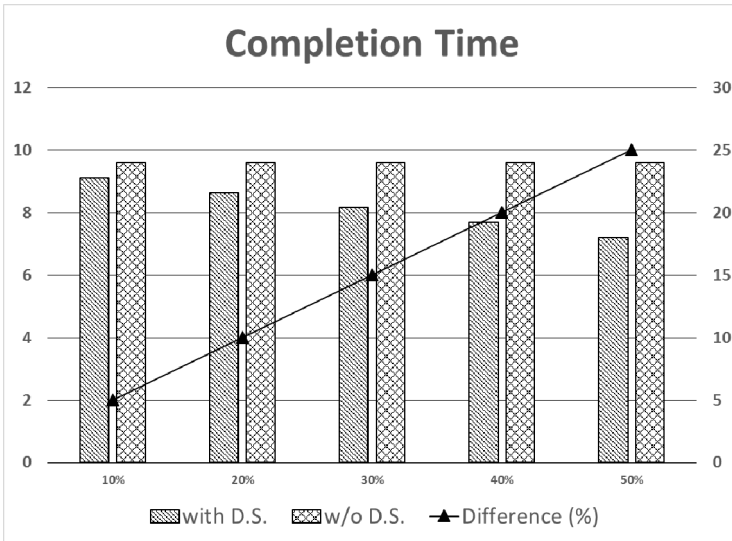


Fig. 3. Comparison of Completion Time

Fig. 3 is the graph which displays the comparison between the proposed and the conventional scheme in terms of the discovery completion time. The x-axis indicates the difference ratio between the largest number of the EDPs of the specific gateway and the smallest number of EDPs in the network. The y-axis of the bar-type represents the absolute time in seconds and that of the line-type shows the difference of time in rate. With this result, we can see that our scheme works better than the conventional EDPs inter-domain gateway and we can also estimate that the increasing rate of the number of EDPs causes the significant reduction of the completion time. The reason is that the non-overflowed operates as a relay of the overflowed, thus the completion time of the overflowed is reduced.

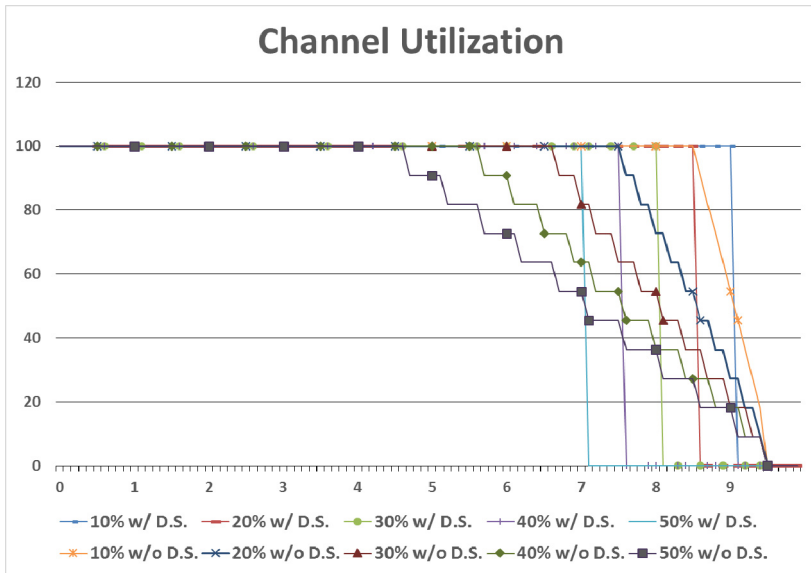


Fig. 4. Channel Utilization

Fig. 4 demonstrates the total channel utilization of the member gateways. As shown in the figure, our proposed scheme maintained the maximum utilization to the end of the discovery. While the utilization of the conventional scheme is partially dropped because of the unbalanced number of EDPs per gateway which makes the completion time late.

5 Conclusions

In this paper, we proposed the server-based discovery support scheme to reduce the completion time of the discovery process in the conventional inter-domain DDS gateway system. We managed all gateways by using the centralized server and used the non-overflowed gateways to relay the excess EDPs. We evaluated our scheme by simulation and confirmed that, in some cases, our scheme made clear reduction of completion time compared to the conventional DDS gateways.

Acknowledgments. This work was supported by the IT R&D Program of MSIP/KEIT [10035708, "The Development of CPS (Cyber-Physical Systems) Core Technologies for High Confidential Autonomic Control Software"].

References

1. Information on, <http://cyberphysicalsystems.org/>
2. OMG, Data Distribution Service for Real-time Systems Specification ver. 1.2 (2007)
3. Park, J., Lee, W., Chung, S., Joe, I., Kim, W.: A fast and scalable service discovery protocol for CPS-based warships. In: Proceedings of IEEE International Conference on Innovative Computing and Communication (CICC 2011), vol. 2, pp. 335–338 (2011)
4. Lee, W., Chung, S., Choi, M., Cho, S., Joe, I., Park, J., Lee, S., Kim, W.: A Robust Inter-Domain DDS Gateway based on Token Passing for Large-Scale Cyber-Physical Systems. In: Proceedings of ICACT 2014, pp. 868–871 (2014)
5. Lopez-Vega, J.M., Povedano-Molina, J., Pardo-Castellote, G., Lopez-Soler, J.M.: A content-aware bridging service for publish/subscribe environments. *The Journal of Systems and Software* 86, 108–124 (2012)
6. Fahmy, S., Kwon, M.: Characterizing Overlay Multicast Networks and Their Costs. *IEEE/ACM Transactions on Networking* 15(2) (2007)
7. Understanding IP Multicasting White Paper, IndigoVision (2008)

Robust Facial States Estimation against Occlusion and Inference of Driver's Drowsiness Using Hidden Markov Model*

In-Ho Choi, Chan-Hee Jeong, and Yong-Guk Kim**

Dept. of Computer Eng., Sejong University, Seoul, Korea
ykim@sejong.ac.kr

Abstract. In tracking facial states of the driver, one of the obstacles is visual occlusion occurred sporadically by his hand, car handle or other objects in the cars. We propose a new system in which the face detector not only detects the face but also handles the occlusion well. Then the detected face is used for tracking the driver's 3D head pose in real-time. A hidden Markov model is adopted in inferring driver's drowsiness. Performance evaluation using the standard and our DB suggests that the system has a commercial potential.

Keywords: Face Detection, Occlusion, Head Pose Estimation, Driving, Drowsiness, Hidden Markov Model.

1 Introduction

Drowsy driving refers to when a driver is half-sleeping after a long period of driving or sleep deprivation. Currently, more than 30% of deaths caused by car accidents are attributed to drowsy driving. In 2008, National Highway Traffic Safety Administration (NHTSA) estimated that 100,000 police reports on vehicle crashes were direct outcomes of driver drowsiness, resulting in 1550 deaths, 71 000 injuries, and \$12.5 billion in monetary losses [1]. For such reason, car manufacturers in the world are eagerly developing diverse systems that can detect drowsiness of a driver and/or prevent drowsy driving.

However, the visual environment within the car provides many challenges for researchers, because the illumination varies very much as the car is moving and the face of the driver is often occluded by his hand or a coffee cup and etc. Although our main purpose is to track driver's head, if the face detector could not provide a detected face properly to the head tracker because of occlusion, the tracker would stop to work. For instance, even the most favored face detector, which is based on AdaBoost (Adaptive Boosting) algorithm [2], is often not able to detect the face when the face is partially occluded by other object as we show in the following sessions.

* 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 (NRF-2013R1A1A2006969) and by IT/SW Creative Research Program through the National IT Industry Promotion Agency (NIPA-2014-H0502-14-3024).

** Corresponding author.

In the present study, we would like to show that a face tracker, which is basically based upon Kalman filter [3], [4], [5] by combining the template matching method, can replace the conventional face detector. This tracker is robust against occlusion and it has some property that it tracks the rotated face well.

2 Face Detection vs. Face Tracking

Among many face detection algorithms, AdaBoost is the best known method. Its performance may depend upon what is the training data and how does trained it. In general, it detects the face well when it is in frontal. However, it is also known that its performance is degraded for illumination variation, occlusion and rotated face. In real driving situation, occlusion often occurs and the face of the driver is often rotated. Therefore, we propose a method of face tracking with occlusion detection using the template matching method [6]. In case of the conventional template matching, it does not deal the occlusion or rotation cases well. In this study, we have used the adaptive template method, that the template is updated by computing the error of the current region.

2.1 Face Tracking

First, we compute the transformation vector for face tracking. The transformation vector is the important value that can measure the velocity and direction in face tracking. The transformation vector $a(t)$ is given as follows:

$$a(t) = \arg \min_a \sum_{p \in R} \rho \left(\frac{\lambda}{\bar{r}} \right) \quad (1)$$

where $\lambda = I(\varphi(p: a), t) - \hat{g}(p, t')$

Where I denotes the frame number. $a(t)$ denotes the parameter vector of the transformation. I denotes the intensity of each pixel at current frame. \hat{g} is the intensity of template and \bar{r} is the standard deviation of λ . It can estimate the adaptive template using robust filter with transformation vector. We use a 4 ways template update method for the template prediction for the various conditions. The robust filter updates the template to measure the degree of obstruction in prediction template.

2.2 Occlusion Detection

The important task in tracking the face is how to detect occurrence of an occlusion. For this, we adopt the block matching method [7] because it allows us to detect the direction of the occlusion occurring in each region.

First, draw an outline of the face within the template employed in the tracking task. And, generate the blocks using the current template within the outline. The watershed algorithm [8] is used in making the outline. Though there are various methods in setting the block size, we use 60 block pieces and its mean size is 120x120. Each

region in the blocks has the weight value. Secondly, apply the weighting method to a pyramid in the basis of the center of the object because it has a decisive effect on the tracking for the occlusion occurred by the center of object. Thirdly, when the occlusion of the blocks occurs, the occlusion rate of the current template is calculated by multiplying these areas by the weight values. We call it the error rate since it calculates the occlusion rate. When the percentage of the occlusion is less than 25%, update the template using the basic method since the occurrence is not significant effect on the face tracking.

The conventional template update method cannot deal well when the percentage of the occlusion is more than 25. To solve that problem, a binary mask is generated for the remaining region except the occluded area. Then, the weight of each block is mortified by the template block to estimate the template area that will appear in the next frame. We call it the weight mask map. Once the weight mask map is generated, the template matching error between the previous and present frames is computed. Then, update the new template to reduce the matching error.

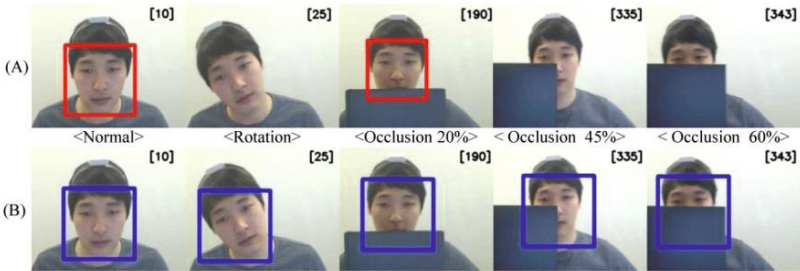


Fig. 1. Comparison between the conventional AdaBoost (A) and our face tracking method (B) for occlusions and the rotated head cases

Table 1. Error Rate (%)

Situation	Face Detection with AdaBoost	Face Tracking with Occlusion
Normal	5	2
Rotation	34	5
Occlusion 20%	24	11
Occlusion 45%	53	16
Occlusion 60%	91	23

2.3 Face Tracking with Occlusion Detection Result

This section presents the result from performance evaluation by comparing the AdaBoost and the face tracker using the video images as shown in Fig. 1. The video consists of 1,000 frames collected 30 fps in speed, containing diverse situation such as translation motion of the face and strong/weak occlusions. The degree of occlusion was 20%, 45%, 60% respectively. In Fig. 1, the top row (A) shows the result from the AdaBoost

algorithm, whereas the bottom row (B) results from the face tracking system. Although the face detector with the AdaBoost method was trained using frontal faces as well as rotated faces up to 15 degree, the face detection rate drops substantially for the rotated face and the strong occlusion where some facial features such as eyes, nose and mouth are disappearing. On the other hand, the face tracker is able to detect the face against diverse situations. Table 1 shows the result between two cases in which the errors for the AdaBoost case are counted whenever it did not detect the face, whereas those for the face tracker whenever the distance from the center of the face to the tracking point became wide. Notice the difference of two error rates for the rotated face case. Moreover, the gap between two error rates for the strong occlusion case became wide. Result suggests that the present face tracker is not only able to replace the face detector but also is actually better than in terms of performance.

3 Head Pose and Eye-Blink Estimation

3.1 Head Pose Estimation

The 3D head pose can be estimated using sequence of 2D images by calculating rotation vectors of the head. From an acquired face using the face tracker as described in previous section; initially facial features are extracted using corner detector. We then calculate optical flow between two pairs of features extracted from two frames. Given that the 3D movement of a head can be described using three moving components: $(\varphi, \theta, \gamma) = (\textit{Tilting}, \textit{Shaking}, \textit{Nodding})$. To estimate head pose, POSIT algorithm is adopted to measure the rotation of a 3D object. Here, an average 3D face model is created by using FaceGen [9].

The standard head pose database from Boston University (BU) [10] is used to evaluate performance of our head tracking system. BU database consists of 72 videos and each video contains 200 frames of images, captured under uniform and varying illumination conditions. It provides the ground truth values acquired by a gyro-sensor mounted on subject's head. We ran our tracking system against all videos from BU database. Performance is evaluated with Mean Average Error (MAE) of head pose.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\theta_i - \hat{\theta}_i| \quad (2)$$

Here, N refers to the entire frames and $\theta_i, \hat{\theta}_i$ the ground truth and the estimation value, respectively. As shown in Table 2, it shows high accuracy comparing with other head pose estimation methods.

Table 2. MAE Of Head pose angles

Degree of Rotation(°)		
Tilting	Shaking	Nodding
1.467	1.856	3.992

3.2 Detection of Eye-Blink

Here, we describe how to detect the eye blink of a driver. The number of blinks can be determined with the following equation. When d in equation (3) is less than 0.8 of the maximum height of eyes, we consider it as a blink state.

$$d = \frac{\sqrt{(Upper_x - Lower_x)} + \sqrt{(Upper_y - Lower_y)}}{2} \tag{3}$$

Eye-blink of the driver was counted based on the shape point around the eyes. When eye blinks, the upper and lower points are getting closer to each other.

4 Driver Drowsiness Detection Using Hidden Markov Model

We have used HMM in quantifying driver’s drowsiness. The head pose estimation and eye-blinking detection were used to create the HMM observation symbols and predict the probabilities of parameter with Baum-Welch algorithm, calculating the symbol state for each state. Driver’s head position and eye-blinking output are used to predict the state most suitable to driver’s state. Each state is composed of normal, warning and alert as shown in Figure 2(Left). In each epoch, 15 frames were used for temporal information to learn and assess the HMM used in the training. The results of nodding and blink state are compared to determine the current state of driver.

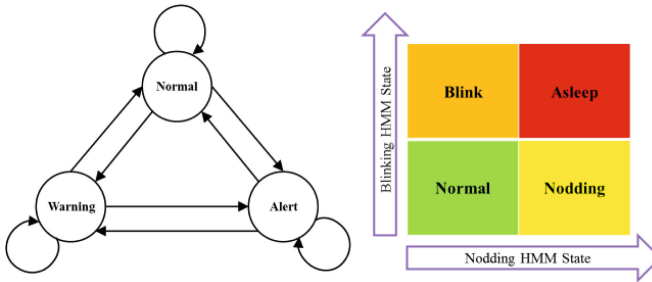


Fig. 2. HMM State Transition for drowsiness detection, having 3 different states (Left) and the criteria for determining driver’s drowsiness according to the nodding and eye-blink HMM States (Right)

Result is illustrated in Figure 3 where there are two examples: the fresh driver (top) and drowsy driver (bottom). Each panel consists of HMM states and drowsiness decision. Notice that whenever the blinking and nodding are overlapped, the red sign, i.e. asleep, is flashed in each panel.

5 Conclusion

We propose that the face tracker based on Kalman filter combining by the adaptive template matching method, can replace the conventional face detector such as the

AdaBoost method, especially in tracking 3D head movements of the driver in this paper. Performance of the conventional face detector deteriorates particularly when the given face is partially occluded by other object or rotated in a great deal. It is shown that our face tracker is more reliable for the rotated face and it is possible to handle well the occlusion cases using the occlusion and disocclusion detection system.

Given that the driver's head in the car is often occluded by either his hand or other objects and is often rotated, it is necessary to handle such cases to track the head in recognizing drowsiness of him. It is believed that the occlusion detection algorithm presented in this study can be applied in dealing the strong shadow, frequently occurring within the car running across the street, cast on driver's face.

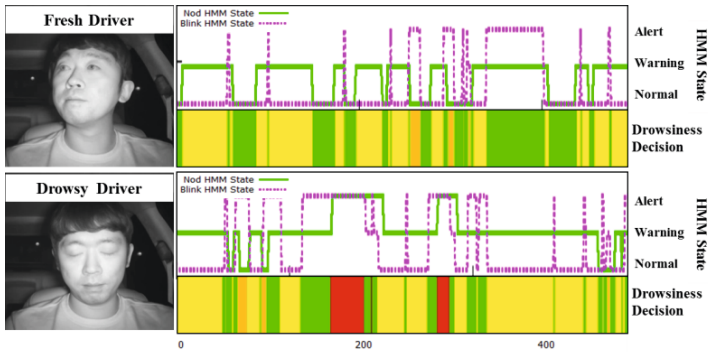


Fig. 3. (Left) Some of the head shots acquired using an infra-red camera; (Right) Driver's drowsiness inference using HMM

References

- [1] Ranney, T.A., Mazzae, E., Garrot, R., Goodman, M.J.: NHTSA driver distraction research: Past, present, and future, Nat. Highway Traffic Safety Admin., East Liberty, OH, USA, Tech. Rep (2000)
- [2] Legters, G.R., Young, T.Y.: A mathematical model for computer image tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6), 583–594 (1982)
- [3] Froba, B., Ernst, A.: Face detection with the modified census transform. In: *Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 91–96 (2004)
- [4] Brunelli, R.: *Template matching techniques in computer vision: theory and practice* (2008)
- [5] Nguyen, H.T., Worring, M., Van Den Boomgaard, R.: Occlusion robust adaptive template tracking. In: *Proceedings of Eighth IEEE International Conference on Computer Vision, ICCV 2001*, vol. 1, pp. 678–683 (2001)
- [6] Pan, J., Hu, B.: Robust occlusion handling in object tracking. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007*, pp. 1–8 (2007)

- [7] Beucher, S., Meyer, F.: The morphological approach to segmentation: the watershed transformation. In: Dougherty, E. (ed.) *Mathematical Morphology in Image Processing*, pp. 433–481. Marcel Dekker, New York (1993)
- [8] Amini, A.A., Weymouth, T.E., Jain, R.C.: Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(9), 855–867 (1990)
- [9] FaceGen Modeller, <http://www.facegen.com>
- [10] La Cascia, M., Isidoro, J., Sclaroff, S.: Head tracking via robust registration in texture map images. In: 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 508–514 (1998)

The Disaster Rescue Robot Design and Implementation Using Open Source

Yung-Hui Chen¹ and Jyu-Wei Wang²

¹ Dept. of Computer Science and Information Engineering, Aisa University, WuFeng
41354 Taichung, Taiwan
skuldchen@gmail.com,
skuldchen@live.asia.edu.tw

² Dept. of Photonics and Communication Engineering, Aisa University, WuFeng
41354 Taichung, Taiwan
jwwang@asia.edu.tw

Abstract. There are many disaster happened around the world. In addition there are unknown space or environment require human seek out it. Especially when disaster could move space for different between in before and after. People don't know any information can assist them make a right policy decision. Also human's perceptions don't very sensitivity, such as optical illusion. The nose can't smell and identify harmful gases, or ear can't identify source of sound. People are need sensors and algorithms to assist them in that scenario. This research presents a method an easily design prototype robot. And robot equipped multiplex sensor used open source hardware. This robot has variety functions. Those functions come from design idea and application. This robot has equipped compass sensor aka M-sensor. The robot equipped Accelerometer, aka M-sensor. Robot equipped other sensors and algorithm could provide navigation and so on function.

Keywords: Robot, Navigation, firmware, human interface software, hardware device driver software, Arduino Mega, Accelerometer, G-sensor, Compass sensor, M-sensor, Spherical trigonometry, Grate circle distance, color space, image processing, canvas, 3D space, optical illusion.

1 Introduction

Nowadays the microcontrollers are powerful usage for several atmospheres, such as wireless sensor network (WSN) [21-24], autonomous system, automation system, robotics [25-27, 31], smart home control, and even can used on hobby, art, education and Medical. Those system components can class to control chip part, sensor part, specific logic part and power energy. Normally the Logic part is hybrid software which is firmware. There are varieties of hardware embedded different microcontroller chip module you can select on your work. Such NI, TI, Silicon and more. Microcontrollers are tiny, signal ship, but there are also expensive. In recent years, Arduino [1] is one of less budget, low cost, common hardware, good for research and open source hardware, mostly easy to used, ability to operator and work with difference Operation Systems such as Windows, Mac OS, Linux, and UNIX-like.

While there are many news, papers, thesis, journals, technology report, and researches to describe the robot, such robot for a life be smart and autonomy. It is beautiful dreams, but the problem is how to design and implement it. Robot has many details in its design. In recent years, NASA in Mars science mission has proposes their building robot, which is “Curiosity”. It is a car-sized robot. It is equipped multi sensor, those sensor let she have power function to face and assist finished her mission. Many papers to describe such as live life or dreams, but normally is only to describe a concept, and without how to implement. Further when student, novice, new researcher would designs and builds their project, problem has be coming and need face it. So that, when we usually design those systems, we widely survey many news, papers, journals, technology report, technology document, and researches papers, which is widely data. However, they are almost describes concept expect to implement that.

We will show you more details of our design, which is include how consider in circuit design of hardware, how software of hardware drive design, aka firmware. Finally we design software of control client, aka User Interface software. Perhaps, in the future we can portable from this system to any platform as soon as possible.

Table 1. End Products compare

Item	Budget	Extern available	Note
Lego NXT [10]	High	None, Black Box	
NAO [9]	High	Fixed	
TraxBot [12]	Middle	Restrict / Limit	
Bioloid [21]	High	Limit	
Arduino MCU	Low	Highly flexible	Our select

Table 2. Develop software compare

Item	Budget	Extern available	Note
LabView [11]	High	Yes, but maybe wait for engineering update	
NI vision Assistant	High	Yes, but maybe wait for engineering update	Need Labview
Matlab	High	Yes, but maybe wait for engineering update	
Visual studio Express edition	Free	Yes.	Our select.
Arduino IDE	Free	Yes.	Our select

Table 3. Equipment and Goods compare

Company	NI	Lego	Arduino
Item price	High	High	Mid
Software	LabView	Library and NXT	Any
Software price	NTD 70000+	NTD 10000+	Freeware
Troubleshooting	wait for engineering update	Wait for engineering update	Yourselfer

2 Methods, System Design

We omit large and heavy related works for readers. Instead of that we list some references [9, 10, 12, 25-27, 31] for readers when they are interest and improvement their research. Table 1, 2, 3 are we compare some robots function, ability and its budgets.

In this method for our design in Hardware side, we consider cheaper, low budget, cost down, and more over we can control chips, device driver, firmware, software by our self. This is means we don't want to at station is what happened and we need wait for some thirty parity to troubleshoot. We can control and handle we self.

We consider our Robot fundament functions include Compass, GPS, Accelerometer, Stand alone Power supply in principle. That we design some circuit with component [3], and use Arduino shield. That means we consider and calculate how many Digital Pins, Communication channel, Analog Pins and so on we need.

We select Arduino Broad, Mega 2560 are we selected. Moreover we need draw out our imagine prototype vision, make the dream be true. And this can help us more clearly to know our goal. Imagine that, using Compass because if we are in unknown and less information environment, compass can guide us direction information. If we can moreover to combine maps, then we could more clearly to know all of information. If we have GPS, then we can more easily to know what position we are, and Robot too. Accelerometer can help the robots to know its attitude, controller can easily to know robot are stand right.

2.1 Platform

Arduino Mega 2560 [1, 2] is a microcontroller production by Atmel, based on ATmega2560 chip. According to datasheet of ATmega2560, this signal chip is very powerful function that let Arduino Mega 2560 have strong application. Arduino Mega 2560 board, work on 5 Voltage, it has 54 digital Input/Output pins, 16 analog pins , 4 UARTs (serial ports) at 0, 1, 19, 18, 17, 16, 15, 14. TTL serial data are 0, 1, which is also used to connected to FTDI USB to TTL. USB connection which is connection to PC, provides ISP and work energy. Power jack which is can connection stand alone power supply or battery to provide more energy, and analog pins are inside A/D convert. Also in digital pins provide 15 PWM pins (2 to 13 and 44, 45, 46). 128kb Flash memory, work on 16MHz crystal oscillator (clock). Moreover it has SPI

communication which is 50 (MISO), 51 (MOSI), 52 (SCK), 53 (SS). Also, TWI pins are Inter-Integrated Circuit (IIC, I2C) communication which is 20 serial data line (SDA), and 21 serial clock line (SCL). One LED (13) on board.

Arduino provide UART communication, which is serial port to communication to PC and Devices. The Arduino Mega 2560 have 4 UARTs that us more flexible to design.

Despite Arduino have its inside DC provide, 3.3V and 5V. But sometimes you may need difference Power supply or stand alone power energy, DC to DC convert solution this is. That allow you design the battery pools, and then used DC to DC Convert to supply you wanted Voltage power is go solution to device need, suggest 7.2V, 12V and so on, respectively. Shield module can provide 3~36Voltage DC, and if battery recharge or change, don't need to warring provide DC energy supply. DC to DC convert allow you to rise up or decline, go down Voltage. LM2596 provide step down voltage, its input 4.5~35 Voltage and output 1.25 ~ 26 Voltage at 2A. LM2577 provide step up voltage, its input 3~34 Voltage and output 4~60 Voltage at 2A. Only one need care is heat processing, use thermal grease, heat sink and cool fins to treatment this is.

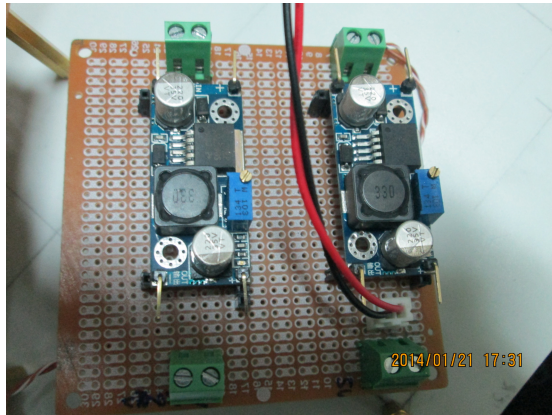


Fig. 1. DC to DC convert our design circuit 1

Battery always was power energy source provide everything of electronics goods to work, thus stand alone power are need careful to consider, include costing of cycle life. Each battery have its reference and normally current/voltage power energy provide [15] to work times dependent how goods used.

2.2 Software Template

This system requires algorithm to process and calculate our method. Here, we used Microsoft Visual Studio Express [5] to design our system application software. Microsoft announce Visual Studio Express for free and student program [4]. This announcement helps engineer, researcher, artists, hobbyist and novice to fulfill their dreams. About its principle datasheet, technology report of how to encode and more detail function, reader can visit Microsoft Developer Network (MSDN) [6].

Microsoft Visual studio has provided C++, C#, Basic and more program language. C# language have Command line mode, Windows Form Application mode (WinForm), Windows Presentation Foundation (WPF) mode, Active Server Pages (ASP) .Net Mode. Each mode has its features and templates. Those various program language and mode means that it's difficult for programmer to study and specialize on it. In this study we are using used C# language, and coding on WinForm mode and WFP mode.

On the Earth, geomagnetic field are covered [28-30]. This Axiom is exist and support man to guide direction in the world. Compass is dependent on those axiom and provide the message for us. Digital Compass is more novel technology.

2.3 Sensors

Compass sensor GY 26[12], aka. device on HMC1022 compass chip [12], operating on 3~5 Voltage DC, -20 to 85 degree C, current 15mA in 5V or 8mA in 3V, provide 0 to 360 degrees and 0.1 degree resolution, 1 degree measuring precision, 9600 Baud rate, provide IIC or RS232 which is Tx, Rx Communication. Navigation are normally application. Include help or assistant Telescope position, hand guide, Guide system, GPS guide, Antenna position, Robot position and guide orientation. Airspace and aerospace model position also are important. We used compass to improve us more information to help decision.

Accelerometer sensor, aka G-Sensor, is provided a vector of moving change. Its operating on 5 Voltage

Now we have to know many sensors and it used, we may compile it into table.

The follow table shows how we design in this project consider power energy, allocate and pins methods.

2.4 Implementation

We describe to implement our system in this session. The system composes hardware, software and algorithm. Now we have to know many sensors and it used on wide field, we may compile it into table. The table shows how we design in this project consider power energy, allocate and pins methods.

Firmware design is to design to control circuit, motors, get and send message. Normally send out control pins signals or reading single from IC message.

Design Arduino firmware is not to hard too. Arduino provide it IDE to develop, named Arduino sketch. This is develop can provide ISP function, plug c library of Arduino, also Arduino already provide some Open Source function on there. The robot firmware principle function is forward, backward, turn-left, turn-right and stop. Those design need consider into main loop. Arduino firmware have 2 necessary function, setup and loop. Setup like to Visual studio c# as WinForm, it is run once in start, beginning. Loop was running until to close Arduino. So that to design need consider with you add sensor. Firmware follow are very important in here. When we add M-sensor as GY-26, or G-sensor as MMA7455, subscript function was import into loop. All system photos we show on Figure 3.

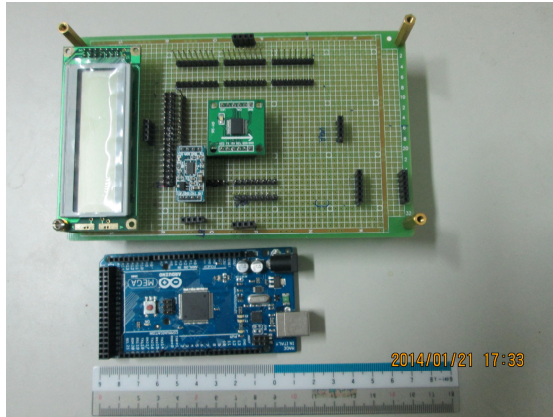


Fig. 2. Arduino Mega 2560 and our design circuit 2

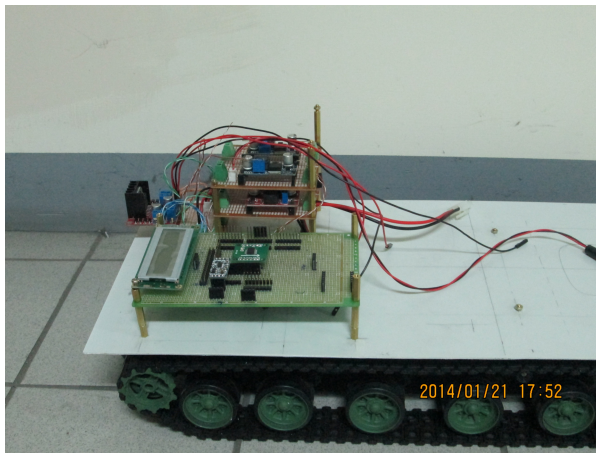


Fig. 3. Our system photo

3 Future Work

This project idea come with author's pass experience on his duty service on military, Taiwan, ROC. He has experience on many Operation mission. Dependent on his experience, he never see one of machine can support enough information to help decision, than after his research, he knows difficult on each part need redesign and he want solution it. After this project, in the future all of this project should need more functional be design. Such as we did not consider work on cool weather because this situation is very harsh environment, need to consider ESD, EMP. ESD and EMP would destroy or burn circuit, chips, and so on. Currently we design on warm weather can work. Future we select more high level chips those can work on cool weather and on more disaster environment.

Secondary, the battery, power pools consumption are very important, how to supervisor Battery furl gauges and send message feedback to controller, user how the robotics still have action time, works time and even provide automatic robots to re-charge its battery is very important. We can design this in future.

Transport message is hardly work, dependent how large data bytes need sends and receives. Video data such as Webcam or Kinect data sender and review. Bluetooth has it limit can not quickly send Video message. Large data transport via wireless is one of important issue. Especially in simultaneous localization and mapping (SLAM) via wireless communication to User Interface need wide band capacity. Finally, dependent on our limit resource, we select and abort some good idea. All of those ides can improve this system performance.

4 Conclusion

We brief our design and improvement, our considering functional robot, we describe way to how to easily modify and true to build invention your own robotics. Perhaps this study can help the readers easily to build there own robotics. This robot functions have compass can navigation you way. It has sensors can detector environment situations. Optical sensor can provide video and depth image. Depth image can measurement distance. Mick array we design algorithm can tracking voice. Accelerator with our algorithm can provide the robot's attitudes. Gases sensors whit our algorithm can identify and provide people what air composition in there.

We wanted this can be interesting more people join and enjoy in science or development, invention. Our robotics provides hardware, logic, software, compass to navigate and prevent loss in new space after disaster. This project is work in progress. We would like this project can help or interesting more people. We would happy to response and troubleshoot when who have question.

NOTE: Because the regular of this conference, the paper length only 8 pages we should follow it and this limit us to cart our paper. Therefore our paper seems not enough. Please, readers can contact us if they have questions.

References

- [1] Arduino Mega 2560, <http://arduino.cc/en/Main/arduinoBoardMega2560>
- [2] Atmel datasheet, <http://www.atmel.com/Images/doc2549.pdf>
- [3] Fritzing, <http://fritzing.org/home/>
- [4] Microsoft Open Source License, OSI approves MS license submission, <http://opensource.org/node/207>
- [5] Microsoft license 1, <http://social.msdn.microsoft.com/Forums/en-US/0782e1b0-db87-4de3-b79d-ba56a481e750/visual-studio-2010-express-is-now-available?forum=Vsexpressinstall>
- [6] Microsoft license 2, students, <http://www.microsoft.com/en-us/student/default.aspx#fbid=fN3b6FzRgC9>

- [7] HMC1022, <http://shop.aiscube.com/datasheet/DigitalCompass/HMC1022%20Digital%20Compass.pdf>
- [8] NAO robot, <http://www.aldebaran-robotics.com/en/>
- [9] Lego robot, <http://www.lego.com/en-us/mindstorms/?domainredir=mindstorms.lego.com>
- [10] NI and its' products LabView, <http://www.ni.com/>
- [11] Traxbot robot, http://wiki.ros.org/traxbot_robot
- [12] TI, Microcontroller chip company and factory, <http://www.ti.com/>
- [13] Battery Reference 7.2 V,
http://www.vexrobotics.com/wiki/7.2v_NiCd_Robot_Battery
- [14] Gearbox design 1 [http://en.wikipedia.org/wiki/Transmission_\(mechanics\)](http://en.wikipedia.org/wiki/Transmission_(mechanics))
- [15] Gearbox design 2, <http://www.khkgears.co.jp/en/>
- [16] Inverse Kinematics Algorithm,
http://en.wikipedia.org/wiki/Inverse_kinematics
- [17] Watt, A., Policarpo, F.: 3D Games Real time Rendering and Software Technology, pp. 400–401. ACM Press, New York (2001) ISBN: 0-201-61921-0
- [18] Tsmots, I., Teslyuk, V., Vavruk, I.: Hardware and Sotdware tools for motion control of mobile robotics system. In: CADSM 2013, Polyana-Svalyava, UKRAINE, p. 368 (February 2013)
- [19] Spatial science, http://en.wikipedia.org/wiki/Spatial_science
- [20] South Korea, Bioloid robots, <http://www.robotis.com/xen/>
- [21] Akyildiz, I.F., Su, W., Sankarasubramaniam, T., Cayirci, E.: A Survey on sensor networks. IEEE Communication Magazine, 102–114 (August 2002)
- [22] Akyildiz, I.F., Su, W., Sankarasubramaniam, T., Cayirci, E.: Wireless sensor networks: a survey. In: Computer Networks, pp. 393–422. Elsevier (2002)
- [23] Akyildiz, I.F., Melodia, T., Chowdury, K.R.: Wireless Multimedia wireless sensor network. IEEE Wireless Communication, 32–39 (December 2007)
- [24] Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. In: Computer Networks, pp. 2292–2330. Elsevier (2008)
- [25] Menzel, P., D'Aluisio, F.: Robo Sapiens- Evolution of a new species. A Material World Book (2000) ISBN:9780262133821
- [26] Menzel, P., D'Aluisio, F.: Chinese language translated edition, 林文源譯 “機器人的進化-人工智慧與機器人學的新世紀, (Robo Sapiens-evolution of a new species)’ 商周出版 ISBN:9867892070
- [27] Robot Wikipedia, <http://en.wikipedia.org/wiki/Robot>
- [28] NASA Earth's Magnetic Field 1, geomagnetic field,
http://www.nasa.gov/mission_pages/themis/news/themis_leaky_shield.html
- [29] Earth's magnetic field 2,
http://en.wikipedia.org/wiki/Earth's_magnetic_field
- [30] Earth's magnetic field 3,
<http://www.appinsys.com/globalwarming/earthmagneticfield.htm>
- [31] Robot, <http://en.wikipedia.org/wiki/Kuratas>

The Direct and Indirect Effects of Perceived Usability of Smartphone Applications on User Satisfaction

Wonjin Jung and Taehwan Kim*

The School of Business and Economics, Dankook University
152, Jook-Jun-Ro, Soo-Ji-Goo, Young-In, Kyung-Ki-Do, Korea, 448-701
{jungw, thkim}@dankook.ac.kr

Abstract. Recently many innovative, technological advances have been introduced in smartphone applications. Nevertheless, not every application is usable because some are neither simple to understand, nor easy to use. In some cases, users are not motivated to use new applications, and this lack of motivation is believed to be related to a lack of satisfaction with less usable applications. A review of the relevant literature reveals little research of the direct and indirect effects that the perceived usability of smartphone applications has on user satisfaction. This study aims to examine: 1) the effects that perceived usability of applications has on both user satisfaction with the given application as well as users' intrinsic motivation, 2) the effect that users' intrinsic motivation has on user satisfaction. Structural Equation Modeling (SEM) was used to analyze data collected through a survey. The results indicate that direct and indirect effects of the perceived usability of smartphone applications influence user satisfaction. This study contributes to the IS literature by presenting empirical evidence relevant to usability and users' intrinsic motivation with respect to using a smartphone application.

Keywords: Smartphone, application, usability, motivation, satisfaction.

1 Introduction

Growth in the global smartphone market accelerated after the introduction of the iPhone in 2007 by Apple. In 2012, there were more than one billion smartphones in use worldwide [11], and the rapid penetration of smartphones has resulted in a wider use of applications [12]. In 2014, there were 1.2 million applications in Apple's App Store and 1.3 million applications in Google Play [8, 9, 10].

Recently, smartphone applications have introduced a number of innovative technological advances. These applications provide greater functionality and better access to information, and consequently, have become larger in size as well as more complicated. In general, complicated applications are less usable. They are neither easy to understand, nor to use. Users become confused when they encounter a complicated application, and at worst, this may have a negative impact on user

* Corresponding author.

satisfaction with a given application. Therefore, applications need to be usable. Usable applications are not only simple enough to use, but are also interesting enough to motivate users to keep using them.

Nevertheless, not every application is usable, and some applications do not even motivate users to keep using them. In the literature, motivation is viewed as not only a natural inclination toward master and interest, but also as an important source of enjoyment and vitality [3]. Several researchers have noted that motivation, as a theoretical construct, can explain the reasons for our behavior [4, 7]. Holbrook and Hirschman [7] asserted that motivated consumers purchase products or services due to the interest, joy, and satisfaction that they experience.

These ideas with respect to motivation are applicable to the smartphone market. When users are motivated to use applications due to a high level of usability, they may be willing to spend more time on the applications. The more time they spend on a given application, the more opportunity they have to be satisfied. Thus, smartphone users' intrinsic motivation seems to play an important role, especially when users make decisions regarding their adoption and use of smartphone applications.

In sum, the literature review and the discussion above indicate that the usability of applications as perceived by smartphone users seems to be critical determinant of not only user satisfaction, but also users' intrinsic motivation. In addition, users' intrinsic motivation may influence their satisfaction with a given application.

Therefore, the following hypotheses are proposed.

H1: The usability of smartphone applications as perceived by smartphone users positively affects user satisfaction.

H2: The usability of smartphone applications as perceived by smartphone users positively affects their intrinsic motivation to use the applications.

H3: Smartphone users' intrinsic motivation to use applications positively affects their satisfaction.

2 Research Methodology, Data Analysis, and Results

The goal of this study is to examine the direct and indirect effects that the perceived usability of smartphone applications has on user satisfaction. A survey was conducted to collect data, and a total of 236 students and practitioners participated in the survey. 33.1% of the participants were practitioners, and 66.9% were undergraduate students majoring in various academic programs including business administration, economics, and computer science at three major universities in Korea. 50.4% percent of the participants were male, and 80.1% were in their twenties. About 61.9% of the applications that participants had used just before answered the questions on the survey were social networking and communication applications.

This study employed a Structural Equation Model (SEM) to test the proposed research model. SPSS Statistics with AMOS ver. 18 was used for the analysis. First, the reliability of the individual instrument items was examined to test the proposed measurement model. In this respect, all reliability measures, that is, the loadings of the items on their respective constructs, should be 0.6 [1, 2]. The results showed that all of the loadings were 0.6 or higher, suggesting adequate reliability (see Table 1).

Table 1. Standardized Regression Weights of Observable Variables, Composite Reliability (CR), and Average Variance Extracted (AVE)

Latent Variables	Estimates	Variance C.R.	Composite Reliability	AVE
Perceived Usability	.840	6.644	.906	.671
	.825	7.098		
	.791	7.951		
Intrinsic Motivation	.704	9.049	.848	.651
	.938	2.210		
	.760	8.023		
User Satisfaction	.777	8.092	.900	.656
	.797	7.645		
	.854	6.000		

Next, the convergent validity of the instrument items was tested. To do so, the composite reliability (CR) and the average variance extracted (AVE) of the latent variables were examined. The values for CR and AVE were manually calculated since AMOS does not provide a function for them. The formulas below, as suggested by Fornell and Larcker [5] and Hair et al. [6], were used for the calculation. The results indicate that the values of CR for all constructs were of 0.8 or higher, well above the recommended tolerance of 0.7 (see Table 1). In addition, all constructs also had estimates for AVE greater than the recommended cutoff of 0.5 (see Table 1). Thus, the instrument items demonstrated satisfactory convergent validity.

$$CR = (\sum \text{Standardized Regression Weights})^2 / ((\sum \text{Standardized Regression Weights})^2 + (\sum \text{Variance})) \quad [5]$$

$$AVE = (\sum \text{Standardized Regression Weights}^2) / N \quad [6]$$

Next, the discriminant validity of the instrument items was examined. To satisfy the requirements for the discriminant validity, all constructs should have the square root of their AVEs be greater than the correlations with other constructs in the proposed model [2], and the results showed that this was the case for each construct in the model (see Table 2). Therefore, the discriminant validity for the items was also confirmed.

Table 2. Correlation Coefficient Value between Constructs and AVE

Constructs	AVE	ρ^2	ρ^2	ρ^2
Perceived Usability	.671	.379	.074	1.000
Intrinsic Motivation	.651	.091	1.000	
User Satisfaction	.656	1.000		

The goodness of fit was evaluated to test the structural model. The indices for the goodness of fit include χ^2/df , GFI, AGFI, NFI, TLI, CFI, and RMSEA, and the results were as follows: $\chi^2/df = 2.422$, GFI = .950, AGFI = .906, NFI = .947, TLI = .951, CFI = .968, and RMSEA = .078, which indicate that the model has a fairly good fit in general.

Finally, the structural model was tested by analyzing the significance and the strength of the relationships between the variables in the model. To do so, estimates of the path coefficients were examined. Table 3 below shows the results of the test. As expected, the perceived usability had a significant influence on user satisfaction ($\beta = .599$, $p = .000$) as well as on intrinsic motivation ($\beta = .324$, $p = .000$). In addition, the intrinsic motivation had a positive impact on user satisfaction ($\beta = .125$, $p = .029$). Thus, Hypotheses 1, 2, and 3 were supported. Figure 1 presents the results of the structural model with R^2 values representing the amount of variance.

Table 3. Hypothesis Test

	Paths	Coeff.	Stand. Coeff.	P	Results
H1	Perceived Usability -> User Satisfaction	.591	.577	***	Accept
H2	Perceived Usability -> Intrinsic Motivation	.324	.272	***	Accept
H3	Intrinsic Motivation -> User Satisfaction	.125	.145	.029	Accept

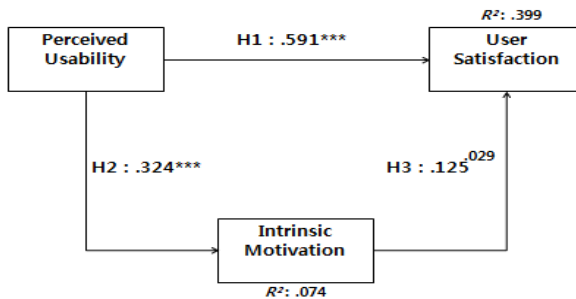


Fig. 1. Structural Model Results

3 Discussion and Conclusion

This study empirically examined the direct effect that the perceived usability of smartphone applications had on user satisfaction as well as the indirect effect due to users' intrinsic motivation, and the results revealed that both the direct and indirect effect significantly influenced user satisfaction.

These results indicate that when smartphone applications are usable, users both experience satisfaction and are motivated to use them. Smartphone users who are motivated continue using the applications and are then more likely to experience

satisfaction with the applications. These findings highlight the importance of usability in smartphone applications. In fact, today's smartphone applications have become larger as more functions and more information are added. At the same time, the applications have become ever more complicated and, consequently, have become less usable. In general, less usable applications are neither simple, nor easy to use. This study suggests that smartphone applications should be usable in some way if they have to satisfy users since direct and indirect effects of usability affect user satisfaction.

In addition to the academic contribution to IS literature, this study also provides a few implications for practitioners. Practitioners can use the findings of this study to improve the usability of their applications. Since user interface is strongly related to usability, effort should be made to improve user interface design in applications as an effective strategy for developers to differentiate their applications. Despite presenting interesting findings, this study is subject to the limitations of empirical research.

References

- [1] Barclay, D., Higgins, C., Thompson, R.: The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration. *Technology Studies* 2, 285–324 (1995)
- [2] Chin, W.W.: The Partial Least Squares Approach for Structural Equation Modeling. In: Marcoulides, G.A. (ed.) *Modern Methods for Business Research*, pp. 295–336. Lawrence Erlbaum, Mahwah (1998)
- [3] Csikszentmihalyi, M.: *Creativity: Flow and the Psychology of Discovery and Invention*. Harper Collins, New York (1996)
- [4] Elliot, A., Covington, M.: Approach and Avoidance Motivation. *Educational Psychology Review* 13, 73–92 (2001)
- [5] Fornell, C., Larcker, D.F.: Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research* 18, 39–50 (1981)
- [6] Hair, J.F., Black, B., Babin, B., Andersong, R.E., Tatham, R.L.: *Multivariate Data Analysis*, 6th edn. Pearson Prentice Hall, Upper Saddle River (2006)
- [7] Holbrook, M.B., Hirschman, E.C.: The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun. *Journal of Consumer Research* 9, 132–140 (1982)
- [8] Keach, S.: Microsoft Says Windows Phone Now Touts 300,000 Apps, <http://www.t3.com/news/microsoft-says-windows-phone-now-touts-300000-apps>
- [9] Perez, S.: Mobile App Usage Increases In, As Mobile Web Surfing Declines (2014), <http://techcrunch.com/2014/04/01/mobile-app-usage-increases-in-2014-as-mobile-web-surfing-declines/>
- [10] Perez, S.: iTunes App Store Now Has 1.2 Million Apps, Has Seen 75 Billion Downloads to Date, <http://techcrunch.com/2014/06/02/itunes-app-store-now-has-1-2-million-apps-has-seen-75-billion-downloads-to-date/>
- [11] Reisinger, D.: Worldwide Smartphone User Base Hits 1 Billion, <http://www.cnet.com/news/worldwide-smartphone-user-base-hits-1-billion/> (retrieved October 25, 2014)
- [12] Yang, H.C.: Bon Appétit for Apps: Young American Consumers' Acceptance of Mobile Applications. *Journal of Computer Information Systems* 53, 85–96 (2013)

Appendix

Latent Variables	Questions
Perceived Usability	The functions that I tried to find in the application were quickly found.
	The information that I tried to find in the application was easily found.
	The application that I recently used provided information quickly.
Intrinsic Motivation	I am interested in learning how to use the smartphone application that I used most recently.
	As much as possible, I want to learn how to use the smartphone application that I used most recently.
	As much as possible, I want to know how to use the smartphone application that I used most recently.
User Satisfaction	The experience of smartphone application I used most recently was very satisfying.
	The use of the smartphone application that I used most recently was a pleasant experience.
	Overall, I was satisfied with the smartphone application that I used most recently in every aspect.

A New Query Integrity Verification Method for Encrypted Data in Database Outsourcing

Miyoung Jang, Min Yoon, Youngho Song, and Jae-Woo Chang

Dept. of Computer Engineering, Chonbuk National University
567 Baekjae-daero, Jeonju, Republic of Korea
{brilliant, myoon, yhsong, jwchang}@jbnu.ac.kr

Abstract. In database outsourcing, two issues of data security emerge: data confidentiality and data integrity. Existing data transformation schemes were widely studied for preserving data confidentiality, but they are vulnerable to data leakage problem because they do not consider data distribution when encrypting data. Meanwhile, query authentication schemes verifying data integrity, suffer from transmission overhead for verification data. In this paper, we propose a privacy-aware query authentication scheme which guarantees the data confidentiality and the query result integrity of sensitive data. To solve the original data leakage problem, we propose a bitmap-based data transformation scheme with anchor selection based on data distribution. Also, we devise a query result authentication index that stores an encrypted signature for each anchor so that it can reduce the amount of auditing data. Through performance evaluation, we show that our scheme outperforms the state-of-the-art method in terms of query processing time and verification overhead.

Keywords: Data confidentiality, data transformation, query result authentication, k-NN query processing algorithm.

1 Introduction and Background

In database outsourcing, because the service provider might be untrusted or compromised, two issues of data security emerge: data confidentiality and data integrity. Data transformation methods[1-2] are widely studied for preserving data confidentiality by transforming the original data domain into another one. They divide the whole space into partitions and encrypt each partition by using space perturbation. However, the existing schemes have two main problems. First, the random selection of anchor objects can cause the skewed distribution of clusters. If an adversary has the background knowledge of the data distribution, the cluster information can be easily revealed. If queries are converged on the densely populated cluster, computation cost is highly increased. Second, because some of the existing techniques employ a tree-based index scheme, they can be only used for ordered plaintext. In addition, the query processing cost highly depends on the tree's depth.

On the other hand, in order to verify the correctness and completeness of outsourced data, query authentication should be provided in database outsourcing environment. The baseline approach for this is to send the verification information signed by the data owner to users with query result so that the result can be verified by the signature. Previous data authentication researches [3-6] can be categorized into three classes: signature-based approaches, authenticated data structures and bucket-based authentication. First, signature-based approaches assign one signature to each data tuple and verify data integrity by comparing all the signatures of data within the results. This leads significant overheads for users both in time and space. Second, authenticated data structure based approaches generate a tree-based data index. The root is signed with data owner's signature as a concatenation of all children nodes' information. However, authentication data structure does not guarantee the data confidentiality, since it cannot be built on the encrypted data. In addition, tree-based index suffer from data update overhead and verification object transmission costs. Finally, J. Wang et al. [3] proposed a bucket-based authentication scheme where a bucket contains a bucket id, data range (upper-lower bound), a checksum and the number of tuples in a bucket. A checksum is similar concept to the data signature and generated by using a Hash function. The limitations of the existing bucket-based authentication scheme are as follows. First, because it generates a bucket with equal width of the data range, the original data distribution can be disclosed. Moreover bucket id is assigned as the ascending order of the data range. The existing bucket-based authentication cannot fully provide the data security. Secondly, because the existing bucket-based authentication methods only consider relational database contents, the distribution of database may not be protected. However, in case of sensitive databases, e.g. physical or mental health details, purchase records and political issues, their distribution can involve meaningful knowledge. To provide thorough protection for the sensitive database, any meaningful information including data distribution should not be revealed.

Motivated by these problems, in this paper, we propose a privacy-aware query authentication scheme which guarantees the data confidentiality and the query result integrity of sensitive data. First, to solve the original data leakage problem, we devise a bitmap-based encryption scheme by selecting anchors based on data distribution. Uniformly distributed partitions prevent the attackers from inferring the original data distribution. We also design an algebraic coding-based hash index that transforms a query to bitmap data and retrieves the anchor information efficiently. Second, to reduce the transmission overhead of verification data, we devise a query result authentication index that stores an encrypted signature for each anchor and compares the anchor signature with the verification data from the data owner. Hence, we can reduce data transfer overhead for query integrity checking while enhancing data privacy.

The rest of this paper is organized as follows. Section 2 presents our bitmap encryption based data integrity scheme is proposed. Section 3 provides an experimental evaluation on the existing and the proposed methods. Section 4 concludes this paper with further research directions.

2 Bitmap-Based Data Transformation and Range Query Processing with Query Result Authentication

2.1 Data Transformation

In this section, we propose a bitmap-based data transformation technique for processing a private nearest neighbor search query. The main advantage of this technique is that the server always returns a constant-sized candidate set. Upon receiving a candidate result set, the client refines them to obtain a final result. In data transformation phase, we select a set of anchors based on the data distribution. For this, we utilize a grid index in order to generate a histogram that shows the distribution of data.

Let P be the original dataset of objects. In the data transformation, we generate a grid index for choosing a set of objects from the set P as anchor objects. By using the grid index, the data owner selects anchors based on the distribution of data. As a result, more number of objects are selected in dense area where as less number of objects are elected as anchors in sparse area. Then, the data owner specifies a parameter r for data grouping. Parameter r means a data range for each anchor. This parameter provides a trade-off between the cost of transformation and the query accuracy. That is to say, a higher value of r leads to better query accuracy, but also higher transformation cost. The data transformation is done in two detailed steps: anchor selection and bitmap data generation.

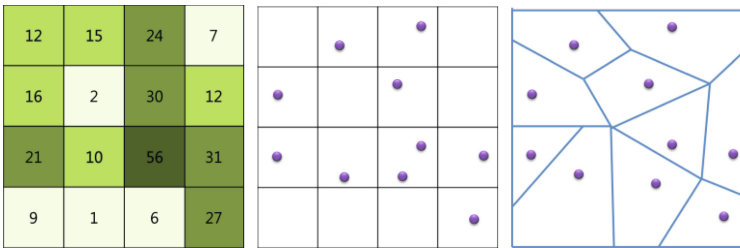


Fig. 1. (a) number of data per cell; (b) selected anchors; (c) Voronoi diagram-based index

Step 1: anchor selection. Anchor selection is performed as follows. First, the original dataset (D) is inserted in the $n*n$ grid index. Secondly, the algorithm counts the number of data within each grid cell, and sorts the cells in the descending order of data count. Thirdly, from the most dense grid cell, the data owner determines the number of anchors based on equation 1, and then randomly selects anchors from the cell.

$$A = \frac{(\#of\ anchors) \times (\#of\ data\ in\ anchor\ range)}{\#of\ data\ in\ D} \tag{1}$$

For example, in figure 1(a), the grid cell that encloses 56 data is the most densely populated cell so that the anchor selection is performed from it. Assume that we divide data into 10 groups, we select 10 anchors and 2 anchors will be randomly selected from the most densely populated cell based on the equation (1) (see Fig. 1(b)).

Step 2: bitmap-based hash index. Let the selected anchor objects be $a_1, a_2, \dots, a_i, a_A$. For each anchor object a_i , we need to calculate a distance value r_i . Given an object $p \in P$, we convert it into an A -length bitmap where the i -th bit of the bitmap is defined in [1]:

$$BM(p)[i] = \begin{cases} i = 0 & \text{if } dist(a_i, p) \leq r \\ i = 1 & \text{otherwise} \end{cases} \quad (2)$$

Once the bitmaps for all anchor groups are generated, we need an index for query processing based on the bitmap information. In order to enhance the query processing performance, we use an algebraic coding-based hash index for retrieving anchor bitmap instantly. The proposed hash index uses single bit of the array as a coefficient of polynomial expression. And the hash table address is calculated based on the prime number that is close to the hash table size (Equation 3). By using this prime number we divide the bit value and the remainder is assigned as the hash table address. In equation 3, bit means the anchor bitmap and Hsize indicates the hash table size. Hence, we insert the anchors' bitmap to the hash table by applying the bitmap to the algebraic coding-based hash function. At query time, we also can directly access the hash table so that improve the query processing performance.

$$hashAddress = bit \% \left(1 + \sum_{m=1}^{2^{Hsize}} \left[\sqrt[Hsize]{Hsize} \left(\sum_{x=1}^n \left[\cos^2 \pi \frac{(x-1)! + 1}{x} \right] \right)^{-\frac{1}{n}} \right] \right) \quad (3)$$

2.2 Range Query Processing with Result Authentication

After generating data group, we generate a private data authentication index in order to provide privacy-preserving range query processing. Each data group is signed by the data owner using Condensed-RSA [7] with data ids within the group. By this means, the private authentication index is generated without revealing the partitioning information to unauthorized accesses. At query processing time, our method employs a novel technique for searching nearest data groups from a query point, in order to maximize the utility of the transformed data. In the literature, the Hamming distance measure has been employed for approximate NN search. Once the nearest anchor group is retrieved, the service provider sends the encrypted data within the group and its signature. Since the query user was given the transformation key and signature function from the data owner, he/she can decrypt the result data and generates signatures of them. If the generated signature is identical to the signature from the service provider, the client confirms that the query result is correct and genuine.

The query processing with query result integrity auditing is performed between two parties: a query user and a service provider. First, the query user transforms a query into bit-array $BM(q)$ and forwards it to the server. Upon receiving a query, the service provider retrieves the hash index and the private authentication index to return

an encrypted original data set for the query and their signatures. Finally, the user generates signatures of query results by using the RSA key sent from the data owner, and compares the signature with result signature. If generated signature is identical to the signature with the query results, the user can verify the correctness and completeness of query results.

3 Experimental Evaluation

In this section, we present the extensive experimental evaluation by comparing the performance of the proposed scheme with the existing work[3]. Our experimental environment includes Intel i3 3.10GH CPU with 4GB memory. We evaluate our scheme by using the real dataset of Northern East America (NE) containing 119,898 point of interests (POIs).

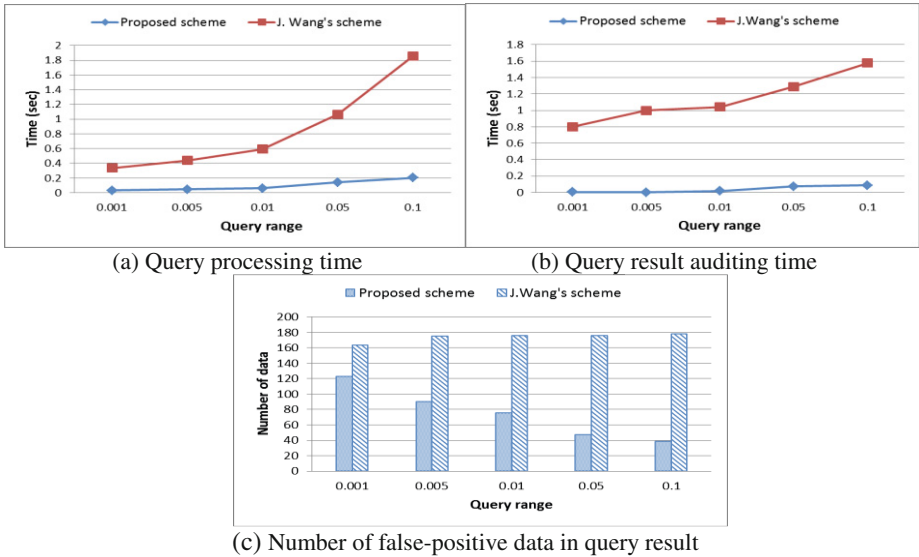


Fig. 2. Query performance with varying query ranges

Figure 2(a) describes the query processing times with varying query range. When the query range is set to 0.05%, the query processing time of our algorithm is 0.048 seconds whereas the existing scheme requires 0.6 seconds. From the result, it is proven that our algorithm outperforms the existing work up to 15 times, in terms of query processing times. This is mainly because our algorithm reduces the data transmission and verification overheads by using bit operation which easily calculates candidate anchors within query range. Also, the hash-based signature index affects to shorten the query processing time.

Figure 2(b) shows the query result auditing time with varying query range. For auditing the query result, the client decrypts the result data and generates a signature for comparison. Therefore, the performance of this term is closely related to the

number of result data. As shown in the figure, the auditing time is increased as the query range increases. In addition, our scheme reduces the number of false-positive data as shown in Figure 2(c). In this figure, our algorithm reduces about 20~70% of the number of false positive data overall query ranges. This is because our algorithm generates data groups based on the efficient data partition policy, such as anchor selection with data distribution and anchor merge/split, whereas the existing scheme clusters them solely based on the data values. Therefore, our scheme is stronger against the data group estimation attacks than the existing scheme.

4 Conclusion

In this paper, we proposed a privacy-aware query authentication scheme which guarantees the data confidentiality and the query result integrity of sensitive data. Through performance evaluation, it was shown that our scheme outperforms the existing method while providing similar performance in returning the number of false positives. As a future work, we will extend our scheme to support various query types, e.g., k-NN, skyline queries.

Acknowledgement. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2014065816).

References

1. Yiu, M.L., Assent, I., Jensen, C.S., Kalnis, P.: Outsourced similarity search on metric data assets. *IEEE Trans. on Knowledge and Data Engineering* 24(2), 338–352 (2012)
2. Sacharidis, D., Mouratidis, K., Papadias, D.: k-Anonymity in the Presence of External Databases. *IEEE Transactions on Knowledge and Data Engineering* (2010)
3. Wang, J., et al.: Bucket-based authentication for outsourced databases. *Concurrency and Computation: Practical and Experience* (2010)
4. Mykletun, E., Narasimha, M., Tsudik, G.: Authentication and integrity in outsourced databases. *Journal ACM Transactions on Storage (TOS)* TOS 2(2) (2006)
5. Hore, B., et al.: Secure multidimensional range queries over outsourced data. *The International Journal on Very Large Data Bases* 21(3) (2012)
6. Balpande, S., et al.: Data integrity and confidentiality in outsourced database. In: *International Conference & Workshop on Recent Trends in Technology* (2012)
7. Merkle, R.C.: A certified digital signature. In: Brassard, G. (ed.) *CRYPTO 1989*. LNCS, vol. 435, pp. 218–238. Springer, Heidelberg (1990)

Preventing Private Information in Secure Dissemination

Hye-Kyeong Ko

Division of Computer Engineering, Sungkyul University, Anyang-city, South Korea
hkko@sungkyul.ac.kr

Abstract. Secure dissemination of XML document is becoming a crucial requirement for many Web-based applications. By secure dissemination, we mean that the delivery of information to users must obey the access control policies. In this paper, we present an approach for the secure dissemination of web contents for users. Our approach uses a labeling scheme that protects the private information of web contents for users. Our experiment results show that the proposed approach is efficient protecting the private information.

Keywords: Security, Data Encryption, Access method, Secure dissemination.

1 Introduction

The success of the Web as a platform for EC and information dissemination has brought an increasing awareness of the fact that document exchange on the Internet should meet precise security requirements such as fine-grained authenticity, secrecy, and access control involving data units at the level of granularity stipulated by the communicating parties [1], [4]. In the Web environment, eXtensible Markup Language (XML) is rapidly becoming the standard for data representation and exchange. Today, more and more applications that utilize XML as the primary data format are being deployed. Generally, the content may have private information that needs protection. For example, if the content is disseminated carelessly, users can infer more information from the disseminated data in terms of inference [2], [8].

The rest of the paper is organized as follows. Section 2 surveys related work. Section 3 presents the principal techniques for the proposed secure dissemination service. Section 4 presents the results of our experiments, and we conclude in Section 5.

2 Related Works

To secure of an XML document, the XML Encryption Working Group of W3C develops a process for encrypting/decrypting XML documents and XML syntax used to represent the encrypted information that enable an intended user to decrypt it [13]. W3C XML Encryption [13] is only capable of encrypting full subtrees, while XML access control can remove sensitive material from the middle of the tree. In W3C XML Encryption, if the contents overlap, the same portions of the XML document could be re-encrypted for multiple users (called "super-encryption"). Access control concerns with who can access which information under what circumstances. XML

access control refers to the practice of restricting access to parts of XML data to only authorized subjects [1]. XML pool encryption approach [5] is able to hide the size and the existence of encrypted contents. In XML pool encryption, nodes containing sensitive information are selected, and are moved into a pool and encrypted. A number of security models have been proposed for XML [2], [3]. Under information push, the system periodically broadcasts documents to users, rather than sending them upon request [3]. The security views approach has recently been proposed in [4], [7]. Sensitive data is effectively protected from access and potential inferences by unauthorized subjects, and authorized subjects are provided with necessary schema information.

3 Proposed Technique

3.1 Proposed Approach

The system will be capable of encrypting arbitrary parts of the content. The work reported in this paper builds on the access control mechanism for content by Christian Geuer-Pollmann in [5]. The idea behind XML pool encryption is to remove private nodes from the tree and encrypt each private node individually. These encrypted nodes are stored in encrypted node sets. After decryption, each node can find its way back to its appropriate location in the content and be reconstructed correctly.

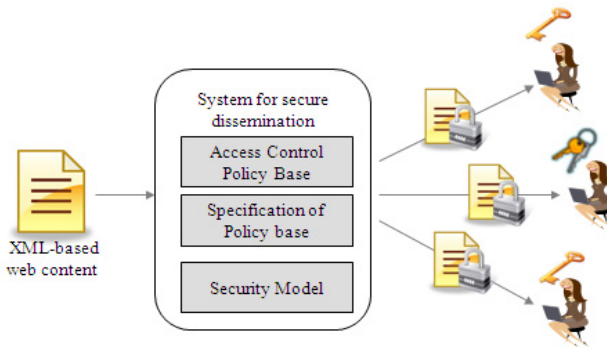


Fig. 1. Example of secure dissemination framework

The overall framework is depicted in Fig. 1. In the subscription phase, a user can be assigned rights, the service returns specific information to the user, which is required to decrypt parts of the XML-based web content source according to the user access rights.

Example 1. News Company provides content to subscribers. Two kinds of content are disseminated: paid content and free content. The paying subscribers can access paid content. The payment records identifying paying subscriber are managed by the company.

The game content in sports, and the story and the music content in culture are all paid content. Subscriber A paying for sports content can access both free and sport content.

Example 1 presents an example of encrypted dissemination of XML-based content. A pool of encrypted nodes (paid content) and unselected nodes (free content) are disseminated to multiple subscribers. Each subscriber can decrypt the content, depending on the access rights.

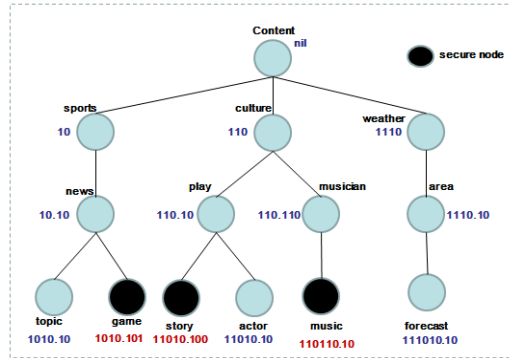


Fig. 2. Example of web content with labeling

3.2 XML Node Labeling

To decrypt the encrypted node, the system requires the location information of the decrypted node; i. e., where the node was in the original content. An efficient labeling scheme is very much required for identifying the location of a node. In this paper, we exploit an IBSL [9], which takes advantage of the lexicographical order of binary strings. The labeling procedure separates public and secure nodes. When the encrypted nodes are removed from the original tree, an adversary can guess the encrypted node location through labeling information and make good assumptions about the structure of the plaintext content. Algorithm 1 gives the details of the operation to label of public nodes with IBSL. In Algorithm 1, if the node is the root node, nil is assigned at the node (Lines 2 to 3 in Algorithm 1). For the child node of the root node, 10 is assigned at a first child for the second child (Lines 4 to 8 in Algorithm 1). Finally, the label of each sibling node is the *n.self label* concatenated with the *parent.label* (Line 10 in Algorithm 1). This method is applied until the full tree is traversed. After the traversal, each node in the tree has a unique value, i.e., a full label of public node. The delimiter "." is employed to assist subjects in figuring out the relationship (i.e., parent-child relationship) between nodes. For example, by looking at node 110.100, one realizes that it is a child of node 110. Algorithm 2 can help to label the secure node efficiently. Algorithm 2 can always label a new binary string between lexicographically ordered binary strings. The label length increases by one bit for the labeling by Algorithm 2. After labeling the plaintext document, the pruning procedure removes secure nodes from the labeled plaintext document. The pruned document is called public document. In fig. 2, the three terms (game, story and music) all refers to

the same node, while it is in different states of the pool encryption procedure and the pool decryption procedure. Besides the basic functionality of being able to reconstruct the document, the system should prevent information leakage to the subject as good as possible. After decryption of the document, the subject has access to the labels of all public nodes and decrypted nodes.

Algorithm 1. Assign label of public node

Input: each public node n in the XML document

Output: label (N)

```

begin
for ( $i = 0$   $i < childnum$   $i++$ ) do
if ( $n$  is the root node) then
  label( $N$ ) = nil;
else if ( $n$  is a child of root node) then
  parent = root.childnum;
   $n.self\ label[1] = 10$ 
   $n.self\ label[i] = 1 \oplus n.self\ label[i-1]$ ;
  label ( $N$ ) =  $n.self\ label[i]$ ;
else  $parent.childnum++$ ;
  label( $N$ ) =  $parent.label \oplus delimiter \oplus$ 
   $n.self\ label[i]$ 
end if
end for
return label ( $N$ )
end

```

Algorithm 2. Assign label of secure node

Input: N_left , N_right

Output: label (N_c)

```

begin
if ( $N\_left$  is empty, but  $N\_right$  is not empty)
then
   $N_c = N\_right \oplus 0$ ;
else if ( $N\_left$  and  $N\_right$  are not empty) then
if ( $len(N\_left) \leq len(N\_right)$ ) then
   $N_c = N\_right \oplus 0$ ;
else if ( $len(N\_left) > len(N\_right)$ ) then
   $N_c = N\_left \oplus 1$ ;
else ( $N\_left$  is not empty, but  $N\_right$  is empty)
   $N_c = N\_left \oplus 1$ ;
end if
return label ( $N_c$ )
end

```

The node reconstruction procedure takes the labeled public document and the decrypted nodes as input and inserts the decrypted nodes into the appropriate locations. In order to reconstruct the decrypted nodes, the reconstruction procedure has to identify which node inside the decrypted document is the parent node of the decrypted node and if that parent node already has child nodes, and determine whether some of the decrypted node's siblings are not siblings, and must be turned into children of the decrypted node.

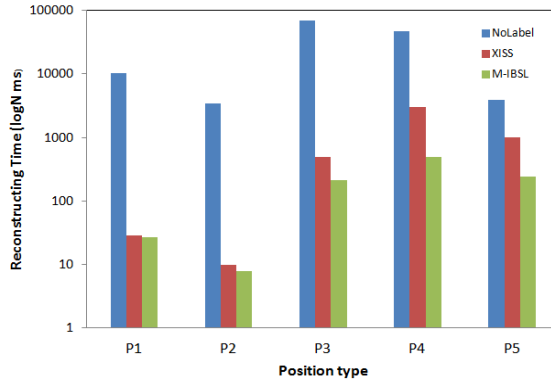
4 Performance Evaluation

We conducted experiments to evaluate and compare the performances of the three schemes, namely NoLabel (baseline), XISS [10], and M-IBSL, used in XML pool encryption [5]. The experiment is against a DOM representation of the baseline without a labeling, namely NoLabel. M-IBSL scheme is implemented using Java 2 and XML Security Suite. Experiments were carried out on a 3.40GHz Pentium processor with 4GB of RAM running Windows 7. In selecting nodes to be encrypted, XPath was used. We conducted experiments 20 times to obtain small confidence intervals.

Table 1 present the XPath expression used to represent nodes to be encrypted.

Table 1. Locations of node to be encrypted

XMark
P1 //item/mailbox/mail/date
P2 /africa/item/*
P3 parlist/listitem/text
DBLP
P4 title//*/sub
P5 sub/sup/tt/ref

**Fig. 3.** Reconstructing time of decrypted nodes (0.1MB)

In Fig. 3, M-IBSL outperformed XISS, on all location types for an XML document with 0.1MB of the XMark and DBLP datasets. In DBLP dataset, the number of sibling node increases when the width of the XML document is increased.

The results demonstrate that the number of nodes to be encrypted is related to the number of compared nodes in reconstructing location, and this affects the location reconstructing time. In M-IBSL, a label is not compared with labels of other nodes because a child node is labeled by extending the parent's label to represent the structural information of the XML document. This labeling scheme supports the easy identification of relationships among nodes. In comparing location reconstructing times for the three encryption schemes, the number of encrypted nodes used in searching for a location in the XML document was determined and the location reconstructing time recorded.

5 Conclusion

In this paper, a labeling scheme for the secure dissemination of web contents for users is proposed. When the encrypted nodes are removed from the tree, an adversary can guess the location of an encrypted node from the labeling information. To solve this problem, we presented an enhanced labeling scheme, named M-IBSL, which separates public and private nodes. The results of the experimental study are

presented for evaluating the performance of the M-IBSL and XISS. The M-IBSL is superior to XISS in terms of the number of nodes used to reconstruct for the proper location of a node, and the number of nodes used to search for the proper location of a node, and the location reconstructing time.

In the future, we will investigate how to reduce the label size and process the internal node encryption.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A3051552), South Korea.

References

- [1] Bertino, E., Carminati, B., Ferrari, E.: Securing xml documents with author-x. *IEEE Internet Computing* 5(3), 21–31 (2001)
- [2] Bouganim, L., Ngoc, F.D., Pucheral, P.: Client-based access control management for xml. In: *Proceedings of the Very Large Data Bases*, pp. 84–95 (2004)
- [3] Damiani, E., De Capitani di Vimercati, S., Paraboschi, S., Samarati, P.: Securing XML documents. In: Zaniolo, C., Grust, T., Scholl, M.H., Lockemann, P.C. (eds.) *EDBT 2000. LNCS*, vol. 1777, pp. 121–135. Springer, Heidelberg (2000)
- [4] Fan, W., Chan, C.-Y., Garafalakis, M.: Xml querying with security views. In: *Proceedings of the ACM SIGMOD*, pp. 587–598 (2004)
- [5] Geuer-Pollmann, C.: Xml pool encryption. In: *Proceedings of the 2002 ACM Workshop on XML Security*, pp. 1–9 (2002)
- [6] Kunda, A., Bertino, E.: An model for secure dissemination of xml content. *IEEE Transactions on Systems Mans and Cybernetics Part C: Applications and Reviews* 38(3), 292–301 (2008)
- [7] Kuper, G., Massacci, F., Rassadko, N.: Generalized xml security views. In: *Proceedings of ACM SIGMOD*, pp. 77–84 (2005)
- [8] Ko, H.-K., Kim, M.-J., Lee, S.: On the Efficiency of Secure XML Broadcasting. *Information Sciences* 177(24), 5505–5521 (2007)
- [9] Ko, H.-K., Lee, S.: A binary string approach for updates in dynamic ordered xml data. *IEEE Transactions on Knowledge and Data Engineering* 22(4), 602–607 (2010)
- [10] Li, Q., Moon, B.: Indexing and querying xml data for regular path expressions. In: *Proceedings of the ICDE*, pp. 361–370 (2001)
- [11] Steele, R., Min, K.: HealthPass: Fine-grained Access Control to Portable Personal Health Records. In: *Proceedings of the AINA*, pp. 1012–1019 (2010)
- [12] Schmidt, A., Wass, F., Busse, R.: XMark: A benchmark for xml data management. In: *Proceedings of the Very Large Data Bases*, pp. 974–985 (2002)
- [13] Imamura, T., et al.: W3C. Xml encryption and processing (2002)

Possibility of Using Embedded Sensors of Smart Devices for Augmented Reality Application

Ondrej Bilek and Ondrej Krejcar

University of Hradec Kralove, FIM, Center for Basic and Applied Research,
Rokitanskeho 62, Hradec Kralove, 500 03, Czech Republic
Ondrej.Bilek@uhk.cz, Ondrej.Krejcar@remoteworld.net

Abstract. Paper deals with use of embedded sensors of mobile devices. While describing a development of an example application using described sensors with emphasis on the widest range of sensors we would like to show possibilities of today smart devices to operate with augmented reality world. We are also focusing on current popular mobile platforms and development options for these platforms, with focus on using of embedded sensors. Most important issues are described with examples of specific procedures and algorithm samples of programming sensors on currently the most widely used mobile platform Google Android. Practical example application for augmented reality is using some described sensors and is based on some principles of augmented reality. There is described initial design of application, implementation with class descriptions and interesting algorithms and testing. Finally we end with testing, evaluation, comparison and conclusion.

Keywords: Sensor, Mobile, Device, Embedded, Augmented, Reality.

1 Introduction

As power and possibilities of mobile devices rapidly grown with massive extension of smart devices such as smartphones or tablet computers, many sensors were integrated to extend possibilities and mainly, to help users with common use of smart devices [1-4].

Smart devices also integrate functions of some other devices [5-8, 12]. Few years ago users need phone for calling or texting, camera for taking photos, MP3 player for listening music, GPS navigation or map to planning their trips [7,11], computer to opening documents etc. Nowadays, all these functions can be handled with single device, which offer great computational power with many possibilities in body of mobile phone or tablet computer. A brief comparison of supported and mandatory sensors is presented inside this paper on one case study example of augmented reality [1-4] application.

While several existing platforms on which the mobile smart devices are based, there is a one leading – Google Android [6, 12] which is described in this paper in more detail, as well as conditions for developers to use these embedded sensors are outlined. Google Android is the most used mobile platforms of these days and it's still

rapidly growing and expanding to other markets. Another great benefit of developing app for Google Android is multiplatform developer tools, relatively cheap app publishing on platform store and much information around web.

After theoretic part of a paper, the practical parts follows. First parts describes proposal of sample application and second describes implementation, functions and testing of sample application. Sample application uses principles of augmented reality, where view from device camera is used to include data from sensors, what can be enhanced to defect management system [2] or some kind of advisory solution [3].

2 Developing Mobile Application Using Embedded Sensors

The following section describes the basic features of the development environment (Software Development Kit - SDK) for major mobile platforms. Currently we can cover these ones: Google Android, Apple iOS, BlackBerry and Microsoft Windows Phone. This section describes the current situation in the late of 2013 and takes into account the current versions of operating systems, which are Android 4.2, iOS 6, BlackBerry 10 and Windows Phone 8.

2.1 Developing Platforms

We compares the possibility of developers for each platform. By developers the friendliest platforms are Android and BlackBerry, which do not burden developer's regular fees and no specific demands on equipment. Both platforms also use Eclipse as development environment. All platforms also cover the possibility for development of performance-intensive applications in C and C ++.

Students can be interesting for Microsoft approach that provides an opportunity for students to have a free developer account and access to a range of development tools under the DreamSpark program. Unfortunately, Microsoft puts great demands on computer equipment intended for development, where for all Microsoft development platforms should be used a computer running Windows 8 Pro and hardware virtualization. Similar restrictions also puts Apple, which supports development only on their computers running Mac OS X.

2.2 Embedded Sensors on Mobile Platforms

Mobile devices can support multiple sensors, but the operating system often supports only some of them, and furthermore, not all are mandatory for all devices with each operating system. The sensors support is the best on Android and BlackBerry platform.

For BlackBerry the broad support for sensors due mainly for supports of applications from Android, but it is not known which sensors are required (currently the only one device is on the market (Z10)). But we can expect a decent sensor support to facilitate routine work.

While Android supports all sensors, including the programming access to them, the equipment varies between devices. Especially cheap tablets are equipped with a

minimum of sensors - accelerometer, geomagnetic field sensor and a microphone. Geolocation using triangulation is supported by all devices. Usually mobile phones also always support the proximity sensor, GPS and camera. Conversely, environmental sensors (except the light sensor) and a gyroscope cannot be found in many devices.

iOS and Windows Phone are in the terms of sensors similar. They support only a few sensors, while a light sensor and proximity cannot be accessed programmatically. For iOS threatens absence GPS (iPod Touch), as well as for Windows Phone it is not mandatory a gyroscope sensor and geomagnetic field sensor. Both platforms also supports environmental sensors.

3 Proposal of Sample Application

Functions of sample application was defined as follows:

1. Show data from GPS and sensors of environment
 - a. (temperature, humidity, atmospheric pressure)
2. Show tilt of device towards water level and direction from north
3. Implement simple GPS navigation, this navigation should show distance and direction to coordinates, which should be user enterable
4. Application should run on majority part of existing smart devices with Google Android operating system

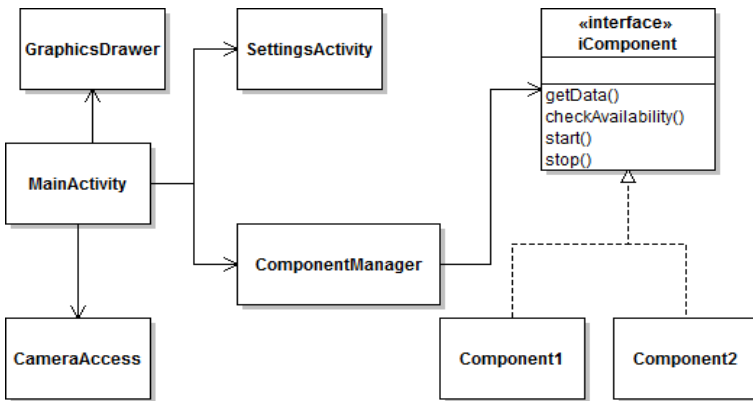


Fig. 1. UML Design of application structure

UML diagram of developing application structure can be seen on (Fig. 1). Proposal of structure application counts on component architecture. There is component manager, interface to define components and components itself. Each component should prepare some data to show to user. Component manager is associated to *MainActivity*, which performs operation of user interface and communicate with rest of application. Other classes are *CameraAccess* for accessing camera, *GraphicsDrawer* for drawing some nonstandard user interface artefacts and *SettingActivity* for application settings.

Sample application is using principles of augmented reality. Basic stone of application is real-time view from device's back camera. This view is enriched by overlay with sensor data. Firstly we focused on practical part in the sense of GUI design prior to final definition of UML and all functions as in classical way of application design. For this reason we are concerned on proposal of application.

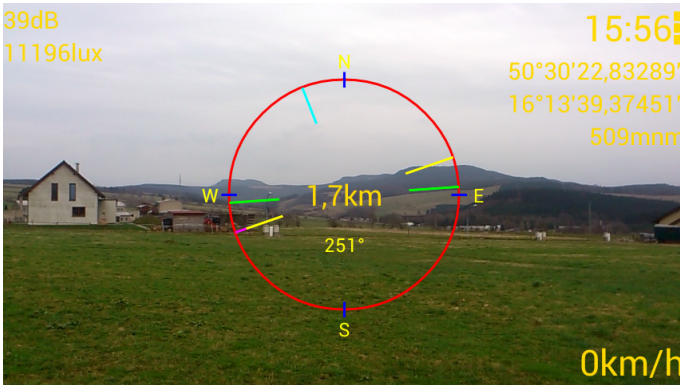


Fig. 2. Final GUI of developed augmented reality application

4 Implementation of Augmented Reality Application

This chapter describes implementation of final version of sample application for Augmented Reality. Design and structure of application was modified against proposal. Design of final application is shown on (Fig. 2) and its structure on attachment (Fig. 3).

Structure of application was modified in these parts:

1. Components were replaced by services. These services extends class *Service*, which is part of Android development kit. This guarantees single structure of services and possibility of sending messages
2. *ComponentManager* and *iComponent* was removed, because it is not needed
3. *SettingActivity* was replaced by fragments package. This package contains definitions of dialogs, which are used for settings and other

Application uses for communication between all parts sending of messages, so services sends messages while the class *MainActivity* receiving all messages sent to individual services and their data. On the service side sending is as follows:

1. Firstly, we declare the ability to sending messages to defined events (by *Intent* class instance). String which play as action is stored in a static variable of the public service.
2. Then the sequence of methods *putExtra(key, data)* attached to the message an individual data to be sent.
3. Finally, a message is sent by method *sendBroadcast()*. So the message is not directly addressed to *MainActivity* class - need to catch the message.

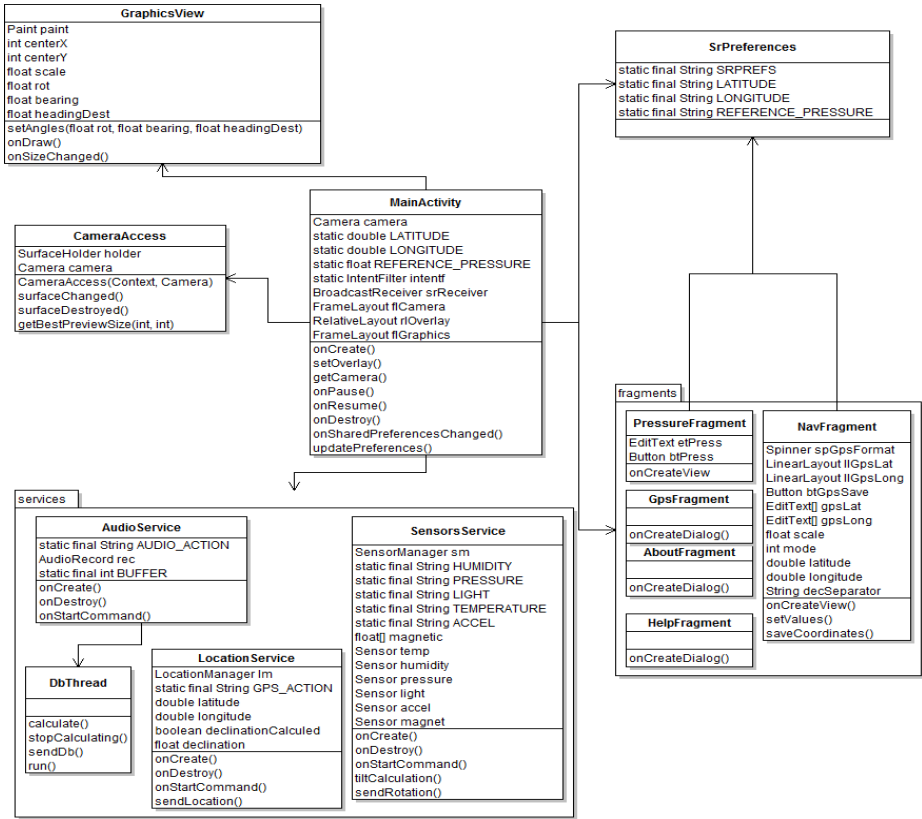


Fig. 3. Simplified UML design of final application

The procedure on the class *MainActivity* is a little more complicated. First, it is necessary to set an instance of *IntentFilter* that determines which messages are received. Action that identifies the message is gradually sets to this instance by calling *addAction()* method. Events are stored for each service as static variables. Receiving of messages is finished by setting of the recipient, which forms the interface instance created a *BroadcastReceiver* interface and *IntentFilter* is created by calling the *registerReceiver (BroadcastReceiver, IntentFilter)*.

5 Conclusions

Main goal of this work was to describe sensors, which are embedded in modern mobile devices, and create a sample application for augmented reality, which is using these sensors. Paper covers main information about sensors, beginning sensors descriptions and ending with practical use of these sensors. Adept of developing for mobile devices find there a possibilities of development tools and conditions on current popular mobile platforms with concern on supported sensors on these platforms. If developer is interested about programming applications which are using

embedded sensors for Google Android platform, he or she can find here instructions how to use sensors and read data from them on this platform.

Acknowledgement. This work and the contribution were supported by project “SP-103-2015 - Smart Solutions for Ubiquitous Computing Environments” Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic.

References

1. Daponte, P., De Vito, L., Picariello, F., et al.: State of the art and future developments of the Augmented Reality for measurement applications. *Measurement* 57, 53–70 (2014)
2. Kwon, O.S., Park, C.S., Lim, C.R.: A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality. *Automation in Construction* 46, 74–81 (2014)
3. Rusch, M.L., Schall, M.C., Lee, J.D., Dawson, J.D., Rizzo, M.: Augmented reality cues to assist older drivers with gap estimation for left-turns. *Accident Analysis and Prevention* 71, 210–221 (2014)
4. Dey, A., Sandor, C.: Lessons learned: Evaluating visualizations for occluded objects in handheld augmented reality. *International Journal of Human-Computer Studies* 72(10-11), 704–716 (2014)
5. Hajovsky, R., Pies, M.: Complex Measuring System for Longtime Monitoring and Visualization of Temperature and Toxic Gases Concentration. *Elektronika ir Elektrotechnika* 122(6), 129–132 (2012)
6. Behan, M., Krejcar, O.: Adaptive Graphical User Interface Solution for Modern User Devices. In: Pan, J.-S., Chen, S.-M., Nguyen, N.T. (eds.) *ACIIDS 2012, Part II. LNCS*, vol. 7197, pp. 411–420. Springer, Heidelberg (2012)
7. Benikovsky, J., Brida, P., Machaj, J.: Proposal of User Adaptive Modular Localization System for Ubiquitous Positioning. In: Pan, J.-S., Chen, S.-M., Nguyen, N.T. (eds.) *ACIIDS 2012, Part II. LNCS*, vol. 7197, pp. 391–400. Springer, Heidelberg (2012)
8. Cimler, R., Matyska, J., Balfk, L., Horalek, J., Sobeslav, V.: Security issues of mobile application using cloud computing. In: Abraham, A., Krömer, P., Snasel, V. (eds.) *Afro-European Conf. for Ind. Advancement. AISC*, vol. 334, pp. 347–358. Springer, Heidelberg (2015)
9. Penhaker, M., Darebnikova, M., Cerny, M.: Sensor Network for Measurement and Analysis on Medical Devices Quality Control. In: Yonazi, J.J., Sedoyeka, E., Ariwa, E., El-Qawasmeh, E. (eds.) *ICeND 2011. CCIS*, vol. 171, pp. 182–196. Springer, Heidelberg (2011)
10. Machacek, Z., Slaby, R., Hercik, R., Koziorek, J.: Advanced system for consumption meters with recognition of video camera signal. *Elektronika Ir Elektrotechnika* 18(10), 57–60 (2012)
11. Jancikova, Z., Kostial, P., Bakosova, D., Ruziak, I., Frydrysek, K., Valicek, J., Farakasova, M., Puchky, R.: The Study of Electrical Transport in Rubber Blends Filled by Single Wall Carbon Nanotubes. *Journal of Nano Research*, No. 16 21, 1–6 (2013)
12. Behan, M., Krejcar, O.: Modern Smart Device-Based Concept of Sensoric Networks. *EURASIP Journal on Wireless Communications and Networking* 2013(1(155)) (June 2013), doi 10.1186/1687-1499-2013-155, ISSN 1687-1499

Economic and Technological Aspects of Social Networks in European Business Sector

Petra Maresova

University of Hradec Kralove, Faculty of Informatics and Management,
Rokitsanskeho 62, Hradec Kralove, Czech Republic
petra.maresova@uhk.cz

Abstract. The Internet has become an inherent part of human life. The number of internet users has been constantly growing. Social networks are a significant tool for supporting the development of learning organization and knowledge management in companies. They become an important communication channel. The aim of this paper is to analyze the use of social network in European business sector. Technological and economic aspects of the use of social networks will be described. The basic methods used are analysis of the external environment and the subsequent SWOT analysis. The analysis shows that, given the corporate internet facilities in Europe, there is great potential for its use that is not currently used.

Keywords: Social network, technological environment, SWOT analysis, Europe.

1 Introduction

The Internet has become an inherent part of human life. The number of internet users has been constantly growing. According to the last surveys, almost 2 billion people worldwide uses the Internet, about half a billion in Europe. With the arrival of a new form of communication and sharing the content of the Internet, the so-called web 2.0, a new trend has become prominent, namely social networks.

Current companies realize the need to integrate tools using the concept of social networks in their business processes. Social networks are a significant tool for supporting the development of learning organization and knowledge management in companies. They become an important communication channel and take over a lot of functions previously provided by e-mail [1]. The term social business more and more frequently applies to the company capable of using the potential of social networks to the full [2].

The international comparison of using social networks by individuals with an access to the Internet from households shows graph no. The EU average is 46% (Fig.1).

The aim of this paper is to analyze the use of social network in European business sector. Technological and economic aspects of the use of social networks will be described. Moreover, in the end they will be summarized in the SWOT analysis.

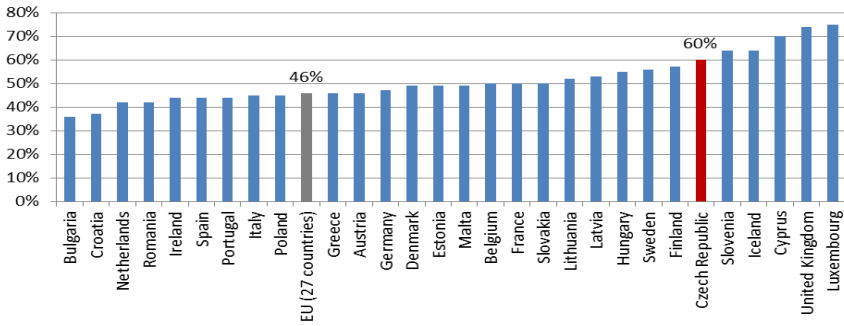


Fig. 1. Use of social network (in 2011, %), source: the author according to Eurostat

2 Methodology

The basic methods used are analysis of the external environment and the subsequent SWOT analysis. The external environment involves forces outside the company or, in this case, outside the whole ICT sector that can potentially influence the use of cloud computing. Analysis of the external environment and the characteristics of the technology has been described in detail in the context of previous work and internal research [2]. Given the scale of this paper is only the current situation of social network accessing described and then directly so-called SWOT matrix (Strengths, Weaknesses, Opportunities, Threats), where are summarized strengths, weaknesses, opportunities and threats of the given segment [3].

3 Technological Background of Social Network

The social network is an interlinked group of people. In a sense, social network is any group of people that communicates in various ways. Currently, there exist a lot of social networks, which vary by their purpose, their members' character, needs, and goals. Social networks exist inside and outside organizations, they are open to anybody or exclusive, i.e. open to a given community only. Social network can be defined as web based services that allow individuals to do the following:

- construct a public or semi-public profile within a bounded system
- articulate a list of other users with whom they share a connection, and
- view and traverse their list of connections and those made by others within the system.

The nature and nomenclature of these connections may vary from site to site [4]. Social network architecture is described in the figure no. 2.

There are so many variations of social networks that they can be utilized in almost any human activity. Concerning their respective focuses, they can be divided into several categories [7], [8], [9].

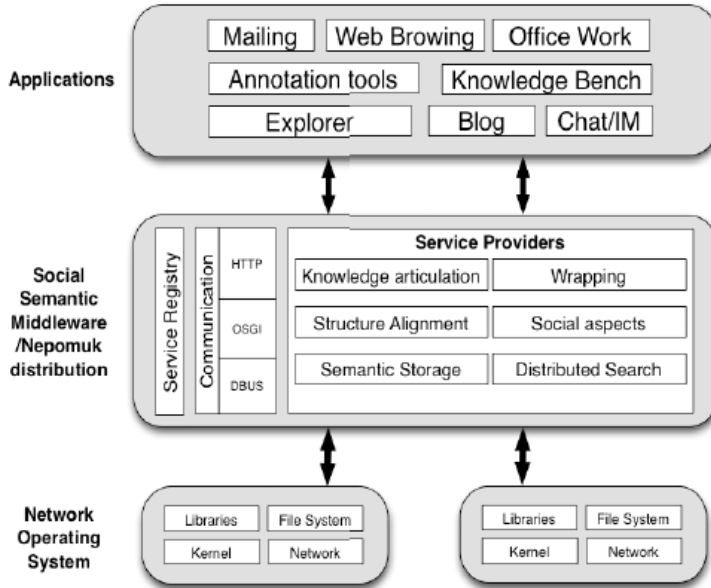


Fig. 2. Social Network Architecture, source: according to [5], [6]

General Social Networks

These networks can offer something to almost any user and above all they enable anybody to be registered. An example of such a social network is Facebook.

Specialist Social Networks

This group includes social networks that associate users interested and/or working in the same branch, industry, or any other activity, no matter whether they are professionals, students, or hobbyists. These networks are established around particular websites focusing on the topic in question. The basic types of branch social networks are as follows:

- Professional social networks – networks of this type associate professionals from the given area, they are only rarely anonymous or open to anybody, they are specifically designed for given professional or interest groups,
- Hobby social networks – this group includes social networks that unify hobbyist users interested in the given activity,
- Student social networks – these networks do not have to unify exclusively students as they focus on the study of one or more fields.

Among the currently most important social networks belong, for instance, Facebook, Twitter, Google+ and LinkedIn. Each network is specific and differs from others [10].

4 Use of Social Networks – SWOT Analysis in Europe Business Environment

The ensuing SWOT analysis provides a summary of benefits and risks in the company's use of social networks. Factors in the SWOT analysis include items that characterize the economic and technological background of social networks in the Europe and on what it brings to those companies that use it (table no.1).

Table 1. SWOT analysis of social networks

Strengths	Weaknesses
<p>Growing accessibility of high quality internet connectivity in households and companies,</p> <ul style="list-style-type: none"> • growing number of PC-equipped households, • growing ICT literacy of the population [12], • companies realize the importance of social networks, • network of contacts surpassing geographical and cultural differences, • it is possible to target customers who show up near the company' location, • procuring information on a specific subject is faster and easier. 	<ul style="list-style-type: none"> • A large amount of companies that use social networks for advertising purposes, • strong competition means it is necessary to stress innovation. • Problem of choosing information and functions of social networks so that the company gets the best of it (it should not be at the expense of employees, etc.). • It is necessary to monitor changes and new possibilities of social networks as they continually change and improve, which means that companies have to adapt effectively and promptly. • It takes time to maintain the social networks updated and full of attractive information. • Some marketing functions are limited to paid services. • Social networks are limited to those from 13 years of age (this is often trespassed).
Opportunities	Threats
<ul style="list-style-type: none"> • Growth in the use of social networks worldwide, • decreasing number of networks that dominate global market – a better opportunity for companies to communicate with more users in the same place. 	<ul style="list-style-type: none"> • Different legislation and regulations in various countries – problem of data protection, • little adaptation to the culture and ways of communication in other countries. • Identity theft,

Table 1. (continued)

Opportunities	Threats
<ul style="list-style-type: none"> • Participation in social networks means that users and friends can become company product sellers, • environment for making new contacts with business partners, or for finding skilful employees, • getting feedback from non-clients, which is valuable information and open new opportunities to attract new clients [11], • professional user base on some social networks means great opportunities for B2B marketing, • increased loyalty of customers to the brand, • possibility to create the company's "personality", • possibility to address a lot of potential customers at relatively low costs. 	<ul style="list-style-type: none"> • possible abuse of personal data, • too many channels of communication through social media in one company can hinder potential customers' orientation, • implementation of a social network can have negative impact on the employees' labour productivity, • dissatisfied customers can spread their experience among a large number of users, • risk of accidental sharing of sensitive information or intellectual property, • neglecting the company website in favour of social networks, • competition can harm the company by e.g. publishing negative reviews.

The above mentioned analysis shows some recommendations. Companies have to consider what information to share with users, how to communicate and what rights they provide their customers with. The threat of losing or abuse of data is palpable. It is also important for every company to choose a sufficiently impressive and personalized approach. Many social network users are inundated with advertisement and necessarily have to choose which of them to respond to.

Social network company profile administration must be done conceptually. First, target groups should be determined, then, based on the features of these groups, the company chooses the most suitable social network to approach them. In order to use the full potential of social networks, it is necessary to understand how they work and how they can be used. The choice of suitable marketing strategy is vital as the potential audience is plentiful and a potential mistake may significantly damage corporate image.

5 Conclusion

The aim of this paper is to analyze the use of social network in European business sector. From that point of view, the onset of social networks made companies revise their marketing strategies and adjust business to people. The age of participative economics, when customers are not satisfied with passive product consumption and

want to be involved, has just started. Apart from that, the companies have higher and higher demands. They want openness, relaxed attitude, originality, and diversification. Social marketing platforms provide tools for immediate sharing of information, photographs, videos, events and comments. Social networks also made companies to move investments from huge mass campaigns to smaller and more segmented ones. Marketing influences by means of social networks above all current customers. New customers are attracted more slowly, in particular by means of the current customers' reference sharing. Owing to a simple way of communication and sharing, it is possible, in case the company did a good marketing job, to leave a large part of promotion to the users themselves. It is not possible to ignore individuals because they can spread their negative experience within minutes among hundreds of people and households. Social networks provide companies with a huge amount of the so-called Big Data for analysing. Nevertheless, it is not easy to understand the data correctly and draw relevant conclusions. It is therefore necessary to educate the employees and keep up with the latest trends.

Acknowledgement. This paper is published thanks to the support of the internal projects of University of Hradec Kralove: Economic and Managerial Aspects of Processes in Biomedicine.

References

- [1] Vávra, T.: Sociální sítě ve firmách pomáhají sdílet informace Social networks in companies to help share information. In Computerworld (2012), <http://computerworld.cz/internet-a-komunikace/socialni-site-ve-firmach-pomahaji-sdilet-informace-49385> (accessed June 22, 2014)
- [2] Marešová, P.: Potential of ICT for Business in the Czech Republic. Professional Publishing (2013)
- [3] Marešová, P., Kuča, K.: Technological environment and SWOT Analysis of Cloud Computing in Europe. International Journal of Information Technology & Computer Science 18(1), 57–77 (2015), http://www.ijitcs.com/volume%2018_No_1/Petra%201.pdf
- [4] Boyd, D.M., Ellison, N.B.: Social network sites: Definition, history, and Scholarship. Journal of Computer-Mediated Communication 13(1), 210–230 (2007)
- [5] Clementking, A., Venkateswaran, J.V.: Study On 4G Communication Architecture Components For Social Networks. International Journal of Reviews in Computing 12(9), 71–75 (2010), <http://www.ijric.org/volumes/Vol2/9Vol2.pdf>
- [6] Choi, S., Kag, G., Shin, A.: Class of Adaptive Hybrid ARQ Schemes for Wireless Links. IEEE Transactions on Vehicular Technology 50, 777–790 (2001)
- [7] Paulraj, A.J., Gore, D.A., Nabar, R.U., Bölcskei, H.: An overview of MIMO communications - a key to gigabit wireless. Proceedings of the IEEE International Journal of Reviews in Computing 92(2), 198–218 (2004)
- [8] Kallel, S., Bakhtiyari, S., Link, R.: An Adaptive Hybrid ARQ Scheme. Wireless Personal Communications 12(3), 297–311 (2000)
- [9] Molnár, Z.: Jak využít sociální sítě v podnikání How to use social networks in business. In Systémová integrace (2011) <http://www.cssi.cz/cssi/jak-vyuzit-socialni-site-v-podnikani> (accessed June 22, 2014)

- [10] Lauschmann, J. Znáte 5 + 1 největších sociálních sítí. Do you know the 5 + 1 largest social networking? (2012), <http://cdr.cz/clanek/nejvetsi-socialni-site-dneska> (accessed June 22, 2014)
- [11] Marešová, P., Kuča, K.: Assessing the Effectiveness of Cloud Computing in European Countries. In: Proceedings of the 4th International Conference on Computer Engineering and Networks, pp. 769–775 (2015)
- [12] Bartuskova, A., Krejcar, O., Selamat, A., et al.: Framework for Managing of Learning Resources for Specific Knowledge Areas. In: 13th International Conference on Intelligent Software Methodologies, Tools, and Techniques (SoMeT). Frontiers in Artificial Intelligence and Applications, pp. 565–576 (2014)

Decision Making Criteria for Cloud Computing Deployment

Petra Maresova

University of Hradec Kralove, Faculty of Informatics and Management, Rokitanskeho 62,
Hradec Kralove, Czech RepublicAffiliation
petra.maresova@uhk.cz

Abstract. Cloud computing represents a potentially very effective solution to current problems linked to the need to decrease business expenditures. The solution is based on sharing the same solution by multiple users as well as on an opportunity of choosing a custom-made solution that would suit any office's particular needs. Use and deployment of this technology in Europe (compared to the USA) is low. The main reasons include lack of knowledge of the technology benefits and decision making criteria, which can help to recommend cloud computing deployment. The aim of this paper is to describe to proposed decision making model of cloud computing. This model is based on a quantitative survey among SMEs in Europe. Its use has also been verified in enterprises.

Keywords: cloud computing, decision making criteria, deployment, business.

1 Introduction

The importance of cloud computing is acknowledged at national and international levels, both in public and private sectors. However, its implementation in European countries is considerably lower contrary to assumptions [1], [2]. Among the frequently mentioned cloud computing benefits for companies belong lower costs of ICT departments, an ability to flexibly change requirements for provided services and access to data from anywhere. There are, however, a lot more benefits. According to many studies the further benefits are:

- utilization of network services and the Internet, the speed of implementation,
- faster access to the market,
- innovative approach,
- elasticity and scalability,
- safety and energy savings.

On the other hand, apart from benefits, cloud computing can pose certain risks, too. A lot of authors divide risks linked to cloud computing into a lot of categories, e.g. Azarnik [3] lists technology development risks, functionality risks, political risks, project risks and financial risks. Last but not least, Shayan [5] divides risks into two

categories, tangible risks and intangible risks. Tangible risks can easily be understood and the user can decide whether or not to run the risk. Tangible risks include access, availability, long-term viability, infrastructure, integrity, vendor lock-in and investigative support. On the other hand, intangible risks are not transparent to the user and only the provider knows them. Intangible risks include access, service continuity, confidentiality, protect data, physical network infrastructure, storage mechanism, government investigation and location of data [6].

The lack of knowledge of the technology characteristics is main reason, why some companies refuse cloud computing deployment [7]. These companies attach great importance to the above mentioned risks. The aim of this paper is to describe to proposed decision making model of cloud computing. This model helps clarify the possibility to use cloud computing in relation to business data nad risk factors.

2 Method

A method of retrospective analysis of documents is used to characterize the basic concepts and model of cloud computing. For the analysis of assumptions deployment of cloud computing in European countries the research among small and medium companies was carried out. Then decision making criteria for cloud computing deployment were proposed and verified.

3 Cloud Computing – Theoretical Background

The basic characteristic of cloud computing infrastructure is accessibility to individual users or companies. The service provider must be an external entity. The same applies to the data centre from which the service is accessible [8]. In relation to this model the so-called Virtual Private Cloud is sometimes mentioned. It is a solution that enables any organization to create a group of separate ICT tools in public space. According to Mell and Grance [8]. Private Cloud is every model which makes services accessible to the only entity. The whole infrastructure can be administered by a third party or the given organization. On the other hand, Armbrust [9] can see the essence of the private solution in internal data centres, which make their services accessible exclusively to their own organization, not to the public or to other entities. are distinguished different service models of cloud computing.

Infrastructure as a Service (IaaS) uses virtualization technology to allow several virtual systems (referred to as virtual machines) to operate on top of a single physical hardware infrastructure in an isolated manner. The key software module in virtualization is the hypervisor that manages and organizes the virtual resources on the physical hardware (memory, processors, storage). In this category, cloud providers can deliver on-demand virtual machines with configurable resources.

Platform as a Service (PaaS) includes all the features provided by IaaS, but in this case the user is able to use the provider's system platform. PaaS allows clients to

develop their own system using the platform tools, without having to install and maintain these tools themselves. In this category, users obtain access to a specific OS (e.g., a version of Windows or Linux) and associated tools (e.g., SQL Server, MySQL, Apache web server, etc.) [9]. Software as a Service (SaaS) - this type of computer cloud eliminates the need to install and run software applications on the client's local computers. With SaaS, cloud providers install, manage, and operate the software application, and the user has neither knowledge nor control of the underlying infrastructure. With this type of cloud service, the end-user has the least flexibility but the cost is dramatically lower [10].

Proposed decision making criteria for cloud computing deployment described below are focused on Infrastructure as a Service (IaaS).

4 Decision Making Criteria for Cloud Computing Deployment – Case Study

In August 2013, a questionnaire survey was carried out in the Czech Republic (the chosen European country). Its topic was „Utilization of ICT in Czech companies“. This survey's aim was to determine current attitudes of companies to using technologies in order to support financial management, customer relation management (CRM) and cloud computing. The total of 200 questionnaires was collected. The respondents were managers and managing directors of companies, or heads of ICT departments. Respondents indicated as the most significant criteria for successful utilization of cloud computing, scalability, elasticity and customization. Problem of data security is dismissed as less significant in cloud computing. It becomes a problem only when the respondents are asked about risks of linked to the implementation of cloud computing. Some managers perceive following obstacles:

- data security,
- dependence of operation on internet access,
- insufficient knowledge of cloud computing and its possibilities,
- low priority of perceiving cloud computing as necessary and the related lack of time for analysing strengths and weaknesses,
- financial burden.

The main reason for low utilization of cloud computing is the fact that companies do not perceive the need to implement this technology 68.5% and do not have sufficient knowledge of its possibilities (14.5%). The proposed model helps clarify the possibility to use cloud computing in relation to business data and risk factors.

4.1 Criterion Questions for the Decision of Cloud Computing

Criterion are divided into three groups: entry information, Company's strategic management and Technical area (table no. 1).

Table 1. Criterion questions for the decision of cloud computing

Entry information	Value
Field of business Size of organization Geographical location of subsidiaries A number of full-time employees A number of part-time employees Field work Online business Type of business relationship Fluctuation of orders during the year What data volume with respect to the whole cannot be transferred on a third party (consider possible barriers: customer's requirements, internal regulations)	CZ NACE small/medium/large CR, Europe, USA, ASIA, in one place in % out of the whole company's employment rate in % out of the whole company's employment rate in % out of all activities yes/no B2B, B2C, B2G, B2A in % out of all activities in % out of the whole data volume
Company's strategic management	
Is your present IT capable of beating its competition or bring such savings you need? (Do you consider criteria such as speed of service supply, price or quality?) Can your IT adjust to the company's expansion, particularly with respect to its needs to store ever growing data volumes and growing computer demands? (Do you consider criteria such as data storage costs, renewal after the breakdown and provision for the continuity of business?) Do the investments into IT help you to differ from your competition, i.e. do you use quite specific solutions to your business or do you, on the contrary, use highly standard solutions? (Do you consider criteria such as an investment share into the standard IT solutions?)	scale 1-4 scale 1-4 scale 1-4
Technical area	
Is the whole day's loading of a server constant from the company's point of view? Is there a situation when I need to react to the current need of loading immediately? How often do you renew HW? How much data do I need to backup during the year? (GB) Do you request monitoring? Which distribution channel are you thinking about to use? When was your last investment into your IT? At what price?	scale 1-4 scale 1-4 less than once in three years/ 3-5 years/ in five years and more TB Yes /no PaaS/ SaaS/IaaS less than two years ago/ 2-4 years ago/ five and more years ago Czech Crown

Source: own processing according to:[5]

With respect to the first set of questions, it can be said that if there is a company which has more subsidiaries in different locations, its employees do their work outside the company or at home, no investment into the IT system have been recently made, cloud computing can be definitely used. Furthermore, it is important to know how the internal regulations are set with respect to the handling of data, whether companies work with sensitive data of their clients and whether it is necessary to solve the question of legislative regulations connected with personal data protection in different countries. This is important for a choice of the model implementation. The questions connected with the company's strategic management aim at a good description of flexibility of company's IT equipment in relation to its strategic targets and their changes. The last set of questions focuses on the technical requirements such as server performance, requirements of loading changes or data storage. At this stage these requirements are only indicative and they are fully discussed in the next step.

5 Conclusion

Since 2009, cloud computing technology has been getting to the foreground. A rapid growth in its utilisation is expected from 2014 onwards. The European Union supports this technology by means of a strategy aiming to encourage the use of cloud computing and create unified rules for its utilisation by European companies. Cloud computing utilisation may help create a competitive advantage. However, its implementation in European countries is considerably low.

The aim of this article therefore was to characterize the criterion model of cloud computing with respect to its use in entrepreneurial practice. The model is based on the conditions of the European business environment. After overcoming the initial barriers and concerns about data safety, Internet connection failures, and data migration, the new technology brings about a lot of benefits.

Acknowledgement. This paper is published thanks to the support of the internal projects of University of Hradec Kralove: Economic and Managerial Aspects of Processes in Biomedicine.

References

- [1] European Commission. The Future of Cloud Computing, Opportunities for European Cloud Computing beyond 2010 (2010), <http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-report-final.pdf> (accessed December 4, 2014)
- [2] Azarnik, A., Shayan, J., Alizadeh, M., Karamizadeh, S.: Associated Risks of Cloud Computing for SMEs. *Open International Journal of Informatics* 1, 37–45 (2012)
- [3] Shayan, J., Azarnik, A., Chuprat, S., Karamizadeh, S., Alizadeh, M.: Identifying Benefits and risks associated with utilizing cloud computing, San Francisco, USA (2013)
- [4] Bandura. Fearful expectations and avoidant actions as coefficients of perceived self-inefficacy (1996)

- [5] Marešová, P.: Deployment of Cloud Computing in Small and Medium Sized Enterprises in The Czech Republic. *E+M Ekonomie a Management* 17(4), 159–174 (2014)
- [6] Mell, P., Grance, T.: The NIST Definition of Cloud Computing. [online]. In National Institute of Standards and Technology (2009), <http://www.sciencedirect.com/science/article/pii/S0167923610002393>(accessed December 4, 2014)
- [7] Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: A View of Cloud Computing. *Communications of the ACM* 53(4), 50–58 (2010)
- [8] Peiyu, L., Dong, L.: The New Risk Assessment Model for Information System in Cloud Computing Environment. *Procedia Engineering* 15, 3200–3204 (2011)
- [9] Sarna, D.E.Y.: *Implementing and Developing Cloud Computing Applications*. CRC Press, New York (2011)
- [10] Bartuskova, A., Krejcar, O.: Evaluation Framework for User Preference Research Implemented as Web Application. In: Bădică, C., Nguyen, N.T., Brezovan, M. (eds.) *ICCCI 2013*. LNCS, vol. 8083, pp. 537–548. Springer, Heidelberg (2013)

Periocular Recognition Based on LBP Method and Matching by Bit-Shifting

So Ra Cho, Gi Pyo Nam, Kwang Yong Shin, Dat Tien Nguyen,
and Kang Ryoung Park*

Division of Electronics and Electrical Engineering, Dongguk University,
Seoul, Republic of Korea
{soracho, oscar1201, skyandla, parkgr}@dgu.edu,
nguyentien.dat@dongguk.edu

Abstract. Periocular recognition requires neither a high-resolution camera nor a zoom lens. It matches using the features extracted from the surrounding area of the eye. In addition, by using a wide-view camera, the constraints to users' head movement decrease. In this research, we newly propose a periocular recognition based on LBP method and matching by bit-shifting. Our research is novel in the following three manners. First, the iris and pupil region in the input eye image are detected. This allows the accurate eye region to be obtained for periocular recognition. Second, the feature code is extracted from the eye region with a local binary pattern method. Third, the proposed system performs matching by bit-shifting to prevent degradation to the matching accuracy caused by head movement. Experimental results show that the high accuracy of periocular recognition is obtained by the proposed method.

1 Introduction

Iris recognition uses iris patterns, and this is known to achieve a higher matching accuracy than other biometric methods [1]. One disadvantage to iris recognition, however, is the need for a high-resolution camera equipped with a zoom lens and near-infrared (NIR) illuminator to observe iris patterns. Moreover, in using a zoom lens, the viewing angle is small, which restricts users' head movement, making them feel uncomfortable. To solve this problem, periocular-based recognition methods have been studied [2-5]. These methods do not require a high-resolution camera with a zoom lens because they match the features extracted from the surrounding area of the eye. In addition, because they use a wide-view camera, the constraints to a user's head movement decrease. Most previous research relied on the eye position provided or manually labeled for cropping the eye region and for normalization [2][3]. Errors in eye detection were therefore not included among the errors in periocular recognition. In addition, most of previous methods do not consider the movement of eye region caused by head movement for matching. To overcome these problems, we newly propose a periocular recognition based on LBP method and matching by

* Corresponding author.

bit-shifting. The iris and pupil region in the input eye image are detected by circular edge detector. This allows the accurate eye region to be obtained for periocular recognition. The binary feature code is extracted from the eye region with a local binary pattern (LBP) method. The proposed system performs matching by bit-shifting to prevent degradation to the matching accuracy caused by head movement.

In the remainder of this paper, the proposed method is explained in Section 2. In Sections 3 and 4, we show the experimental results and conclusion, respectively.

2 Proposed Method

2.1 Overview of the Proposed Method

The overview of the proposed method is shown in Fig. 1. The eye region in an inputted facial image is detected using adaptive boosting (Adaboost) and sub-block-based template matching. Then, the center of the pupil is detected using circular edge detection (CED) and the eye region for recognition is defined. Periocular feature codes are extracted from the eye region with an LBP method. Subsequently, the extracted feature codes are matched to enrolled ones with bit-shifting in order to prevent a degraded matching accuracy caused by the movement of eye region due to head movement.

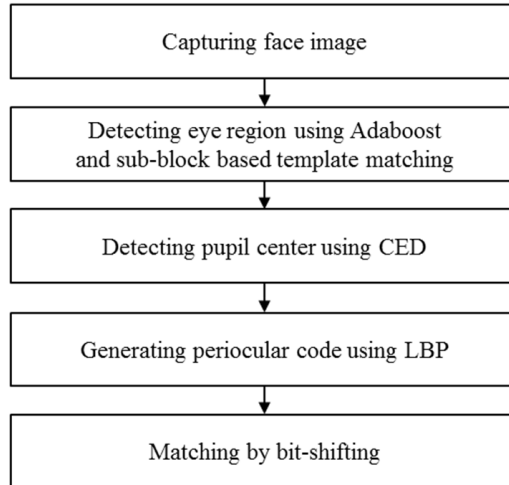


Fig. 1. Flowchart for the proposed method

2.2 Detection of the Periocular Region

Previous research relied on the eye position being either provided or manually labeled for cropping the eye region and for normalization [2][3]. Hence, errors resulting from eye detection were not included among the errors in periocular recognition. Therefore, such methods for manually detecting the eye region are not feasible in real-time

periocular recognition. The method we propose automatically locates the eye region for periocular recognition. To detect the periocular region, a rough region of interest (ROI) where both eyes are found is detected using Adaboost [6]. Then, the left and right eyes are detected using a sub-block-based template matching method within the ROI. Based on the positions for the left and right eye, the ROI for each eye is redefined, and the accurate pupil center is located with CED within the newly defined ROI. In general, each constituent in the eye region—such as the pupil, iris, sclera, and skin—has a different gray level due to the differences in reflectance. Because the pupil has a lower reflectance than other constituents, it is darker than the iris, sclera, and skin areas. Based on this, the pupil region can be defined as the area having the biggest difference between the average pixel value for the corresponding region and that of neighboring regions in the image. Thus, we detect the pupil using sub-block-based template matching based on five sub-blocks, as shown in Fig. 2 [7].

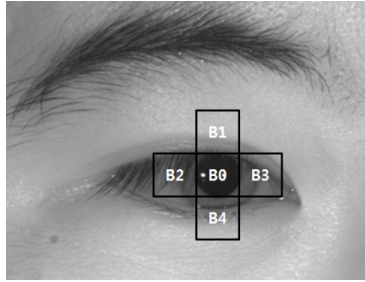


Fig. 2. Pupil detection using sub-block-based template matching

In Fig. 2, B0 is the average gray level for the center sub-block. B1, B2, B3, and B4 are the average gray level in the neighboring sub-blocks. In order to enhance processing speeds, we use an integral image method for calculating the average gray level [8]. The matching score for sub-block-based template matching is calculated using Eq. (1).

$$\text{Score} = (B1 - B0) + (B2 - B0) + (B3 - B0) + (B4 - B0) \quad (1)$$

In this manner, the score is calculated according to the position of each sub-block in Fig. 2 that is moved within the ROI by overlapping one pixel. Finally, the eye region is selected where the score is highest. To detect the periocular region more precisely, the pupil center in each eye is located with CED in the detected eye region with a sub-block-based matching method. CED detects the iris and pupil boundaries simultaneously, as shown in Eq. (2) [7]. (x_0, y_0) and (x'_0, y'_0) are the center coordinates for the iris and pupil, respectively. r and r' are the radii for the iris and pupil, respectively. In general, the iris is covered by the eyelid, so it is detected in the range $-45 \sim +30$ degrees and $+150 \sim +225$ degrees, whereas the pupil is detected in a range of 360 degrees. Fig. 3 shows a resulting image from the pupil region detected with CED. Based on the detected position of pupil, the accurate eye region for periocular recognition is defined.

$$\arg \max_{\substack{(x_0, y_0), r, \\ (x'_0, y'_0), r'}} \left[\frac{\partial}{\partial r} \left(\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}} \frac{I(x, y)}{5\pi r / 12} ds + \int_{\frac{5\pi}{6}}^{\frac{5\pi}{4}} \frac{I(x, y)}{5\pi r / 12} ds \right) + \max \left(\frac{\partial}{\partial r'} \int_0^{2\pi} \frac{I(x', y')}{2\pi r'} ds \right) \right] \quad (2)$$



Fig. 3. Detected pupil region with CED

2.3 Feature Extraction from the Periocular Region and Matching by Bit-Shifting

Features in the periocular image are extracted with LBP, which is a commonly used method for extracting local texture features. Features are represented as binary codes from eight bits using a 3×3 pixel mask that compares the gray level value from the center in the mask with those in the neighboring eight pixels [9]. If the gray level value for the center pixel is equal or greater to that of the neighboring pixels, the center pixel from the mask is represented with the binary code 0. Otherwise, it is given the binary code 1. Eq. (3) provides the method for calculating the hamming distance (HD) between the enrolled LBP code and the input code. n is total number of binary feature codes. BCA and BCI are the enrolled and input codes, respectively, and “ \oplus ” is an exclusive OR operation [7]. If BCA and BCI are identical, the HD is 0; otherwise, the HD is 1.

$$HD = \sum_{i=1}^n (BCA_i \oplus BCI_i) \quad (3)$$

In order to cover the misalignments of the input and enrolled codes caused by the movement of eye region due to the head movement, we perform the matching by bit-shifting. The matching by bit-shifting is performed only in the horizontal and vertical directions of the eye image within the limited ranges of pixels, respectively. The case of obtaining minimum HD while matching by bit-shifting is determined as the case of final matching.

3 Experimental Results

To evaluate the proposed method, we used the CASIA-Iris-Distance database (CASIA-IrisV4) which includes both eyes and faces in images 2352×1728 pixels in size [10]. The images in the CASIA-Iris-Distance database were acquired with the self-developed long-range multi-modal biometric image acquisition and recognition system (LMBS) [10]. Among the 2,567 face images (from 284 classes of 142 persons) in the CASIA-Iris-Distance database, we used 2,068 images (from 141 persons) by excluding images containing an abundance of noise. Thus, the total number of periocular images (from both the right and left eyes) was 4,136 (2068×2). In order to enhance the credibility of our experiment, an additional 8,272 images were produced by rotating the original images by $+5$ and -5 degrees-rotation levels that are common with people in front of a camera. The expanded database, therefore, contained 12,408 images ($4136 \text{ images} \times 3$ (± 5 and 0 degrees of rotation)). Because our research did not include a training procedure, all the images were used for testing. To measure the recognition accuracy, we used an equal error rate (EER). The EER is the error rate

averaging the false acceptance rate (FAR) and false rejection rate (FRR) when the difference between FAR and FRR is minimized. The FAR refers to the error rate from incorrectly accepting an un-enrolled person as an enrolled one. The FRR refers to the error rate from incorrectly rejecting an enrolled person as an un-enrolled one.

For the first experiment, we compared the accuracy in periocular recognition, as shown in Table 1. We compared the accuracy with and without bit-shift matching. Matching by bit-shifting is performed in both horizontal and vertical directions of the image of eye region, in the ranges $-1 \sim +1$, $-2 \sim +2$, and $-3 \sim +3$ pixels, respectively. We measured the minimum EER, as shown in Table 1. Table 1 shows the recognition accuracies. Without matching by bit-shifting, the EER was 13.1870%. With bit-shift matching, on the other hand, the minimum EER was 11.0697%, in case of the bit-shifting $-2 \sim +2$ pixels. From these results, we can confirm that the accuracy of the proposed method with bit-shifting is comparatively higher than that without matching by bit-shifting.

Table 1. Accuracy comparison of periocular recognition using LBP (Unit: %)

Without or with matching by bit-shifting	FAR	FRR	EER
None	13.4181	12.9558	13.1870
$-1 \sim +1$ pixels	11.7381	11.6543	11.6962
$-2 \sim +2$ pixels	11.1963	10.9430	11.0697
$-3 \sim +3$ pixels	12.2329	12.1431	12.1880

For the next experiment, we used the histogram features extracted with a uniform local binary pattern (uLBP) method and chi-square distance for calculating dissimilarity in order to check the applicability of our method with different methods of feature extraction and matching. This method of using the histogram features by uLBP was used by the previous research of periocular recognition, also [3]. Although they excluded the whole eye region for periocular recognition [3], we included this region for the fair comparison of our method. In another research [2], they also used these histogram features by uLBP extracted from the periocular region for the classification of gender and ethnicity.

Table 2. Accuracy comparison of periocular recognition using LBP and uLBP (Unit: %)

Recognition method	EER	Explanation
uLBP [2][3]	24.5713	1×1 sub-block, radius of 5, 16 pixels of uLBP mask
LBP (Proposed method)	11.0697	With matching by bit-shifting in the range of $-2 \sim +2$ pixels

Experimentation was performed with varying numbers of sub-blocks for the uLBP ($1 \times 1 \sim 5 \times 5$), radii of the uLBP mask ($1 \sim 5$), and the number of pixels for the uLBP mask (8, 12, 16). The cases with the lowest errors are shown in Table 2. As shown in Table 2, the EER by our method is lower than those by previous ones [2][3].

4 Conclusions

We proposed a novel periocular recognition method that is robust to the movement of eye region. To detect the periocular region, we used Adaboost, a sub-block-based template matching system, and CED. Then, the feature codes for periocular recognition are extracted using an LBP method. The recognition accuracy is enhanced with matching by bit-shifting in horizontal and vertical directions. Experimental results showed that the accuracy using our method was higher than other methods.

In future work, we plan to adapt our method for other applications, such as periocular-based age and gender recognition.

Acknowledgments. This study was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) Support Program (NIPA-2014-H0301-14-1021) supervised by the NIPA (National IT Industry Promotion Agency).

References

1. Daugman, J.: New Methods in Iris Recognition. *IEEE Trans. Syst. Man, Cybernetics - Part B* 37, 1167–1175 (2007)
2. Lyle, J.R., Miller, P.E., Pundlik, S.J., Woodard, D.L.: Soft Biometric Classification Using Periocular Region Features. In: *International Conference on Biometrics: Theory Applications and Systems (BTAS)*, Washington, DC, USA, September 27-29 (2010)
3. Woodard, D.L., Pundlik, S., Miller, P., Jillela, R., Ross, A.: On the Fusion of Periocular and Iris Biometrics in Non-ideal Imagery. In: *International Conference on Pattern Recognition*, Istanbul, Turkey, August 23-26, pp. 201–204 (2010)
4. Bharadwaj, S., Bhatt, H.S., Vatsa, M., Singh, R.: Periocular Biometrics: When Iris Recognition Fails. In: *International Conference on Biometrics: Theory Applications and Systems (BTAS)*, Washington, DC, USA, September 27-29 (2010)
5. Park, U., Jillela, R.R., Ross, A., Jain, A.K.: Periocular Biometrics in the Visible Spectrum. *IEEE Trans. Inf. Forensics, Secur.* 6, 96–106 (2011)
6. Viola, P., Jones, M.J.: Robust Real-time Face Detection. *Int. J. Comput. Vis.* 57, 137–154 (2004)
7. Shin, K.Y., Kim, Y. G., Park, K. R.: Enhanced Iris Recognition Method Based on Multi-Unit Iris Images. *Opt. Eng.* 52, 047201-1–047201-11 (2013)
8. Viola, P., Jones, M.: Rapid Object Detection Using a Boosted Cascade of Simple Features. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, December 8-14, pp. I-511–I-518 (2001)
9. Ahonen, T., Hadid, A., Pietikäinen, M.: Face Recognition with Local Binary Patterns. In: Pajdla, T., Matas, J(G.) (eds.) *ECCV 2004*. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
10. CASIA-Iris-Distance database (CASIA-IrisV4), <http://www.cbsr.ia.ac.cn/china/Iris%20Databases%20CH.asp>

Residual Histogram Shifting Technique Based on Cascading Prediction for Reversible Data Hiding

Yu-Chen Hu¹, PiYu Tsai², Jieh-Shan Yeh¹, and Wu-Lin Chen¹

¹Department of Computer Science and Information Management, Providence University,
TaiChung, Taiwan (Republic of China)

²Department of Computer Science and Information Engineering,
National United University, Miaoli, Taiwan (Republic of China)
{ychu, jsyeh, wlchen}@pu.edu.tw, pytsai@nuu.edu.tw

Abstract. To reversibly embed the secret data into the host image, the histogram shifting (HS) technique was proposed by embedding 1-bit secret data into the pixels in the peak points. To improve the hiding capacity of the histogram shifting technique, the residual histogram shifting (RHS) technique had been proposed. The prediction errors in the peak points of the residual histogram are used to embed the secret data. To further improve the hiding capacity, an improved block-based residual histogram shifting technique is proposed in this paper. The cascading prediction mechanism is employed in the proposed technique to generate the prediction errors for each image block. Experimental results reveal that the proposed technique provides higher hiding capacity than HS and RHS while keeping good image qualities of the embedded images.

1 Introduction

In the literature, the data hiding techniques can be classified into two categories: irreversible data hiding techniques [1]-[2] and reversible data hiding techniques [3]-[10]. In the reversible data hiding technique, the secret data can be extracted and the original host images can be recovered simultaneously. That is why the reversible data hiding techniques are often called the lossless data hiding techniques. Basically, two main approaches of the reversible data hiding techniques had been introduced. They are the difference expansion (DE) approach [6] and the histogram shifting (HS) approach [7]-[10]. Some reversible data hiding techniques that use the mix of these two approaches had also been designed.

In the literature, Ni et al. first introduced the concept of HS [7] for reversible data hiding in 2006. The occurrences of all possible pixel values in the cover image are calculated to generate the image histogram. Secret data are embedded into the pixels in the peak points. Basically, the hiding capacity of HS equals the number of pixels in the peak points used. To increase the hiding capacity, more pairs of peak and zero points can be used in HS. However, it is impossible to find out more pairs of peak and zero points in most nature images when the number of pairs is greater than 2.

To increase the hiding capacity of HS, several RHS techniques had been proposed. Different prediction ways were introduced in these RHS techniques to increase the occurrences of the peak points in the residual image histogram.

Among these RHS techniques, the block-based prediction histogram shifting (BPHS) technique [8] proposed by Tsai et al. achieves good results and resists to error propagation. In BPHS, the cover image is partitioned into non-overlapped image blocks. The center pixel in each block is selected as the basic pixel for predictive coding. For every block, the difference between each pixel and the basic pixel is calculated to generate the prediction error. No prediction error is generated for the basic pixel in each block. By sequentially processing each block, the prediction errors of the host image are generated. The residual image histogram is then constructed. The prediction errors in the peak points of the residual histogram are used to embed the secret data in BPHS. To further improve the hiding capacity of BPHS while keeping good image quality of the embedded image, an improved residual histogram shifting technique is proposed in this paper.

2 The Proposed Technique

In the proposed technique, the cover image is partitioned into non-overlapped image blocks of $n \times n$ pixels. Recall that the center pixel in each block is selected as the basic pixel to generate the prediction errors in BPHS. To take the image block of 4×4 pixels as shown in Figure 1(b) for example, the center pixel p_{10} in Figure 1(a) is taken as the reference pixel.

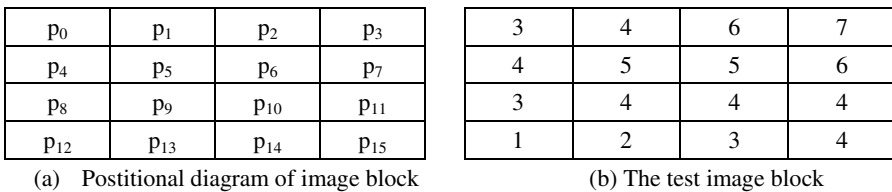


Fig. 1. Illustrative example for 4×4 image block

To clearly describe the proposed technique, the prediction rule of the 4×4 block by using BPHS is listed in Figure 2(a). A total of 15 prediction errors are generated by using BPHS. Let us take the test image block as shown in Figure 1(b) for example, and the prediction errors of this block by using BPHS are shown in Figure 2(b).

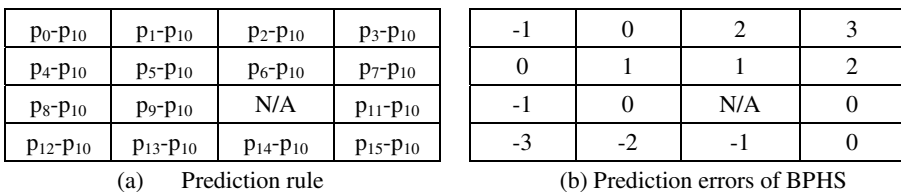


Fig. 2. Example of BPHS

It is obvious that the center pixel will have a lower degree of similarity to the pixels which are far from it when the size of the block increases. To increase the occurrences of the peak points, the cascading prediction mechanism is employed in the proposed technique. An illustrative example to describe how it works on the 4×4 block is depicted in Figure 3. Similarly, the center pixel p_{10} is taken as the reference pixel of the image block. First, eight prediction errors as listed in Figure 3(a) of these directly adjacent pixels to the center pixel are computed. Then, the remaining seven prediction errors are computed as shown in Figure 3(b). Here, p_5 is taken as the reference pixel for p_0 , p_1 , and p_4 . Besides, p_6 , p_7 , p_9 , and p_{13} are taken as the reference pixels for p_2 , p_3 , p_8 , and p_{12} , respectively.

-	-	-	-
-	p_5-p_{10}	p_6-p_{10}	p_7-p_{10}
-	p_9-p_{10}	N/A	$p_{11}-p_{10}$
-	$p_{13}-p_{10}$	$p_{14}-p_{10}$	$p_{15}-p_{10}$

(a) Prediction errors of directly adjacent pixels

p_0-p_5	p_1-p_5	p_2-p_6	p_3-p_7
p_4-p_5	-	-	-
p_8-p_9	-	N/A	-
$p_{12}-p_{13}$	-	-	-

(b) Prediction errors of the others

Fig. 3. Proposed cascading prediction rules

To take the test image block as shown in Figure 1(b) for example, the prediction errors of this block by using the proposed cascading prediction rule are shown in Figure 4(b). A total of 15 prediction errors are generated by the proposed cascading prediction rule for the image block of 4×4 pixels. The resultant image histograms of BPHS and the proposed technique are listed in Figure 5(a) and Figure 5(b), respectively.

-2	-1	1	1
-1	1	1	2
-1	0	N/A	0
-1	-2	-1	0

Fig. 4. Example of the proposed cascading prediction mechanism

When the prediction errors of each $n \times n$ image block are computed by using the cascading prediction rule, the occurrences of all possible residual values are calculated to generate the residual image histogram. Secret data are then embedded into the residual values in the peak points. Multiple pairs of peak and zero points can be used to increase the hiding capacity. After the secret data are embedded into the residual values, the embedded image can be generated by performing the reverse cascading prediction.

In the proposed technique, the cascading prediction rule can be easily extended to the image block of $n \times n$ pixels where n is greater than 4. In Figure 6, the cascading prediction rule for the 5×5 image block are listed. A total of 24 prediction errors are generated by using the cascading prediction rule for the image block of 5×5 pixels.

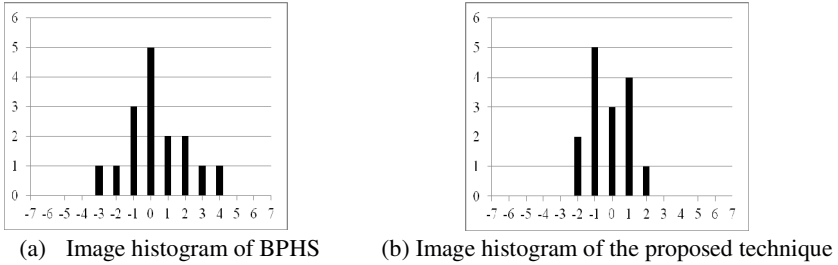


Fig. 5. Resultant image histograms

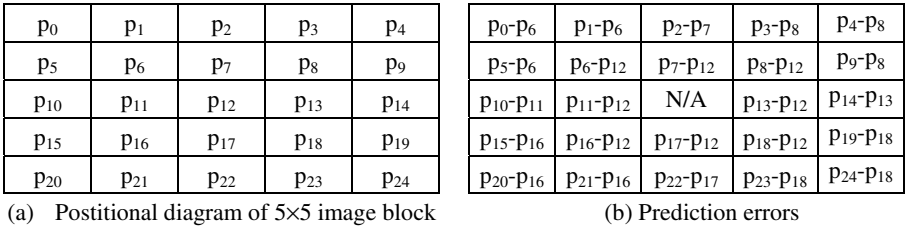


Fig. 6. Proposed cascading prediction rules for 5×5 image blocks

3 Experimental Results

Our experiments are performed on windows 7 PC with an Intel Core i5 2.8 GHz CPU and the 2 GB RAM. The testing programs are implemented by using Bloodshed Dev C++. In our experiments, six grayscale images of 512×512 pixels, “Airplane”, “Boat”, “Girl”, “Goldhill”, “Lenna”, and “Pepper” as shown in Figure 7, are used. In the simulations, the peak signal-to-noise-ratio (PSNR) measurement is employed.



Fig. 7. Grayscale test images

To verify the efficiency of the proposed technique, comparative results of the hiding capacity are listed in Table 1. Four grayscale images of 512×512 pixels are used in the simulations. The block size is set to 4×4. Here, p denotes the pairs of the peak and zero points used in each technique. According to the results, the hiding capacity of each technique increases with the increase of the p value. The proposed techniques provides the highest hiding capacity among these comparative techniques. Average hiding capacities of 9479.5, 52100.75, and 59155.75 bits are achieved when p is set to 2 by using HS, BPHS, and the proposed technique, respectively. Compared to BPHS, average hiding capacity gain of 7055 bits is achieved by using the proposed technique.

Table 1. Results of the hiding capacity (unit: bits) of HS, BPHS, and the proposed technique

Images	HS		BPHS		Proposed Technique	
	$p = 1$	$p = 2$	$p = 1$	$p = 2$	$p = 1$	$p = 2$
Airplane	9002	17042	36532	63616	41328	71406
Lenna	2908	5760	24447	46009	28785	53776
Tiffany	5120	10153	28299	54169	32074	57419
Zelda	2565	4963	22646	44609	27690	54022

Comparative results of the embedded image quality are listed in Table 2. The peak-to-noise ratio (PNSR) measurement is used in the simulation. It is shown that the embedded image quality of each technique decreases as the p value increases. Average image qualities of 48.211 dB, 49.437 dB, and 47.819 dB are achieved when p is set to 2 in HS, BPHS, and the proposed technique, respectively. Compared to BPHS, average image quality loss of 1.618 dB is incurred by using the proposed technique.

Table 2. Results of the embedded image quality (unit: dB) of HS, BPHS, and the proposed technique

Images	HS		BPHS		Proposed Technique	
	$p = 1$	$p = 2$	$p = 1$	$p = 2$	$p = 1$	$p = 2$
Airplane	53.237	48.272	52.150	49.713	49.540	48.063
Lenna	53.849	48.179	51.924	49.313	49.238	47.551
Tiffany	52.172	48.219	51.990	49.438	49.466	48.199
Zelda	53.317	48.172	51.880	49.283	49.148	47.462



(a) HS



(b) BPHS



(c) Proposed technique

Fig. 8. Results of the embedded images of the test image 'Lenna'

Figure 8 shows the embedded images of HS, BPHS, and the proposed technique. To generate the embedded images of HS, BPHS, and the proposed technique, the p value is set to 2. According to the results, the visual qualities of the embedded images are achieved. It is hard to distinguish the difference between each embedded image and the original one by using the human visual system.

4 Conclusions

An improved residual histogram shifting technique based on BPHS are proposed in this study. This technique provides greater hiding capacity than HS and BPHS while keeping good image quality of the embedded image. Compared to BPHS, about 13% of hiding capacity improvement is achieved by using the proposed technique. The embedded image quality of the proposed technique is greater than 47 dB. In addition, good visual quality of the embedded image is achieved according to the experimental results.

Acknowledgement. This research was partially supported by the Ministry of Science and Technology, Taipei, R.O.C. under contracts MOST 103-2410-H-126-009-MY3 and MOST 103-2632-E-126-001-MY3.

References

1. Bender, W., Gruhl, D., Morimoto, N., Lu, A.: Techniques for Data Hiding. *IBM Systems Journal* 35(3-4), 313–336 (1996)
2. Hu, Y.C.: High Capacity Image Hiding Scheme Based on Vector Quantization. *Pattern Recognition* 39(9), 1715–1724 (2006)
3. Cheddad, A., Condell, J., Curran, K., Kevitt, P.M.: Digital Image Steganography: Survey and Analysis of Current Methods. *Signal Processing* 90(3), 727–752 (2010)
4. Chang, C.C., Yu, Y.H., Hu, Y.C.: Hiding Secret Data in Images via Predictive Coding. *Pattern Recognition* 38(5), 691–705 (2005)
5. Chang, C.C., Wu, W.C., Hu, Y.C.: Lossless Recovery of a VQ Index Table with Embedded Secret data. *Journal of Visual Communication and Image Representation* 18(3), 207–216 (2007)
6. Tian, J.: Reversible Data Embedding Using a Difference Expansion. *IEEE Transactions on Circuits and Systems for Video Technology* 13(8), 890–896 (2003)
7. Ni, Z., Shi, Y.Q., Ansari, N., Su, W.: Reversible Data Hiding. *IEEE Transactions on Circuits and Systems for Video Technology* 16(3), 354–362 (2006)
8. Tsai, P.Y., Hu, Y.C., Yeh, H.L.: Reversible Image Hiding Scheme Using Predictive Coding and Histogram Shifting. *Signal Processing* 89(6), 1129–1143 (2009)
9. Huang, L.C., Tseng, L.Y., Hwang, M.S.: A Reversible Data Hiding Method by Histogram Shifting in High Quality Medical Images. *The Journal of Systems and Software* 86(3), 716–727 (2013)
10. Chang, I.C., Hu, Y.C., Chen, W.L., Lo, C.C.: High Capacity Reversible Data Hiding Scheme Based on Residual Histogram Shifting for Block Truncation Coding. *Signal Processing* 108, 376–388 (2015)

Color and Depth Image Correspondence for Kinect v2

Changhee Kim¹, Seokmin Yun¹, Seung-Won Jung², and Chee Sun Won^{1*}

¹ Dongguk University-Seoul,

Dept of Electronics and Electrical Engineering,

² Dongguk University-Seoul, Dept of Multimedia Engineering,

Dongguk University-Seoul, 30, Pildong-ro 1gil, Jung-gu, Seoul 100-715, Korea
Kimchda210@naver.com, {smyun, swjung83, cswon}@dongguk.edu

Abstract. Kinect v2, a new version of Kinect sensor, provides RGB, IR (Infra-Red) and depth images like its predecessor Kinect v1. However, the depth measurement mechanism and the image resolutions of the Kinect v2 are different from those of Kinect v1, which requires a new transformation matrix for the camera calibration of Kinect v2. In this paper, we correct the radial distortion of the RGB camera and find the transformation matrix for the correspondence between the RGB and depth image of the Kinect v2. Experimental results show that our method yields accurate correspondence between the RGB and depth images.

Keywords: Kinect v2, registration, camera calibration.

1 Introduction

Kinect from Microsoft is a very popular depth sensor. Thanks to its low price, it has been widely adopted for various computer vision applications, including 3D reconstruction, object recognition, and object tracking. To adopt the Kinect for more sophisticated computer vision applications, however, RGB and depth sensors of Kinect need to be precisely calibrated. The calibration for the first version of the Kinect (Kinect v1) has been proposed [1].

Recently, the second version of Kinect (Kinect v2) has been released. The Kinect v2 adopts a different sensing method for the depth measurement and provides higher image resolutions. Specifically, the Kinect v2 uses ‘Time of Flight’ (TOF) method instead of ‘light coding’ of the Kinect v1 for the depth measurements. Also, the image resolutions for both RGB and depth of the Kinect 2 are higher than those of Kinect v1. See Table 1 for more comparisons between Kinect v1 and v2 [2]. Because of the differences listed in Table 1, the calibration parameters developed for the Kinect v1 are not applicable for the Kinect v2 sensors. In this paper, we correct the image distortions caused by the lens of the Kinect v2 and provide its calibration matrix for the correspondence between the RGB, depth, and IR (Infra Red) images.

This paper is composed of the following sections. In Section 2, the radial distortions of the Kinect are corrected. The holes in the depth image are filled and the correspondence matrix between the RGB and the depth images is provided. In Section 3 we show the accuracy of our method by comparing with the results by SDK tools.

* Corresponding author.

Table 1. Comparative specifications of Kinect v1 and Kinect v2

	Kinect v1	Kinect v2
Resolution of color image	640x480 (pixel)	1920x1080 (pixel)
Resolution of IR and depth image	320x240 (pixel)	512x424 (pixel)
Field of view of color image	62°x48.6°	84.1°x53.8°
Field of view of IR and depth image	57.5°x43.5°	70.6°x60°
Maximum skeletal tracking	2	6
Method of depth measurement	Light coding	Time of Flight
Working range	0.8m~3.5m	0.5m~8m

2 Kinect v2 Calibration

2.1 Distortion Correction

Camera lens causes image distortions, where Kinect v2 is not an exception. Both IR and color cameras have distortions. Calibration tools such as Camera Calibration Toolbox for Matlab are available to determine the intrinsic parameters of the cameras. A set of checkerboard images from both color and IR cameras are used to identify the corners of the checkerboard pattern in RGB and IR images. Then, by solving the equations from the correspondences of the corner points and by using the non-linear optimization technique to reduce the reprojection errors, the intrinsic camera parameters such as focal length, principal points, and skew can be determined [3]. Then, the radial distortions of the color camera can be corrected. The results of our calibration parameters for the Kinect v2 are listed in Table 2.

Table 2. Camera parameters for Kinect v2

Camera Parameter	RGB	IR
Focal Length (f_c)	[1053.622 1047.508] ± [4.6884 4.5323]	[376.6518 371.4936] ± [1.8265 1.8015]
Principal points (c_c)	[950.3941 527.3442]	[265.5583 206.6131]
Skew (α_c)	[0.000] ± [0.000] →angle of pixel=90.00 degrees	[0.000 ± [0.000] →angle of pixel=90.00 degrees
Distortion (k_c)	[0.0042 -0.0019 -0.0038 -0.0026 0] ± [0.0038 0.0033 0.0007 0.0008 0]	[-0.0094 -0.0431 0.0004 -0.0003 0] ± [0.0094 0.0144 0.0015 0.0017 0]

Zoomed images before and after the correction of the Kinect v2 distortions are shown in Fig. 1. As one can see in Fig. 1 (a) and (c), the images captured by the Kinect v2 sensor suffer from the radial distortions. These distortions are corrected by the calibration parameters in Table 2 (see Fig. 1 (b) and (d)).

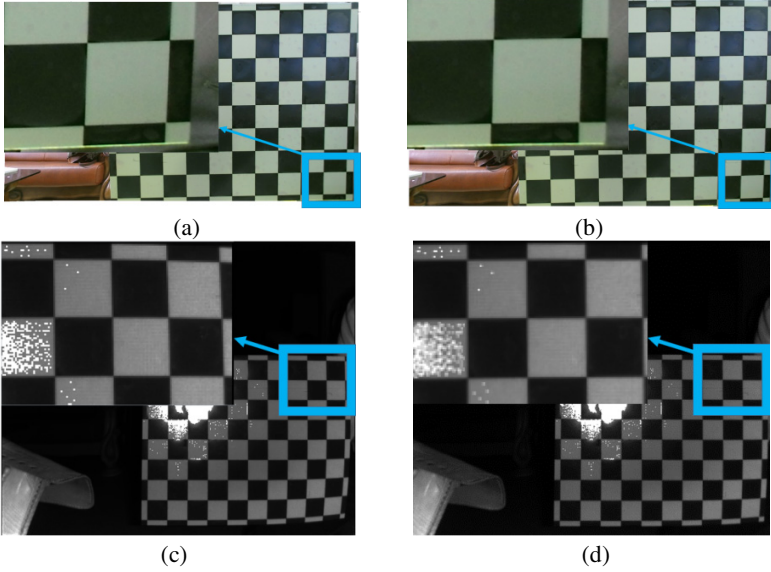


Fig. 1. Images with distortions (a) and (c) and after distortion correction (b) and (d). Above images are taken by color camera and below ones are IR images.

2.2 Correspondence of Color and IR (Depth) Images

The correspondence between the RGB sensor and the IR (depth) sensor can be done by a transformation matrix between them. Since the depth image is generated by the IR sensor, we can use either the depth image or the IR image for the registration. Here, since the depth image cannot show the pattern on the planer checkerboard, our registration is based on the checkerboard images of the RGB and the IR images.

The projective matrix which converts the RGB image coordinate (x,y) into the IR image coordinate (X,Y) is given as follows [4].

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

Matrix equation (1) with eight unknown parameter values can be rewritten as follows

$$X = a_1x + a_2y + a_3 - a_7xX - a_8yX \quad (2)$$

$$Y = a_4x + a_5y + a_6 - a_7xY - a_8yY \quad (3)$$

Four corresponding pairs of (x,y) and (X,Y) must be known to solve equations (2) and (3). If we use more than four pairs, the parameters can be found using a least square method like the direct linear transform. We use the checkerboard to find the pairs of corresponding points between the RGB and IR images as shown in Fig. 2. The corresponding points are selected manually. Using these points, we calculate the eight parameters for the transformation matrix. Specifically, we use 515 pairs of points extracted from indoor images of limited distances to calculate equations (2) and (3), and we get the transformation matrix as shown in (4). So, the matrix in (4) is good for the calibration of Kinect v2 images captured in the near distances, say up to 3m.

$$T = \begin{bmatrix} 0.3584 & 0.0089 & 0.0000 \\ 0.0031 & 0.3531 & 0.0001 \\ -101.5934 & 13.6311 & 0.9914 \end{bmatrix} \quad (4)$$

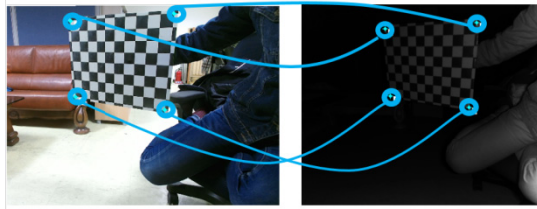


Fig. 2. Corresponding points in RGB and IR images



Fig. 3. Above: original color and IR images. Below: registered color and IR images obtained by the transformation and cropping.

Note that the size of RGB image is bigger than the IR image in Kinect v2. Also, they have different field of views (FOV). Therefore, color images need to be cropped after registration. Fig. 3 shows an example of registration results. Above images are raw images and below ones are the results after the registration and cropping. Color image is cropped at both sides to fit the FOV of the IR image. The IR image is also cropped at top-bottom parts because the top-bottom FOV of color image is bigger than IR image. After the calibration and cropping the size of IR image and color image is changed to 512x360. Original size of color is 1920x1080 and that of IR is 512x424.

2.3 Hole Filling

Since the depth image is captured by the IR sensor in Kinect v2, the registration between the RGB and the IR image automatically yields the registration among the RGB, IR, and depth images. So, after the registration and cropping, we can generate the pixel-by-pixel correspondence images of RGB, IR, and depth from the Kinect v2 sensor. Here, to make a perfect correspondence, the holes in the depth image are to be filled.

As in Kinect v1, Kinect v2 has depth holes with missing depth measurements. Although the holes of the Kinect v2 along the object boundary are usually thinner

than those of the Kinect v1, the holes near the object boundary in the Kinect v2 depth image can be still filled by the method of Kinect v1. In particular, since the RGB image is already aligned with the depth image, we can exploit the edge information in the RGB image to determine the direction of the hole filling in the depth image [5].

Table 3. Comparison of pixel distance

Image set \ Method	SDK	Ours
A	32.5038	5.7118
B	21.3204	3.0104
C	24.8821	7.8142
Average	26.3254	5.5121

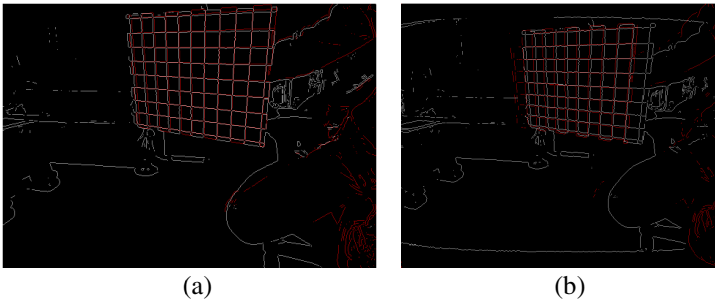


Fig. 4. Superimposed images (red lines are edges of IR image and white ones are edges of color image): (a) our method, (b) the SDK function

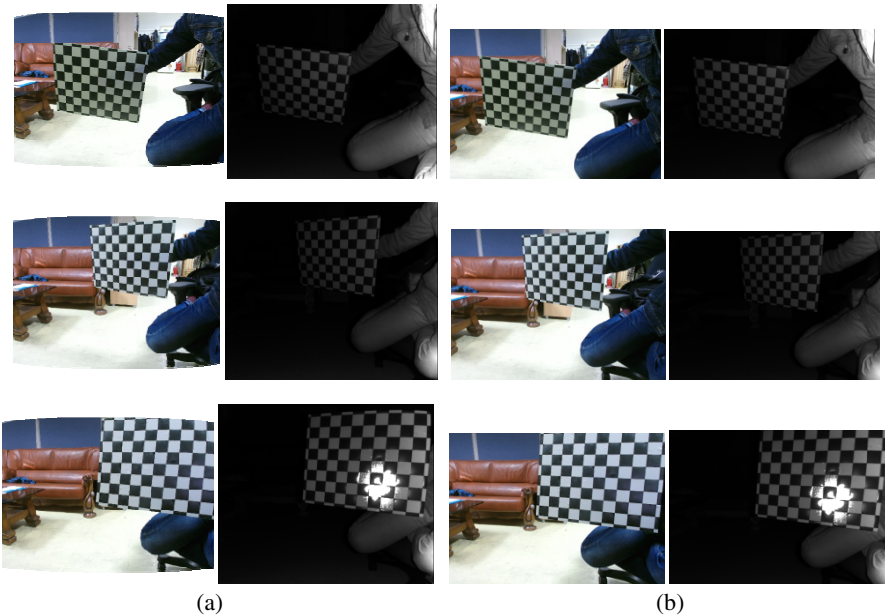


Fig. 5. Three sets of color and IR images that are used in pixel position error calculation: (a) Result images from the SDK and (b) Result images from our method

3 Results

Fig. 4-(a) shows the superimposed images after the correction of the radial distortion and the registration. Compared to the result of Fig. 4-(b) obtained by the software development kit (SDK) [2] function of “*MapColorFrameToDepthSpace*”, the mismatch errors of our method are much smaller than those of the SDK method. Specifically, the checkerboard lines in the left image (Fig. 4-(a)) are well aligned compared to the checkerboard lines in right image (Fig. 4-(b)).

We also calculate the pixel distances of correspondence points of checkerboard in the registered color and IR images to compare the pixel position errors. The results using three sets of color and IR images are shown in Table 3. Also there are images used for calculating pixel position errors in Fig. 5. The average distance of the SDK method is 26.2354 and that of our result is 5.5121. We can notice that our method is about five times more accurate than the SDK method in terms of the pixel mismatches.

4 Conclusion

In this paper, we performed a case study of aligning color, IR, and depth images using Kinect v2 sensor. We obtained the intrinsic parameters for Kinect v2 and calibrated the RGB and IR sensors via the transformation matrix between the two sensors. High accuracy of the proposed registration was confirmed by comparing to the results of the SDK function visually and numerically. As a future work we need a universal calibration matrix for all near and far distances covered by the Kinect v2.

Acknowledgments. This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2013R1A1A2005024) and by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2015-H0301-14-4007) supervised by the NIPA(National IT Industry Promotion Agency).

References

1. Smisek, J., Jancosek, M., Pajdla, T.: 3D with Kinect. Consumer Depth Cameras for Computer Vision, pp. 3–25. Springer, London (2013)
2. Microsoft, Kinect for Windows, <http://www.microsoft.com/en-us/kinectforwindows>
3. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1330–1334 (2000)
4. Rothwell, C., Forsyth, D.A., Zisserman, A., Mundy, J.L.: Extracting Projective Structure from Single Perspective Views of 3D Point Sets. In: Fourth International Conference on Computer Vision, pp. 573–582. IEEE Press (1993)
5. Le, A.V., Jung, S.W., Won, C.S.: Directional Joint Bilateral Filter for Depth Images. *Sensors* 14, 11362–11378 (2014)

Reduced Reference Quality Metric for Depth Images

Thanh-Ha Le¹, Seongjo Lee², Seung-Won Jung², and Chee Sun Won³

¹ University and Engineering and Technology, Vietnam National University, Hanoi

² Department of Multimedia Engineering, Dongguk University, Seoul, South Korea

³ Department of Electronics and Electrical Engineering, Dongguk University,
Seoul, South Korea

Abstract. In this paper, a new quality metric for depth images is proposed. Unlike the conventional depth metrics which require the additional information such as the ground truth depth image or a stereo image pair, the proposed quality metric demands only a single camera image and its corresponding depth image. In this work, we first empirically observe that the depth distortion is closely related to the local image characteristics. Based on the observation, we introduce a method to assess the local depth distortion for the edge and non-edge regions. Then, the local distortion is adaptively weighted by the Gabor filter and added up to the quality metric for the depth image.

Keywords: Depth image, Gabor filter, image quality assessment, reduced reference, quality metric.

1 Introduction

Depth image, often called the depth map, plays a fundamental role in 3-D video and free viewpoint video applications [1]. In the stereo vision, a right-view image can be reconstructed by using a left-view image and its corresponding depth image. Besides, in the multi-view video, an arbitrary viewpoint image can be generated by interpolating or extrapolating the given images using the depth images. Due to the importance and applicability of the depth image, the estimation of the depth image has been extensively researched in the last decades [2].

In the classical computer vision field, the depth image estimation is formulated as a stereo matching problem. Although the state-of-the-art stereo matching techniques [3, 4] could significantly improve the estimation accuracy of the depth image, inaccurate depth images are often produced because of occlusion and large homogeneous regions. In order to alleviate the inherent difficulty of stereo matching, the depth camera is used to capture the depth of the scene. Owing to the advances in the hardware of the depth camera, an accurate depth image can be achieved but the noise is inevitable and the depth range is limited.

Regardless of the depth image acquisition methods, the quality assessment of the depth image is required. Specifically, to evaluate the performance of the stereo matching algorithms or the depth cameras, the objective depth quality metric (DQM) should be defined. One simple solution is to compare the estimated depth image with the ground truth depth image [2]. This full reference DQM (FR-DQM) can ideally measure the accuracy of the depth image; however, the ground truth depth image is not available in practice. Another possible DQM is to gauge the quality of the

reconstructed image obtained by the depth image. For instance, the right-view image is approximated by warping the left-view image using the depth image, and the warped image is compared with the original right-view image. However, such image pairs are not always obtainable in the depth-image-based rendering (DIBR) [5] and the depth camera applications. Although the above two DQMs have a certain limitation, a more general DQM has not been widely studied.

Recently, an FR quality metric for the geometrically distorted images showed that the severe geometric degradation of the image structure is originated from the distortion of the displacement field into the orthogonal direction of the image bars and edges [6]. Motivated by this work, a new reduced reference DQM (RR-DQM) is proposed in this paper. The proposed RR-DQM requires only one camera image and its corresponding depth image. Since the objective is to measure the accuracy of the depth image, the camera image is considered as the side or reduced information of the depth. Throughout the experimental studies, the major sources of depth distortion are found and a suitable measurement is developed using the Gabor filter and the smallest univalue segment assimilating nucleus (SUSAN) detector [7].

The rest of this paper is organized as follows. The proposed RR-DQM is presented in Section 2. The conclusions are given in Section 3.

2 Proposed Depth Quality Metric

In general, the depth image is used to render or synthesis images for multi-view and 3-D video applications. Thus, the same local distortion of the depth image does not

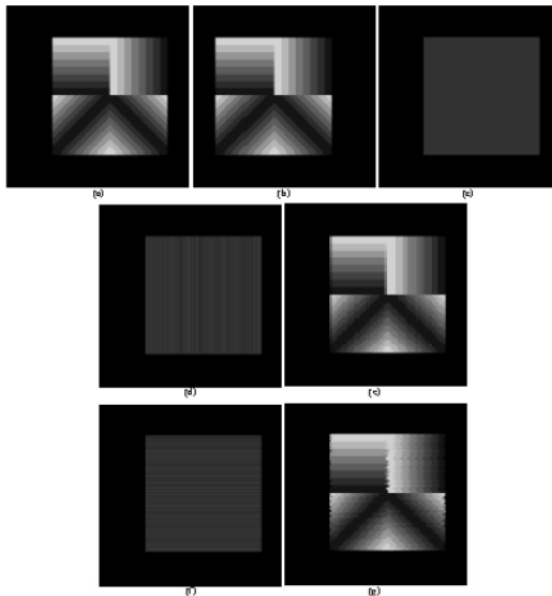


Fig. 1. Synthetic example: (a) left-view image, (b) right-view image, (c) ground truth depth image, (d) depth image with distortion along horizontal direction, (e) compensated left-view image obtained using (b) and (d), (f) depth image with distortion along vertical direction, (g) compensated left-view image obtained using (b) and (f).

equally affect the resultant images. In other words, the local distortion of the depth image should be jointly considered with the local image characteristics.

To demonstrate this supposition, a pair of simple synthetic stereo images of the size 400x400 is generated as shown in Figs. 1(a) and (b). Here, the square of the size 256 x256 containing four representative directional edges as shown in Fig. 1(a) is shifted by 50 pixel distances to the left direction as shown in Fig. 1(b). In other words, the pixels inside the square have the same horizontal disparity as shown in Fig. 1(c). The background is assumed to be located in the infinite distance, and thus there is no horizontal disparity. Since the other directional disparities can be eliminated by using the image rectification technique, the assumption of the horizontal disparity case only is simple but widely adopted [2].

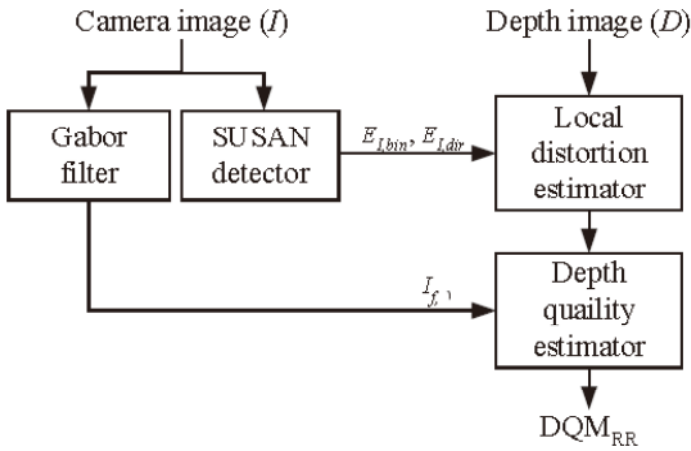


Fig. 2. Flowchart of the proposed method

The disparity values inside the square are then distorted in the horizontal and vertical directions, as shown in Figs. 1(d) and (f), respectively. Specifically, for the depth image with distortion along horizontal direction, one row of the zero-mean uniformly distributed random noise with a variance of 10 is constructed. Its length is the same as the width of the square, 256, and this row is added to the all rows in the square region of the depth image. Thus, the depth values are the same in the vertical direction and only depth distortion exists along the horizontal direction as shown in Fig. 1(d). The depth image with distortion along vertical direction is constructed in a similar manner.

Given the depth image and the right-view image, the left-view image can be reconstructed. Specifically, the pixels in the left-view image are found from the pixels in the right-view image in which the pixel positions are determined according to the horizontal disparity values in the depth image. From the compensated left-view images as shown in Figs. 1(e) and (g), we can first see that the horizontal image edges are not visually deteriorated. Since only the horizontal disparity is assumed, the different types of distortion can only change the start and end positions of the horizontal edges. Thus, the local distortion of the depth image in the horizontal edge

regions can be disregarded. It is also confirmed that the distortion in the compensated images is prominent when the depth value varies along the image edges. For instance, the vertical image edges are severely damaged when the depth image has distortion along vertical direction as shown in Figs. 1(f) and (g).

From the above observations, it is found that the effect of the local depth distortion is strongly dependent on the local image characteristics. Thus, the relation between the depth distortion and image characteristics should be exploited to measure the accuracy of the depth image. Figure 2 shows the flowchart of the proposed RR-DQM. In the proposed method, Gabor filter is used to differently weight the local image structures. In addition, the SUSAN edge detector [7] is employed to attain the edge information of the image. In particular, the SUSAN detector is known to robustly estimate image edges and their edge direction.

Let $E_{I,bin}$ and $E_{I,dir}$ denote the binary edge map and the edge direction map of the image I obtained by the SUSAN detector, respectively. For the simplicity, $E_{I,dir}$ is quantized to represent only the horizontal, vertical, left diagonal, and right diagonal directions. Then, for the non-edge positions, the local depth distortion is measured by the average difference of the depth values in the local neighborhood. On the other hand, for the edge positions, the depth variation along the edge direction is assessed to take the edge distortion or deformation into account.

Let $D_{I,D}$ represent the depth distortion map obtained from the binary edge map $E_{I,bin}$ and the depth image D , defined as

$$D_{I,D}(x,y) = \begin{cases} \frac{1}{n(N)} \sum_{(u,v) \in N} |D(x,y) - D(x+u,y+v)| & ; \text{if } E_{I,bin}(x,y) = 0 \\ |D(x,y) - \frac{1}{2}(D(x+u_1,y+v_1) + D(x+u_2,y+v_2))| & ; \text{otherwise} \end{cases}, \quad (1)$$

where N is the set of 8-neighborhood positions and $n(N)$ is the cardinality of N . Thus, for the non-edge positions, the mean absolute difference (MAD) of the depth values is used to describe the local depth distortion. For the edge positions, the average of the two neighboring depth values along the edge direction is compared with the depth value of the current position. Here, (u_i, v_i) is set according to the edge direction. For instance, $(u_1, v_1) = (0, 1)$ and $(u_2, v_2) = (0, -1)$ for the vertical edge. Since the gradual change along edge direction tends to be the actual depth variation, the central difference is adopted to assess the local depth distortion.

Until now, the methods of obtaining the Gabor energy and the depth distortion map have been described. The rationale behind the proposed quality metric of depth images is that depth discontinuities along the image edges are the main factor of the depth distortion. Thus, both the local image characteristics (encoded in the Gabor energy) and the local depth distortion (encoded in the depth distortion map) are used to define the final distortion map, $\overline{D}_{I,D}$. To this end, $\overline{D}_{I,D}$ is defined by combining the depth distortion map and the Gabor energy as follows:

$$\overline{D}_{I,D}(x,y) = \sum_{\theta \in \Theta} w_{\theta} \cdot I_{f,\theta}(x,y) \cdot D_{I,D}(x,y), \quad (2)$$

where $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ and w_θ is the weight of the direction θ . Similar definition for the image geometric distortion can be found in [6]. Figure 3 shows the final distortion maps obtained by using Figs. 1(d) and (f), where $w_\theta = \{1, 0.5, 0, 0.5\}$ for four directions in the same order of θ . Since the vertical edge is the most sensitive and the horizontal one is less vulnerable to the local depth distortion as we confirmed, w_θ is determined accordingly. By comparing Figs. 1 and 3, it can be seen that the final distortion maps correspond with the visual geometric deterioration resulted from the depth distortion.

The remaining problem is to determine the metric, RR-DQM, of the depth image distortion. The RR-DQM is defined by pooling the all distortion values in $\overline{D}_{I,D}$ except for the outlier regions,

$$DQM_{RR} = \frac{1}{n(\Psi_1)} \left(\sum_{(x,y) \in \Psi_1} |\overline{D}_{I,D}(x,y)| \right), \quad (3)$$

where Ψ_1 is a set of all pixel positions in the image excluding outlier regions and $n(\Psi_1)$ is the cardinality of Ψ_1 .

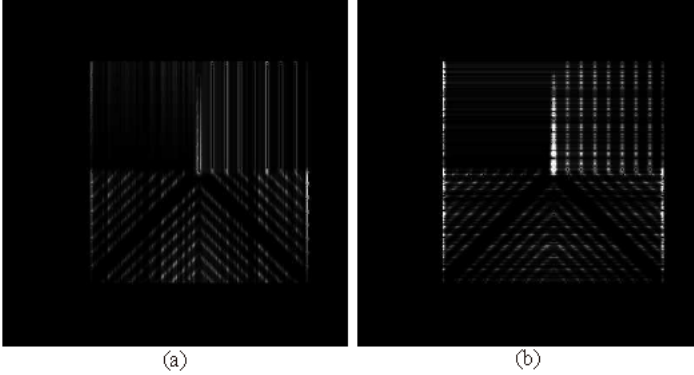


Fig. 3. Final distortion maps corresponding to (a) Fig. 1(d) and (b) Fig. 1(f)

3 Conclusion

We proposed a depth quality assessment technique that does not require the ground truth depth image or the stereo image pair. Based on the analysis using the synthetic image, the strong correlation between the local depth distortion and the local image characteristic is verified. Then, the depth distortion is measured depending on the edge directions. In addition, the Gabor filter is used to adaptively weight the local depth distortion. The experimental results show that the proposed metric closely approximates the conventional depth quality metrics that necessitate the additional information.

Acknowledgement. This work was supported by the basic research projects in natural science in 2012 of the National Foundation for Science & Technology Development (Nafosted), Vietnam (102.01-2012.36, Coding and communication of multiview video plus depth for 3D Television Systems). This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2015-H0301-15-1021) supervised by the NIPA (National IT Industry Promotion Agency).

References

1. Smolic, A., Mueller, K., Merkle, P., Kauff, P., Wiegand, T.: An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution. In: Proc. 27th Conference on Picture Coding Symposium, pp. 389–392 (2009)
2. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision* 47(1-3), 7–42 (2002)
3. Klaus, A., Sormann, M., Karner, K.: Segment-based stereo matching using brief propagation and a self-adapting dissimilarity measure. In: Proc. IEEE Conference on Pattern Recognition, pp. 15–18 (2006)
4. Wang, Z.-F., Zheng, Z.-G.: A region based stereo matching algorithm using cooperative optimization. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
5. Fehn, C.: Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. In: Proc. SPIE, vol. 5291, pp. 93–104 (2004)
6. D’Angelo, A., Zhaoping, L., Barni, M.: A full-reference quality metric for geometrically distorted images. *IEEE Trans. Image Process.* 19(4), 867–881 (2010)
7. Smith, S.M., Brady, J.M.: SUSAN—a new approach to low level image processing. *Int. J. Comput. Vision* 23(1), 45–78 (1997)
8. Scharstein, D., Szeliski, R., <http://vision.middlebury.edu/stereo>
9. Yang, Q., Yang, R., Davis, J., Nister, D.: Spatial-depth super resolution for range images. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)

Sustainable Operation Algorithm for High Availability with Integrated Desktop Storage Based on Virtualization

Hyun-Woo Kim¹, HwiRim Byun¹, Eun-Ha Song², and Young-Sik Jeong^{1,*}

¹ Department of Multimedia Engineering, Dongguk University, Seoul, Korea
{hwkim, hazzzly, ysjeong}@dongguk.edu

² Humanitas Collage, Wonkwang University, Iksan, Korea
ehsong@wonkwang.ac.kr

Abstract. Recently, following the rapid development of IT, diverse virtualization-based studies have been conducted on big data storage. Among the diverse studies, desktop storage virtualization integrates unused storage in distributed legacy desktops using virtualization and provides integrated storage to users for other purposes. In the case of this desktop storage virtualization, high availability is regarded as very important to providing reliability to storage users. In addition, although studies on hierarchical structures and resource incorporation in desktop-based integration environments have been conducted, studies on efficient operations following the occurrence of server faults are insufficient. To achieve this operational high availability, in the present paper, a Sustainable Operation Algorithm (SOA) that can actively respond to the occurrence of desktop server faults in desktop-based storage integration environments is proposed. If the SOA is applied in desktop-based integration environments, desktops can be easily added or removed. The alternative server actively operates following the occurrence of faults. Finally, the SOA provides high QoS to storage users.

1 Introduction

Recently, thanks to the advanced computing technologies of IT, manual job processing, into which great amounts of manpower have been inputted, was automated. In addition, thanks to the enhanced performance and convenient portability of diverse smart devices, leisure time has increased and work efficiency has improved. As data are rapidly created in diverse smart devices, many virtualization-based studies are conducted to store these big data. Desktop storage virtualization integrates unused storage into distributed legacy desktops using visualization. Integrated storage is provided when other users request storage [1, 2, 3, 4, 5, 6, 7].

As desktop storage virtualization is connected to many distributed desktops, reliability for the storage user is regarded as being very important. Therefore, functions to respond to the occurrence of exceptions, such as flaws and faults in the

* Corresponding author.

desktop operating as a server or the server being down due to an incorrect operation by the desktop user, are necessary. Among existing studies on desktop storage virtualization, studies on hierarchical structures and resource integration in desktop-based integration environments have been conducted. However, studies are insufficient in terms of such an efficient operation [3, 4, 5, 8, 9, 10].

For this reason, in the present paper, a Sustainable Operation Algorithm (SOA) that can actively respond to the occurrence of desktop server faults in desktop-based storage integration environments is proposed.

2 Related Works

Among existing studies on desktop-based resource integration techniques, DRV-CS [3], RISBD [4], and CSTORE [5] were analyzed, as shown in Table 1.

Table 1. Resource Integration Mechanism based on Desktop

Mechanism	Description
DRV-CS[3] (desktop resource virtualization-clustering simulator)	A DRV-CS, which is a clustering technique to integrate efficiently distributed desktop resources, was proposed. For efficient hierarchical structures, the DRV-CS brings the performances of distributed desktops to the surface and clusters them. Although optimum desktops can be selected according to required performance, the DRV-CS has a weakness in that clustering is stopped following the occurrence of a server fault. In the present paper, a scheme of operation even when a server fault has occurred is proposed.
RISBD[4] (resource-integrated system for big data)	The RISBD was proposed as a scheme to analyze unused resources among desktop resources being used independently and to integrate the unused resources to provide private cloud-based storage. However, the large amounts of unnecessary overhead required due to IP-based server selection act as a factor for performance degradation. In the present paper, servers are more efficiently selected by considering desktop performance, location, and stored resources.
CSTORE[5] (desktop-oriented distributed public cloud storage system)	The CSTORE was proposed as a scheme to guarantee users' independent spaces and data security in terms of the operation of distributed desktops. Although the CSTORE shows expandability and performance enhancements over public storage systems, it causes a problem in that users cannot gain access following a server fault.

3 Sustainable Operation Algorithms

The SOA proposed in the present paper includes both server and client functions. According to the configuration of operation, it operates as a server or a client. It actively responds to the occurrence of a server fault by automatically selecting a server from among connected clients and operating it.

3.1 Server Side

The server side SOA operation in basic operations is shown in Fig. 1. A major operation is creating metadata when a new desktop has been connected. The created metadata is updated regarding information on the performance, distances, and operating conditions of individual desktops for server replacement, and optimum desktops are prioritized based on the updated information. Thereafter, the prioritized metadata are synchronized to all connected desktops.

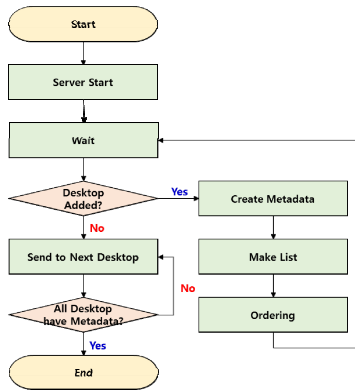


Fig. 1. Server side of SOA

3.2 Client Side

The SOA operation of clients that accessed the server is shown in Fig. 2. Any occurrences of faults in the server are sensed through periodic heartbeats. If the

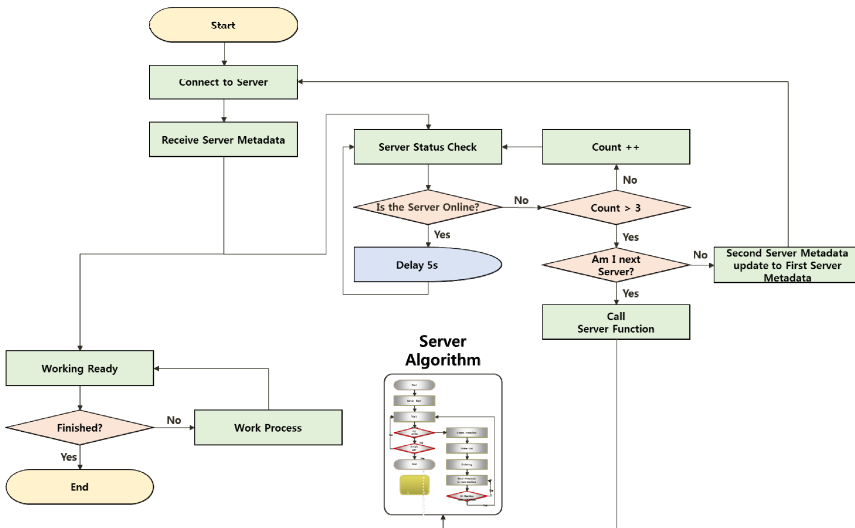


Fig. 2. Client side of SOA

number of faults sensed exceeds a certain number of faults, an alternative server will be recognized through desktop metadata received by the server. If the IP of the alternative server is the same as its IP, the desktop will prepare for operation as a server. If the IP of the alternative server is not the same as its IP, the desktop will prepare for connection to the alternative server. The SOA is operated until at least two desktops are connected to each other. This technique enables active responses to the occurrence of server faults in desktop storage virtualization environments where many desktops are connected to each other.

4 Design of the *ids*-SO

Integrated desktop storage systems applied with the SOA proposed in the present paper are defined as integrated desktop storage system for sustainable operation (*ids*-SO). The *ids*-SO is functionally divided largely into a User Interface that sets the storage and basic information of the host providing storage, a Host Manager that analyzes and manages the connected host, a Resource Manager that integrates resources and creates and manages the metadata of the integrated resources and connected host, a Coordinate Converter that processes system information for visualization on the Viewer, and a Viewer that visualizes the conditions of the host and resources. Each host consists of a Server Check Manager (SCM) that checks the server periodically, a Job Worker (JW) that performs work as requested by the user, and a Server Metadata List (SML) that stores metadata received from the server.

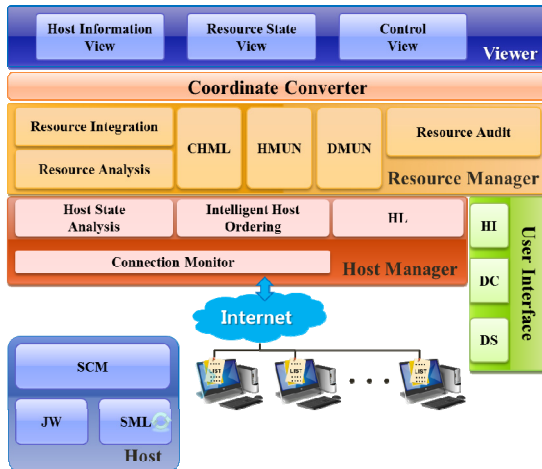


Fig. 3. Architecture of the *ids*-SO

5 Implementation of the *ids*-SO

Figure 4 shows the operation screen in cases where 120 hosts have been connected to the *ids*-SO. Figure 4① shows the total number of hosts, the IP of the server in operation,

and a list of alternative servers that will operate following the occurrence of a server fault. The server in operation is shown at the top of the alternative server list. Figure 4② shows the conditions of integrated CPU, memory, and storage, and Fig. 4③ shows the present situation of the provision of resources by the connected host.

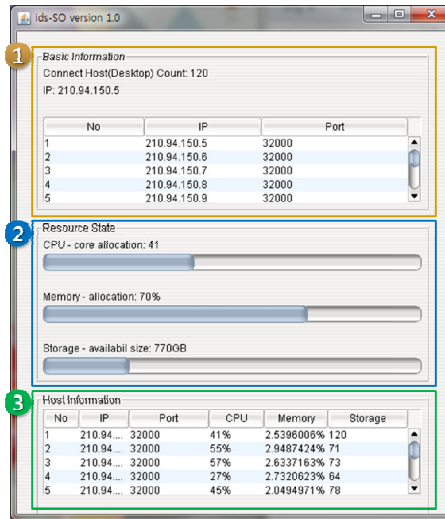


Fig. 4. Operation view for replacement server of *ids-SO*

6 Conclusions

In the present paper, an SOA was proposed that could integrate and operate efficiently distributed desktop resources. With the SOA, when a desktop server fault has occurred, another connected desktop actively operates as an alternative server. This algorithm provided a high availability of the server by being applied even to cases with a minimum two desktops, which were a metanode and a storagenode, connected to each other. In addition, it provides high QoS to storage users by enabling the easy addition and removal of desktops.

Later, algorithms that can respond to the occurrence of server faults in advance of and not following an occurrence will be studied. The algorithms to be studied are to measure the risk of the occurrence of faults based on the state of the desktop operations and prepare the alternative server in advance to enable a more efficient operation.

Acknowledgments. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2014R1A1A2053564). And also this research was supported by the MSP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2014-H0301-13-4007) supervised by the NIPA(National IT Industry Promotion Agency).

References

- [1] Hu, Y., Wang, L., Zhang, X.: Cloud Storage Virtualization Technology and Its Architecture. *Applied Mechanics and Materials* 713-715(2015), 2435–2439 (2015)
- [2] Kleineweber, C., Reinefeld, A., Schütt, T.: QoS-aware storage virtualization for cloud file systems. In: *Proceedings of the 1st ACM International Workshop on Programmable File Systems, PFSW 2014*, Vancouver, BC, Canada, June 23, pp. 19–26 (2014)
- [3] Park, J.H., Kim, H.-W., Jeong, Y.-S.: Efficiency Sustainability Resource Visual Simulator for Clustered Desktop Virtualization Based on Cloud Infrastructure. *Sustainability* 6(11), 8079–8091 (2014)
- [4] Kim, H.-W., Park, J.H., Jeong, Y.-S.: Human-centric storage resource mechanism for big data on cloud service architecture. *Journal of Supercomputing* (online published, February 2015)
- [5] Duan, H., Yu, S., Mei, M., Zhan, W., Li, L.: CSTORE: A desktop-oriented distributed public cloud storage system. *Computers and Electrical Engineering* (online published, November 2014)
- [6] Fu, Y., Xiao, N., Liu, F., Bao, X.: Deduplication Based Storage Optimization Technique for Virtual Desktop. *Journal of Computer Research and Development* S1, 65–70 (2012)
- [7] Gil, J.-M., Park, J.H., Jeong, Y.-S.: Data center selection based on neuro-fuzzy inference systems in cloud computing environment. *Journal of Supercomputing* 66(3), 1194–1214 (2013)
- [8] Buyya, R., Murshed, M.: GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for Grid computing. *Concurrency and Computation: Practice and Experience* 14(13-15), 1175–1220 (2002)
- [9] Jeong, Y.-S., Kim, H.-W., Jang, H.J.: Adaptive resource management scheme for monitoring of CPS. *Journal of Supercomputing* 66(1), 57–69 (2013)
- [10] Song, E.-H., Kim, H.-W., Jeong, Y.-S.: Visual Monitoring System of Multi-Hosts Behavior for Trustworthiness with Mobile Cloud. *Journal of Information Processing System* 8(2) (June 2012)

Analysis on Temporal Sparsity of Human Face for Various Emotion Stimuli

Unsoo Jang¹, Byoung Cheul Kim¹, Yoonkyoung Kim¹,
Min Woo Park¹, and Eui Chul Lee^{2,*}

¹Department of Computer Science, Graduate School, Sangmyung University,
Seoul, Republic of Korea

{dnstn222, soccer0751, kim_yk2606}@naver.com,
nogood79dle@gmail.com

²Department of Computer Science, Sangmyung University, Seoul, Republic of Korea
eclee@smu.ac.kr

Abstract. Recently, many researchers have been focused on partial region of single frame image such as mouth, eyes in order to recognize facial expression. However, according to the recent psychological theory, the other regions of face included important factors in terms of facial expression. To solve the problem of previous works, our proposed method verifies facial temporal movement in the face region. For that, one hundred regions (10×10) are defined onto the detected face region. In addition, temporal sparsity of each region was calculated on specific frequency band well known to express micro-movement of human body. To perform the calculation, adjacent frame subtraction method was used. Consequently, sparsity data against four emotion stimuli was obtained. At result, we confirmed that spatial characteristics of temporal sparsity were different accordance with different emotions.

Keywords: facial expression, facial temporal movement, sparsity.

1 Introduction

Recently, many researches have been performing for emotion classification by extracting and calculating feature of physiological characteristics. For that, signals for measuring intrinsic parameters such as CNS (Central Nervous System) and ANS (Autonomic Nervous System) responses and images for measuring extrinsic responses such as human behavior and face can be used.

In a recent research, SVM (Support Vector Machine) was used to study solution for recognizing facial expression [1]. However, the SVM based method requires well-sampled training data collected from well-defined area. That is, classification performance of SVM was depending on the quality of collected training data. In the different method, Eigen-points based method was performed by measuring distance between two feature vectors extracted from the specific expression and expressionless image [2]. However, because the method is to require excessive facial expression, the method is not suitable in the case of using subtle facial expression.

* Corresponding author.

To solve these problems, in our previous work, facial temporal sparsities were analyzed in order to classify facial expression [3]. Facial temporal sparsities between opposite emotions were compared based on successive frame subtraction method. In detail, we defined five ROIs (Region Of Interests) and calculated the corresponding sparsities.

In our proposed method, adjacent frame subtraction method was used to obtain temporal sparsity data caused by facial expression for four emotion stimuli. According to "action unit" of facial muscle by Ekman [4], because the other sub-regions of face excluding mouth and eyes are also moved along the emotion, we define one hundred ROIs as 10×10 on the entire face region. In addition, we analyze entire facial movement at a specific frequency band (12.5Hz) without transform method to frequency domain such as FFT (Fast Fourier Transform) and DCT (Discrete Cosine Transform). In here, the used frequency band is known for meaningful frequency band in terms of measuring micro-movement [5]. Consequently, the facial temporal sparsities of the four types of emotional stimuli were comparatively analyzed in entire facial movement.

2 Proposed Method

2.1 Face ROI Definition

Our First, we detect face region by using Adaboost algorithm [6]. Then, we define one hundred ROIs (10×10) in the detected face region as shown in Fig. 1.

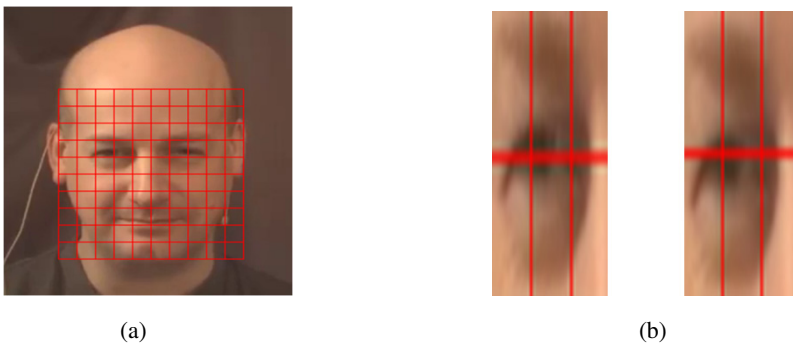


Fig. 1. The used ROI definition and an example of position variation. (a) The face region detected by using Adaboost algorithm and divided into 10×10 sub-regions. (b) Position Variations in detected face region by using Adaboost algorithm.

Although face is not moved, when detecting face region using Adaboost algorithm, two problems can be occurred. The first problem is about position, which means the center of detected region is unstable. The second problem is about size variation. The size of detected face region is not constant in every frame. However, because subjects sit and watch toward camera in the video used in experiment, the facial movement is comparatively little. To solve first problem, we fix the detected size

which is determined from the first frame. Then, we defined ROIs based on detected face-region. To guarantee stable ROIs definition, we divided detected face-region into one hundred ROIs as shown in Fig. 1.

2.2 Sparsity Calculation

In our method, sparsity can be defined by the amount of movement in each sub-region on the detected face region. Higher and lower sparsities mean less and greater movements, respectively. For example, if there is the region whose sparsity is low in case of a specific emotion stimulus, the region is determined as meaningful one in terms of recognizing the emotion. To measure the degree of sparsity, subtraction method between two temporal adjacent frames is used in which the amount of movement is calculated in each ROI.

A previous research reported that the human micro-movement was well observed at about 10Hz [4]. Based on that, we analyze the sparsity at temporal frequency bands as fixed 12.5Hz. However, the traditional transform methods into frequency band such as FFT and DCT are not used. Instead, the interval between two adjacent frames is manipulated. The used video is temporally sampled at 50 FPS (Frames Per Second). In here, the interval between two adjacent frames intends temporal movement at 50Hz. Therefore, the amount of movement at 12.5Hz can be calculated by subtracting between a frame and the next fourth frame.

At the stage of calculating the movement between two frames mentioned in section 2.1, the solution about position problem is required. Actually, despite of no movement, difference between two pixels can be appeared because of variation of the position of detected face. To solve this problem, we adopt shift-matching scheme as used in our previous work [3]. Because the used video has high temporal resolution (50 FPS), the large amount of facial movement cannot be occurred. Actually, we confirmed that the amount of face position's variation is maximum horizontally ± 3 pixels and vertically ± 5 pixels, respectively. Therefore, the range of shift-matching is determined as the mentioned size. The difference between overlapped regions is calculated by extracting LBP (Local Binary Pattern) and calculating HD (Hamming Distance) [7]. At result of shift-matching, the minimum HD value is regarded as the difference between overlapped ROIs.

3 Experimental Result

For experiment, we used dataset provided by the International Conference on Multimedia and Expo (ICME 2010), which is a subset of the SEMAINE corpus [8]. The frame rate of video is 50 frames per second, and the spatial resolution is 640×480 pixels. Total 42 videos were included about four types of emotions in the used dataset in which the analyzed length of each video was 90 seconds.

We analyzed the amount of movement in each ROI for four type emotions. And the comparison for each corresponding ROI was performed. As a result, the features for four type emotions such as neutral, happy, angry, and sad are determined.

As the above mentioned, high sparsity region means less movement. Therefore, small amount of difference was extracted in the high sparsity region. In contrast, low sparsity region has the large amount of difference. Consequently, the region with large difference is regarded as meaningful one in terms of the corresponding emotion.

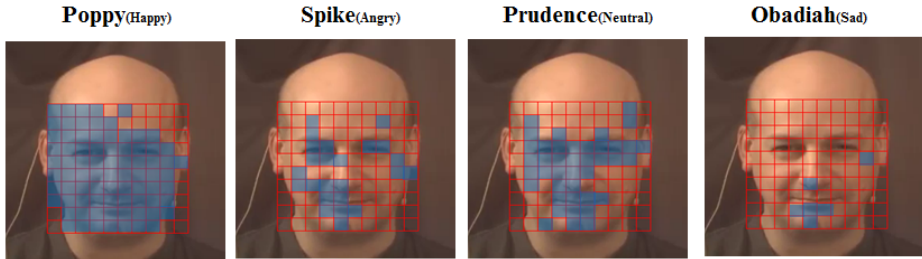


Fig. 2. Significant low sparsity regions for four emotions

In Fig. 2, greater differences were appeared around eye, nose, and mouse regions in case of Prudence (Neutral) emotion. Spike (Angry) emotion generated less differences compared with the Prudence emotion. In case of Poppy (Happy) emotion, significant differences were totally appeared. Among all cases of emotion stimuli, Obadiah (Sad) has the least movements. Consequently, we found that the order of greater facial movements was “Happy > Neutral > Angry > Sad”.

4 Conclusion

In this paper, we proposed a new method which determined significant regions of face in terms of given four types of emotions such as Neutral, Happy, Angry, and Sad. We extracted sparsity features of face regions caused by each emotion stimuli. Face region was detected by using Adaboost, and two problems such as position and size variations were solved. Then, we defined one hundred ROIs (10×10) in the detected face region in order to analyze partial facial movements. At the stage of sparsity calculation, shift-matching scheme was adopted to solve position variation. The difference between overlapped regions was calculated by extracting LBP and calculating HD. At result of shift-matching, the minimum HD value was regarded as the difference between overlapped ROIs. Consequently, facial sparsities of four type emotions were comparatively analyzed.

In future work, we will analyze our experimental results based on Ekman’s facial action unit.

Acknowledgement. This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1014) supervised by the IITP(Institute for Information & communications Technology Promotion).

References

1. Heisele, B.: Face Recognition with Support Vector Machines: Global Versus Component-based Approach. In: 8th IEEE International Conference on Computer Vision, pp. 688–694 (2001)
2. Hong, S., Byun, H.: Facial Expression Recognition using Eigen-points. In: Spring Conference of Korean Institute of Information Science and Engineering, vol. 31, pp. 817–819 (2004)
3. Kim, Y., Kim, H., Lee, E.: Emotion Classification using Facial Temporal Sparsity. *International Journal of Applied Engineering Research* 9(24), 24793–24801 (2014)
4. Ekman, P., Friesen, W.: *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto (1978)
5. Park, S., Ko, D., Whang, M., Lee, E.C.: Vision Based Body Dither Measurement for Estimating Human Emotion Parameters. In: Kurosu, M. (ed.) *HCI/HCI 2013, Part V. LNCS*, vol. 8008, pp. 346–352. Springer, Heidelberg (2013)
6. Viola, P., Jones, M.J.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
7. Yang, H., Wang, Y.: A LBP-based Face Recognition Method with Hamming Distance Constraint. In: 4th International Conference on Image and Graphics, pp. 645–649 (2007)
8. McKeown, G., Valstar, F., Cowie, R., Pantic, M.: The SEMAINE Corpus of Emotionally Coloured Character Interactions. In: 2010 IEEE International Conference on Multimedia and Expo, pp. 1079–1084 (2010)

Monocular Eye Tracking System Using Webcam and Zoom Lens

Kun Ha Suh¹, Yun-Jung Kim², Yoonkyoung Kim¹, Daejune Ko¹, and Eui Chul Lee^{3,*}

¹Department of Computer Science, Graduate School, Sangmyung University,
Seoul, Republic of Korea

{tjrjsgk, kim_yk2606}@naver.com, Kodeajune@gmail.com

²Initiativesix Co., LTD., Seoul, Republic of Korea

yjkim_ceo@initiativesix.com

³Department of Computer Science, Sangmyung University, Seoul, Republic of Korea
eclee@smu.ac.kr

Abstract. Many researches about eye tracking have been done in field of human computer interface, virtual reality, and so on. However, most of eye tracking devices are priced high because of its complicated hardware. Hence, in this paper, we propose a simple real-time eye tracking system using web camera. In our system, we use a zoom lens for web camera, which cover the low resolution of eye region by wide-view capturing. In addition, an infrared light illuminator is required because the gaze position is calculated by using the positional relationship between the pupil and specular reflection caused by the illuminator. Our system measures the gaze position using just one eye, and does not require to wear any devices onto user's head. The average error of gaze point estimation was about 0.9 degree at the Z distance of 70cm on the 23inch size monitor. Experimental results showed that the proposed system can be used enough for navigation or pointing interaction.

Keywords: monocular, eye tracking, infrared illuminator, specular reflection.

1 Introduction

Eye tracking is as a technology to track the gaze point of user, which has wide range of applicability. Many applications using eye tracking have been utilized in areas such as computer interface for hand disabled, human computer interface, immersive game, measure the effect of PPL(Product PLacement), and so on. For more accurate and convenient, many kinds of methods have been researched [1][2][3][4].

Previous eye tracking methods can be categorized into two groups: (1) wearable camera based method and (2) remote camera based method. In the wearable camera based methods, some equipments are required to wear for eye tracking [5]. The method of this type has the advantage that a large and accurate eye image for calculating gaze point can be obtained by capturing the eye at close distance. Also, facial movements are allowed because the device moving along the face. However, wearable device may provide discomfort to the users. In addition, when the worn equipment is

* Corresponding author.

moved by nature or mistake after initial calibration, it is difficult to detect the correct gaze point.

In the other category, remote camera based methods is used to track the eye without wearing any equipment. In this case, to guarantee high spatial resolution in far distance complex device or algorithm is needed such as an unusual camera which has PTZ(Panning & Tilting & Zooming) mechanism. These types of method are convenient but cost is high.

According to a different point of view, eye tracking method can be categorized into two groups: (1) binocular gaze detection and (2) monocular gaze detection methods. Commonly, binocular gaze detection method is more accurate than monocular one because the more information for deciding gaze point can be extracted [6]. In monocular gaze detection method, whereas, the gaze point is calculated by only one eye but gaze estimation time is decreased by processing one eye compared to the by binocular gaze detection [7].

To solve problems of wearable and binocular methods, in this paper, a new eye tracking system is proposed. To capture an image which has enough spatial resolution for eye tracking from USB(Universal Serial Bus) webcam, a default lens is replaced to a zoom lens. Eye region is detected by using the properties of the dark pupil and bright SR(Specular Reflection) and the center positions of the pupil and SR is sequentially detected through several processes. A detailed process is provided in Sec. 2.2. After initial user-dependent calibration by gazing only 4 points, gaze positions are successively calculated through adopting the mapping function between the pupil-SR vector and the four corner coordinates of the monitor which extracted by the calibration.

The rest of this paper is explained as follows. The proposed system and method are described in Sec. 2. The experimental results and conclusions are presented in Sec. 3 and 4, respectively.

2 Proposed System and Method

2.1 Proposed System

The proposed system consists of one webcam, one infrared light illuminator and one zoom lens for webcam as shown in Fig. 1.

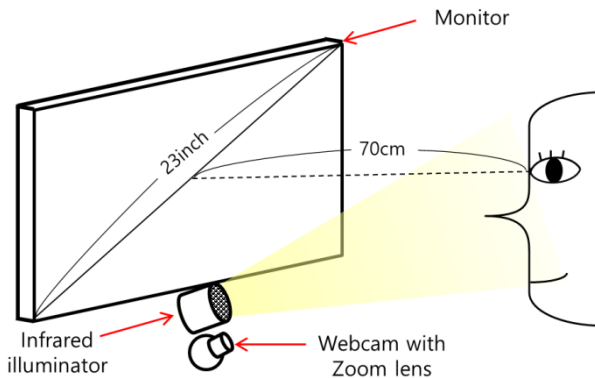


Fig. 1. Proposed eye tracking system

The used infrared light illuminator includes 50 IR-LEDs (Infrared-Light Emitting Diode) whose wavelength and illumination angle are 850nm and 30 degrees, respectively [8].

In our experiments, we used a Logitech Webcam C600 whose spatial resolution is 1600×1200 pixels [9]. To obtain the image of dark pupil and bright SR, we converted the camera to an infrared one. The infrared cutting filter of the camera is removed and infrared passing filter is attached to its place. However, since the used camera alone is not enough to ensure sufficient spatial resolution to calculate position accurately, the zoom lens is attached to the camera. As a result, it is possible to detect the accurate pupil and SR at a distance of 70cm. Also, infrared illuminator and camera are located on the bottom of the monitor as shown in Fig. 1 to avoid pupil occlusion by eyelid.

2.2 Proposed Method

The processing stage of proposed system includes three main steps. First, detecting eye region is an essential prerequisite to detect the location of pupil and SR. In our method, we use the left side eye for eye tracking. To detect an eye region, the input image is divided into sub-blocks of 64×60 size (in Fig. 3 (a)) [7]. In each sub-block, the values of maximum and minimum pixel level are measured. Then, the difference between them is calculated. Based on the difference, the sub-blocks are sorted in descending order and upper three regions are selected as an eye candidate region (in Fig. 3 (b)). Histogram stretching is applied into each candidate region then the number of black pixels is counted. Consequently, the region of including the most black pixels is regarded as eye region.

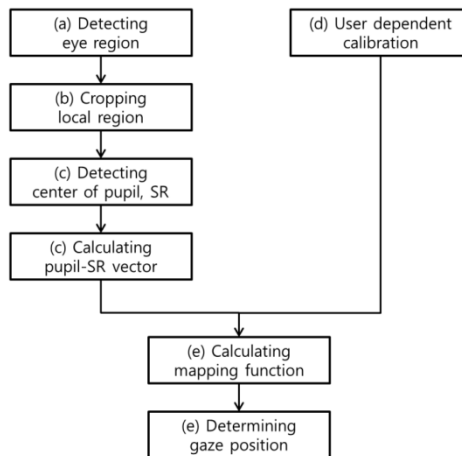


Fig. 2. Flow chart of the proposed method

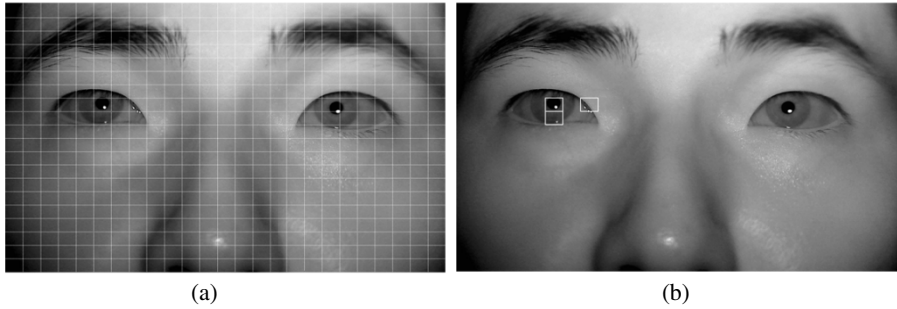


Fig. 3. Example of the eye region detection process. (a) Divided sub-blocks on the original image, (b) Three candidate regions of eye.

Next, based on the detected eye region, local region of 192×120 size is cropped from the original images. And then, calculating the center of the pupil and SR is divided sub-steps again as shown in Fig 4. For that, the binarization, morphological operation, and component labeling are sequentially performed for pupil and SR, respectively after histogram stretching. Then, the final positions of pupil center and SR center are obtained through ellipse fitting algorithm [10].

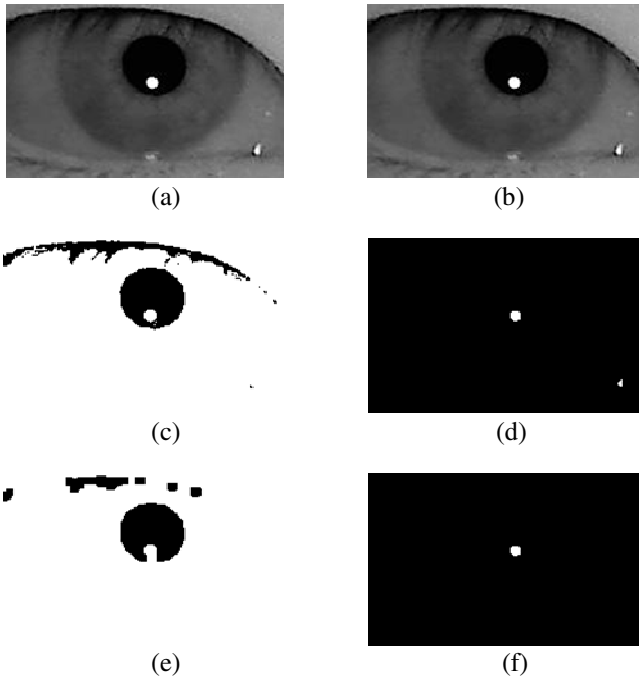


Fig. 4. Results of each step for calculating center of the pupil and SR. (a) Original local-region. (b) Histogram stretching result of (a). (c) Binarization result of (b) for the pupil. (d) Binarization result of (b) for the SR. (e) and (f) Each morphological operation result of (c) and (d), respectively. (g) and (h) Each component labeling and conditional filtering results of (e) and (f), respectively. (i) Detected center of the pupil and SR using ellipse fitting algorithm.

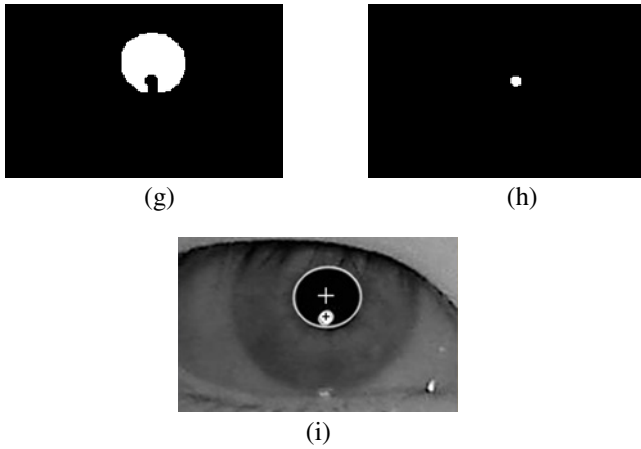


Fig. 4. (continued)

Finally, we use the just one eye for eye tracking by analyzing magnitude and direction of the pupil-SR vector as shown in Fig. 5 [8]

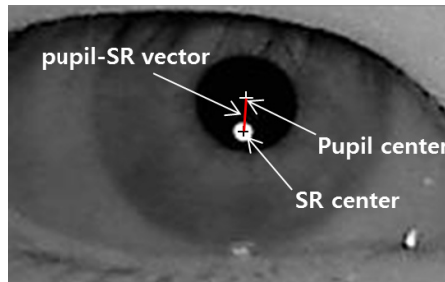


Fig. 5. Example of the pupil-SR vector

Before calculating the gaze position, user-dependent calibration should be performed as the step (d) of Fig. 2. In our method, 4-points calibration scheme is adopted. When a user gaze four corners of monitor $((m_{x1}, m_{y1}), (m_{x2}, m_{y2}), (m_{x3}, m_{y3}), (m_{x4}, m_{y4}))$, four pupil-SR vectors $((x_{pr1}, y_{pr1}), (x_{pr2}, y_{pr2}), (x_{pr3}, y_{pr3}), (x_{pr4}, y_{pr4}))$ of the eye can be defined. Then, by using the above coordinates, the gaze position on the monitor plane (G_x, G_y) is calculated through following two formulas.

$$M = T \cdot V \tag{1}$$

$$\begin{bmatrix} m_{x1} & m_{x2} & m_{x3} & m_{x4} \\ m_{y1} & m_{y2} & m_{y3} & m_{y4} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{pr1} & x_{pr2} & x_{pr3} & x_{pr4} \\ y_{pr1} & y_{pr2} & y_{pr3} & y_{pr4} \\ x_{pr1}y_{pr1} & x_{pr2}y_{pr2} & x_{pr3}y_{pr3} & x_{pr4}y_{pr4} \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} G_x \\ G_y \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{pr} \\ y_{pr} \\ x_{pr}y_{pr} \\ 1 \end{bmatrix} \tag{2}$$

3 Experimental Results

The proposed eye tracking system was tested with a Intel Core i5(2.30Hz) CPU and 4GB RAM. The program was implemented using Visual C++ with OpenCV library [11]. The size and the spatial resolution of used monitor in experiment were diagonally 23 inch (16:9) and 1920×1080 pixels, respectively. A total of 5 persons participated in the experiments and to validate the accuracy of the proposed system, we did a test to gaze the 9 reference points. Each person performed the test 5 times.

Experimental results showed that the RMS (root mean square) error is approximately the angular error of 0.9 degrees by 46 pixels, which despite using the webcam, it has fairly high accuracy.

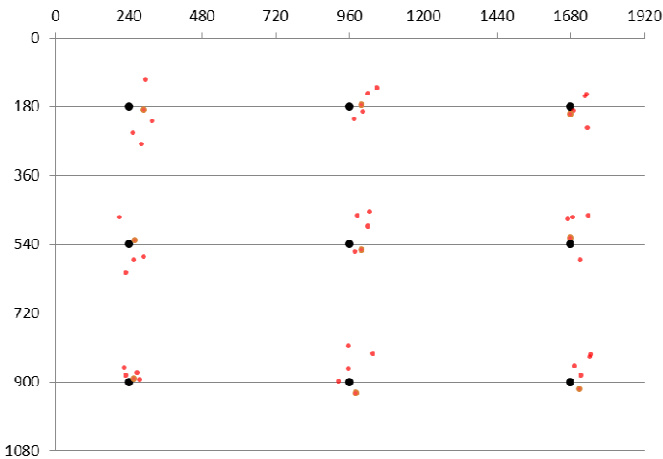


Fig. 6. Five examples of the test for measuring gaze estimation accuracy

4 Conclusion

In this paper, we proposed a new eye tracking system using webcam. The proposed system is very cheap and simple because they do not use any complex or expensive hardware. Besides, it is very convenient because there is no need to wear any equipment.

In future works, we will study a method for increasing the accuracy of the gaze point and simplify the procedure of user-dependent calibration.

Acknowledgement. This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1014) supervised by the IITP(Institute for Information & communications Technology Promotion).

References

1. Matsumoto, Y., Ino, T., Ogsawara, T.: Development of intelligent wheelchair system with face and gaze based interface. In: 10th IEEE International of Robot and Human Interactive Communication, pp. 262–267 (2001)
2. Zhu, Z., Ji, Q.: Eye and gaze tracking for interactive graphic display. *Machine Vision and Applications* 15, 139–148 (2004)
3. <http://www.tobii.com/en/eye-experience/eyex> (accessed on March 18, 2015)
4. Kules, B., Capra, R., Banta, M., Sierra, T.: What do exploratory searchers look at in a faceted search interface? In: 9th ACM/IEEE CS Joint Conference on Digital Libraries, pp. 313–322 (2009)
5. Lee, E.C., Park, K.R.: A study on eye gaze estimation method based on cornea model of human eye. In: Gagalowicz, A., Philips, W. (eds.) *MIRAGE 2007*. LNCS, vol. 4418, pp. 307–317. Springer, Heidelberg (2007)
6. Cho, C., Lee, H., Gwon, S., Lee, J., Jung, D., Park, K., Kim, H., Cha, J.: Binocular gaze detection method using a fuzzy algorithm based on quality measurements. *Optical Engineering* 52(5), 053111 (2014)
7. Lee, E., Park, K.: Non-wearable eye tracking method using single camera with single illuminator. In: 2nd International Conference on Internet, pp. 170–173 (2010)
8. <http://www.koditec.com/sub211.html?cid=142&pid=254> (accessed on March 18, 2015)
9. <http://support.logitech.com/product/2-mp-webcam-c600> (accessed on March 18, 2015)
10. <http://www.sciencedirect.com/science/article/pii/S0169260798001059> (accessed on March 18, 2015)
11. <http://sourceforge.net/projects/opencvlibrary> (accessed on March 18, 2015)

Many-to-Many Data Collection for Mobile Users in Wireless Sensor Networks

Chi-Fu Huang* and Wei-Chen Lin

Department of Computer Science and Information Engineering,
National Chung Cheng University
Chia-Yi, Taiwan, R.O.C.
cfhuang@cs.ccu.edu.tw

Abstract. Data collection is one of the fundamental functions in Wireless Sensor Networks (WSNs). Different from traditional data collection mechanisms which only consider a single and stationary sink, this paper studies this problem in a scenario with multiple mobile sinks. The motivation is for a WSN to support future applications, such as Internet of Things (IoT). In this case, a WSN is required to be able to deliver sensing results to multiple users moving around the network. There are two difficulties in this problem: sink mobility and multiple sinks. Since sinks are mobile, data delivery paths need to be updated frequently, which causes huge maintenance cost. To resolve this problem, we propose a hop-count based data collection architecture together with an efficient mobility management scheme. On the other side, sensing results from a large number of sensors are sent to multiple sinks, which causes lots of packet transmissions. To resolve this problem, we combine the idea of multicast and data aggregation. We first prove that the optimal multicast decision is a NP-hard problem and then propose a distributed heuristic solution. In addition, we further integrate data aggregation into multicast and propose a distributed many-to-many aggregation mechanism. Simulations are constructed to show the efficiency of the proposed schemes. The results show that both our multicast method and many-to-many aggregation method can efficiently reduce communication cost when delivering data to multiple mobile sinks.

Keywords: sensor networks, data collection, mobile sink, multiple sink, multicast, data aggregation.

1 Introduction

A wireless sensor network (WSN) consists a large number of tiny, inexpensive, and limited power sensor nodes. They are randomly deployed in a specific area to monitor and collect the environmental information. In the last decade, WSNs have become one of the most popular research areas because of its wide range of developing applications, such as habitat monitoring, healthcare system, and forest surveillance. Due to the limited power of sensor nodes, there are many constraints when designing a WSN, such as energy, transmission range, memory, and computing power.

* C.-F. Huang's research is sponsored by NSC under Grant No.103-2221-E-194 -023 -MY2.

Data collection [1][2][3] is one of the fundamental functions in a WSN. Mostly, it is assumed that there are many sensor nodes and a single sink, who are all stationary, together constructed a data collection network. Some sensor nodes act as Source Nodes (SNs) if they sense and report events to the sink. All the data packets are transmitted hop by hop via wireless transmission. The sink collects, analyzes, and saves results to an external storage server for further applications. A data collection tree [4] rooted at the sink is the most widely accepted solution when constructing collection routing paths. It is efficient since all the sensor nodes and sink are stationary. Once a tree is created, it can be used for a long period of time.

In literatures, sink mobility and multiple sinks are proposed to relieve the hotspot problem [1][2][5]. Sensor nodes closer to the sink exhaust their own energy more quickly. In this paper, we consider multiple mobile sinks from the application aspects. Many applications require a WSN allow mobile users moving around the network to have abilities to access the network, such as a firefighter system [6]. Besides, the rise of Internet of Things (IoT) increases the demand of mobile sinks. Data demanders in IoT, such as humans or vehicles, may be roaming arbitrarily in the network. Efficient data access mechanisms to support mobile sinks are expected

To support multiple sink mobility, we construct a hop-count based data collection architecture and propose an efficient management scheme modified from [1]. Based on this architecture, we then investigate many-to-many communication issues. In our scenario, there are multiple mobile sinks and multiple SNs in a data collection network. If traditional tree structure is adopted, multiple trees rooted at each sink are established. Data packets produced by each SN are transmits individually. Here we propose a many-to-many aggregation scheme to efficiently deliver data packets to multiple mobile sinks. The proposed scheme combines the ideas of two well-known technologies: multicast [7] and data aggregation [3]. We first design a distributed multicast scheme, and then propose a data aggregation mechanism to enhance aggregation probability. Final, we combine multicast and data aggregation into a distributed many-to-many aggregation.

The rest of the paper is organized as follows. Section 2 presents our system model. The system design is presented in Section 3. Section 4 contains our simulation results. Section 5 concludes this paper.

2 Network Model

Given a sensor network, there are n static sensor nodes and m mobile sinks in the network. Let S denote the sensor set, where $S = \{s_i | i = 1, 2, \dots, n\}$, and D denote the sink set, where $D = \{d_n | n = 1, 2, \dots, m\}$. Sensor nodes are assumed to be the same, i.e. identical hardware and ability. The transmission range of a sensor is denoted as R . N_i is the neighboring set of sensor s_i , where $N_i = \{s_j | dist(s_i, s_j) < R\}$ and $dist(s_i, s_j)$ is the Euclidean distance between sensor s_i and s_j . We assume under-layer transmissions are reliable and focus our design on the data collection mechanism of routing layer. Each sensor is unaware of its own location. Only neighboring nodes in N_i are possible to be selected to relay data packets. S' is a subset of S , which are used

to denote source nodes (SNs). Each mobile sink in D is assumed to move arbitrarily in the network. In addition, it selects a nearby sensor node as its agent. An agent is responsible for relaying sensing data collected from S' for its sink. Whenever a sink moves, the agent changes accordingly. The whole process is, when an event is detected by a SN s'_i in S' , this event is requested to be delivered to all sinks in D . s'_i first relays the sensing packet to all agents through the network in a multi-hop manner and then agents forward the packet to their sinks.

Our goal is to design a data collection mechanism to deliver all sensing packets from S' to D with the minimum cost.

3 System Design

3.1 A Hop-Count Based Architecture and Mobility Management

To support our many-to-many aggregation, a hop-count based architecture is established. At beginning, each sink floods an initiation message which consists of $h_{i,j}$, that is the hop-count distance between the current receiving sensor node s_i and sink d_j . $h_{i,j}$ field is set to zero initially and broadcasted by agent node of d_j . It is increased by one once the initiation message is rebroadcasted. After flooding through the network, sensor nodes are aware of their shortest hop-count distances to each sink. By exchange of periodical beacon messages with neighbors, each sensor s_i can obtain hop-count information to each sink of its neighbors, and then establish a parent set $M_{i,j}$. $M_{i,j}$ is a subset of N_i and consists of neighbors who have less hop-count distance to d_j . When delivering a data packet to d_j , s_i can only choose the next relay node in $M_{i,j}$. This architecture is much more flexible. Each sensor node has more choices to relay packets compared with tree structure. As a result, data transmissions can be more efficient and energy consumption can be more balance.

Once a sink move, the agent change and the hop-count distance needed to be update accordingly. After a sink moves and chooses a nearby sensor node s_v as a new agent, s_v broadcasts an update message. The update message includes a hop-count distance, $h_{u,v}$, from the old agent s_u to s_v , and a hop-count distance, $h_{u,v}$, directly from s_v to the received node s_i . When a node s_i receives an update message received from an intermediate node s_j , it compares the old path, delivery to s_u to and then forward to s_v , and the new path, directly delivery to s_v , by calculation of the following equation,

$$\frac{h_{u,v}+h_{u,i}}{h_{v,i}} > \lambda, \quad (1)$$

where $\lambda \geq 1$ is a predefined constant. If this equation is satisfied, i.e. the new path is λ times shorter than the old path, the update message is rebroadcasted. Node s_i abandons the old parent set and includes s_j as its parent set. Further intermediate nodes through different paths with the same hop-count distance are included in the new parent set. Otherwise, if the equation is not satisfied, the message is discarded. No more downstream node will be updated. By different settings of λ , we can control the update range and frequency.

3.2 Many-to-Many Aggregation

We summarize the overall process of the proposed many-to-many aggregation. When a sensor node s_i receives a data packet, the packet is put into the buffer first. Periodically, s_i tries to aggregate and transmit data packets in its buffer. s_i first establishes a destination set D' to record all destinations of packets by parsing all sessions records of all data packets in the buffer. The proposed greedy multicast decision scheme is used to decide relay nodes. If there is more than one neighbor who can serve the same maximum sinks, s_i decide relay node by the proposed probabilistic aggregation decision scheme. After deciding relay nodes, s_i aggregates all the sessions served by the same relay node and transmit to it.

3.2.1 Multicast

Due to the nature of sink mobility, establishment of a global static multicast tree is impractical. We focus our design on local decision. Whenever a sensor receives data packets to be relayed to a certain set of sinks, this sensor needs to decide a subset of neighbors as its next-hop relay nodes. We first give a problem definition. A distributed greedy scheme is then proposed to resolve this multicast decision problem.

At first, we define a metric, *progress*, to represent the benefit of multicast. The progress is defined as,

$$\text{progress} = \text{unicast path length} - \text{multicast path length} \quad (2)$$

When delivering a sensing packet from a sensor to multiple sinks, the unicast length is a summation of lengths of unicast paths, and the multicast path length is a summation of degrees of nodes in the multicast tree. *Progress* can be used to represent improvement of the number of packet transmissions.

Definition 1. One-Hop Maximum Progress (1HMP) problem: Given a sensor node s_i and a subset of sinks $D' \subset D$, the one-hop maximum progress problem is to decide a subset M'_i of N_i to relay packets to D' , so that the progress is maximum.

To prove the 1HMP problem is NP-hard, a well-known NPC problem, the *set cover problem*, is reduced to 1HMP problem. Due to the limitation of pages, details of proof are omitted and can be found in the extension version of this paper.

Greedy Multicast Decision Scheme: When a node s_i receives a data packet to sinks D' , it chooses neighbor node s_j as a relay node first if s_j can serve the most destinations. More specifically, each neighbor's parent set is checked. Who has paths to the most sinks in D' is picked up first. D' is then updated to exclude those sinks has been served. This process repeats until D' is empty.

3.2.2 Data Aggregation

To increase the chance of aggregations, the proposed routing decision is based on two observations. (1) A relay node may have more than one neighbor with the maximum progress. (2) If a node was chosen to relay more data packets, it has more aggregation chances.

Probabilistic Aggregation Decision Scheme: To increase aggregation chances, each node s_i maintains a *session table* which records the number of relays, $r_{j,k}^i$, for each session $L_{j,k}$ from SN s_j to agent d_k . $r_{j,k}^i$ is initiated to zero, and increased by one once node s_i relays a data packet for session $L_{j,k}$. By periodically broadcasting the session table, each node knows session tables of its neighbors. We then calculate the total relay times, $t_i = \sum_{\forall j \in S', k \in D} r_{j,k}^i$. Once there is more than one node with the same maximum progress, they are chosen based on the choosing probability p_i , defined as

$$p_i = \frac{t_i + 1}{\sum_{s_j \in M_i} (t_j + 1)} , \tag{3}$$

where M_i includes nodes who have the same maximum progress.

4 Simulation Results

A simulator is implemented to evaluate the efficiency of the proposed schemes. We compare the proposed multicast and many-to-many aggregation with the tree based scheme [1]. Test results are average of 50 rounds of each network setting and each round of 200 seconds. Performance metrics are communication cost, latency and energy consumption.

We first observe the influence of network density. 250-500 sensor nodes are randomly deployed in the field. As shown in Fig. 1(a), we can find that the communication cost decreases when the network density increases because sensor nodes can find better routing path in a higher network density. We find the propose

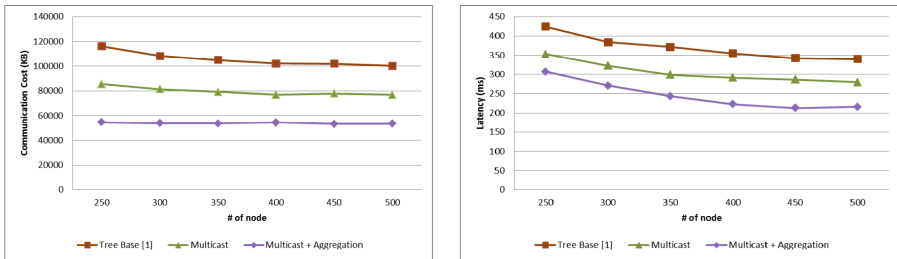


Fig. 1. Communication cost and latency under different networks densities

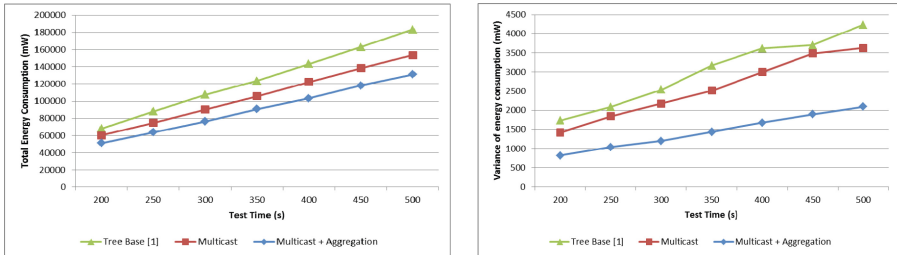


Fig. 2. Total energy consumption and variance of energy consumption

many-to-many aggregation has the lowest communication cost, and the multicast scheme also has lower communication cost than tree based scheme. Both our multicast and many-to-many aggregation can efficiently reduce the communication cost. In Fig.1 (b), latency of the many-to-many aggregation is the lowest because it can reduce the most number of data packets transmitted in the networks.

Next, energy consumptions of different schemes are evaluated. In Fig. 2(a), the many-to-many aggregation consumes lowest total energy, and the multicast also has lower total energy consumption because both the proposed schemes can efficiently reduce network traffic. In figure Fig. 2(b), we compare the energy variance between the maximum and minimum energy consuming sensor nodes. Although our many-to-many aggregation seemingly selects the same group of sensor as relay nodes, the results show the energy variance is still low. That is because the probabilistic aggregation decision scheme can balance energy consumption between nodes.

5 Conclusions

In this paper, we study the data collection problem in a multiple mobile sink scenario. We first propose a hop-count based architecture and an efficient management scheme. To reduce network traffic, we propose a greedy multicast decision scheme to choose next-hop relay nodes to perform multicast. To increase the aggregation chances and further reduce network traffic, a probabilistic aggregation decision scheme is proposed to select between relay nodes. Multicast and data aggregation are integrated into many-to-many aggregation. Simulation results show that both multicast and many-to-many aggregation can efficiently reduce communication cost, latency, and energy consumption.

References

1. Li, Z., Li, M., Wang, J., Cao, Z.: Ubiquitous data collection for mobile users in wireless sensor networks. In: IEEE INFOCOM (2011)
2. Mottola, L., Picco, G.P.: MUSTER: Adaptive energy-aware multisink routing in wireless sensor networks. *IEEE Transactions on Mobile Computing* 10(12), 1694–1709 (2011)
3. Weng, H.-C., Chen, Y.-H., Wu, E.H.-K., Chen, G.-H.: Correlated data gathering with double trees in wireless sensor networks. *IEEE Sensors Journal* 12(5), 1147–1156 (2012)
4. Gnawali, O., Fonseca, R., Jamieson, K., Moss, D., Levis, P.: Collection tree protocol. In: ACM Conference on Embedded Networked Sensor Systems, SenSys (2009)
5. Wu, X., Chen, G.: Dual-Sink: Using mobile and static sinks for lifetime improvement in wireless sensor networks. In: International Conference on Computer Communications and Networks, ICCCN (2007)
6. Salam, H.A., Rizvi, S.R., Ainsworth, S., Olariu, S.: A durable sensor enabled lifeline support for firefighters. In: IEEE INFOCOM Workshops (2008)
7. Sanchez, J.A., Ruiz, P.M., Liu, J., Stojmenovic, I.: Bandwidth-efficient geographic multicast routing protocol for wireless sensor networks. *IEEE Sensors Journal* 7(5), 627–636 (2007)

Method of Path Computation Using National Research Network Federation

Jinhyung Park¹, Joon-Min Gil², Woohyun Kim¹,
Jeongwook Park¹, and Hyunhun Cho^{1,*}

¹ KREONET Center, Korea Institute of Science and Technology Information, Korea

² School of Information Technology Engineering, Catholic University of Daegu, Korea

{ntosk, woohyun, jwpark, hhcho}@kisti.re.kr,

jmgil@cu.ac.kr

Abstract. The legacy networks do not open information about a management domain due to scalability, manageability and commercial reasons. Therefore, it is very hard to compute an optimal path to the destination. This study has aimed to investigate federation to build a united network led by the National Research Network for the independent management and sharing of network resources among countries. For this, ICE (Information Control Element) which manages information in each domain as a top-level PCE (Path Computation Element) which supports traffic engineering and the architecture of Global Federation Organizer which controls the ICE have been designed and applied to a multi-domain network.

This paper performs the selection of optimal path among multi-domains and optimization of end-to-end connection as well as the efficient use of resource.

Keywords: NRN federation, PCE, Global Federation Organizer, ICE.

1 Introduction

Today networks consist of individual autonomous system for scalability and manageability, and commercial reasons. The administrative domains consist of the interior gateway protocol area and the exterior gateway protocol area and information of each domain is not disclosed. Thus computing the optimal path to the destination is very difficult. Due to the low mutual-sharing of the path information, rerouting the optimal path is particularly difficult in case of a failure in the middle of the path. Therefore, the node needs to obtain the information of each domain, and the architecture, which is configured as the top-level path computation node for controlling the information of the node, is necessary.

This paper goal is the network federation by sharing network resources among the national research network in each country while managing the resources of the national research network independently. We designed the architectures of the Information Control Element and the Global Federation Organizer as top-level path

* Corresponding author.

computation elements. These compute an optimal network path to support network traffic engineering. The paper is studied of applying the architectures on the multi-domain network and verifying the federation for independent resource management and resource information sharing among national research networks on a multi-domain network.

2 Path Computation Element (PCE)

This chapter describes network route computation on a multi-domain network. A path computation element is that computes a network path or route based on network topology considering the constraints for the computation.

Most IP-based network traffic engineering solutions operate on a single routing domain. These solutions halt when the path stays the routing area between an ingress node and an egress node, or the AS of an ingress node. For this case, the path computation problem causes complexity since complete routing information on the network is not possible to obtain. Service providers are unwilling to disclose the routing information to the outside of the routing area due to scalability constraints or confidentiality concerns.

There are a number of different PCE methods for a single domain, a multi-layer, a multi-domain or a multi-layer. Here we explain the need of a PCE considered a multi-domain and the dynamically manageable domain-information to utilize the network resource of a multi-domain efficiently.

3 Multi-domain Path Computation

3.1 Per-Domain Path Computation

A per-domain path computation approach progresses at an entry point. Path computation progresses in the domain, and a good-enough path is selected by a destination. Typically it is assumed that the connection configuration between domains is recognized. Information such as IP routing is used for selecting an exit point.

However a PCE1 does not recognize the destination domain for per-domain path computation. Therefore to compensate the defect of per-domain path computation, a signaling approach such as crankback can be employed. However the signaling approach is more complex and takes a longer time compared with per-domain path computation.

3.2 Simple Cooperating PCEs

Simple cooperating PCEs compute an optimal path between adjacent domains using communication between PCEs. First, the ingress node of a source requests path computation to a relevant PCE. The PCE computes the path not only using its own domain information but also querying about the path selection to another PCE that controls the adjacent domain. The adjacent PCE selects the optimal path of its own domains, and responds about the adjacent egress.

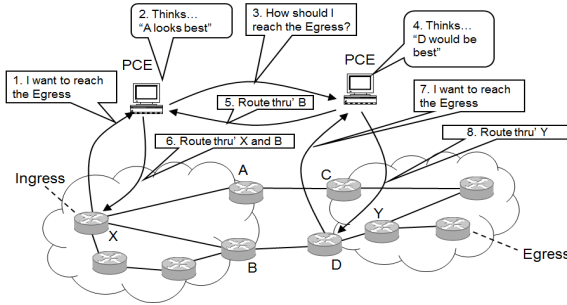


Fig. 1. Example of simple cooperating PCEs

However a simple cooperating PCEs approach cannot provide an optimal path practically when there are more than two adjacent domains. In other words, the proposed optimal path may not be the optimal path from the perspective of another domain.

3.3 Backward Recursive Path Computation

Backward recursive path computation employs the cooperation between path computation elements. Even if the full visibility of a network is not retained, an optimal path can be provided using crank-back signaling for path computation. The path computation of this approach starts from a destination domain and transfers the set of potential paths to an adjacent path computation element. Given that each path computation element computes an optimal path from an entry point to an exit point, and configures the tree with the destination as the root. However this method must recognize a destination domain necessarily.

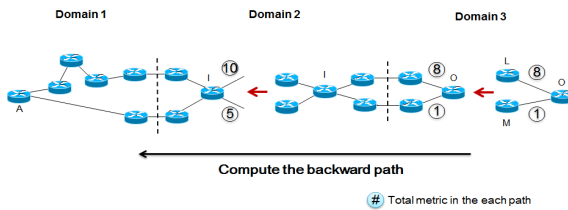


Fig. 2. Backward recursive path computation

4 Global Federation Organizer

We propose Global Federation Organizer, the top-level path computation node and the relevant ICE (Information Control Element), the path computation node in each domain.

4.1 Terminology

Fig. 3 shows the top-level path computation node, Global Federation Organizer architecture, based network. The network collects node information for path computation in a multi-domain environment, and obtains all information between the ICEs.

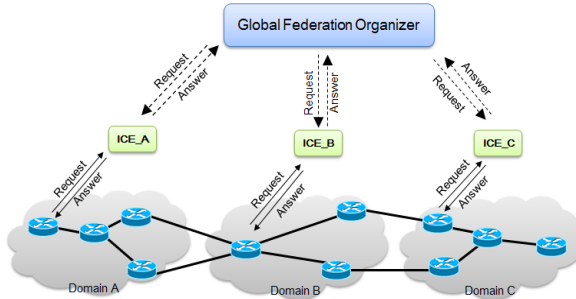


Fig. 3. Overview of Global Federation Organizer & ICE

A. The architecture of Global Federation Organizer

- ICE Information Repository: The ICE Information Repository in Information control element stores the basic domain information of each registered ICE. It transfers the information of ingress/egress nodes and the network performance elements that are required for path computation to the ICE of each domain. The ICE Information Repository then requests the ICE to compute all the available paths. It only stores the minimal information for path computation.

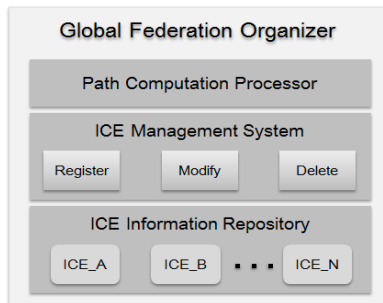


Fig. 4. Architecture of Global Federation Organizer

- ICE Management System: The ICE Management System registers, deletes, or modifies the ICE information in the ICE Information Repository.

- Path Calculation Processor: The Path Calculation Processor calculates an optimal path using the combination of the path result data of the domain requested to the each ICE by the ICE Information Repository. The result is transferred back to the ICE, which requests the path computation.

B. The architecture of ICE

- Route Computation Client: The Route Computation Client sends a route request to the ICE in accordance with the route initializing request from a user or a particular application. In general, the Route Computation Client is an edge/ingress node of a network.
- Route Computation Element: The Route Computation Element computes all path and transfers to the ICE Information Repository based on the node information of ingress/egress, and the information of network performance elements that are transferred from the ICE Information Repository by the path computation request of the ICE Information Repository in the Global Federation Organizer.
- Traffic Engineering Database: The Traffic Engineering Database is configured based on network domain resources and the information of network topology. These information include bandwidth, delay time (delay, jitter). The ICE selects an optimal path that satisfies the requests from ICE based on the information.
- Node/Device Information Database: The database stores the information of the basic network connection of the relevant domain and renews the information regularly for more efficient path computation.

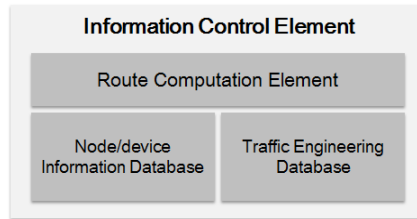


Fig. 5. Architecture of ICE

4.2 Application of Global Federation Organizer/ICE in Multi-domain

The ICE_A requests the path from the starting-point router A to the destination router K. The Global Federation Organizer, received the optimal path computation request from ICE, transfers the required information for the optimal path computation (e.g. destination information) to the ICE_A, B and C. The Global Federation Organizer also requests the ICEs to compute the entire available paths between the ingress and egress node, and transfer the result back to the Global Federation Organizer. The Global Federation Organizer recognizes the domain of the destination router K, and the information that the ICE_C controls the path administrative system of the domain. Given that the Global Federation Organizer computes the path using a backward recursive approach based on the path computation result provided from the ICE3. The Global Federation Organizer then transfers the computed optimal path to the ICE1.

Ultimately the Global Federation Organizer computes the optimal path by combining the required path information for the computation provided by all the ICEs. The Global Federation Organizer selects the optimal path that has the minimum value of the cost or delay by the parameter selections for the transfer characteristics among all the available candidates structured in a tree form like the Fig. 6.

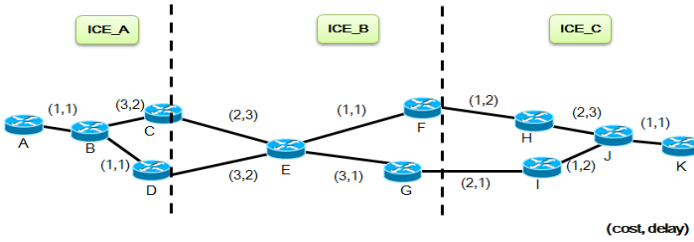


Fig. 6. Metric information in multi-domain

Fig. 7 shows the result of the cost and delay computation for a multi-domain by combining the information of the ICE1, the ICE2 and the ICE3 suggested in Fig. 6.

- ICes calculate available paths in each domain and send the results to Global Federation Organizer.
- Global Federation Organizer computes the optimal path from source(A) to destination(K).
 - (ABC) (EF) (HJK)
 (4, 3) (3, 4) (4, 6) → **(11, 13)**
 - (ABD) (EF) (HJK)
 (2, 2) (3, 4) (4, 6) → **(9, 12)**
 - (ABC) (EG) (IJK)
 (4, 3) (6, 3) (4, 4) → **(14, 10)**
 - (ABD) (EG) (IJK)
 (2, 2) (6, 3) (4, 4) → **(12, 9)**
- Global Federation Organizer returns the result to ICE_A
- ICE_A choose a path with Min(cost or delay or hop) by data characteristics

Fig. 7. Final Path Computation in GFO

Fig. 8 shows the optimal path that selected the cost (throughput) as a transfer parameter among all the computed path candidates for a multi-domain.

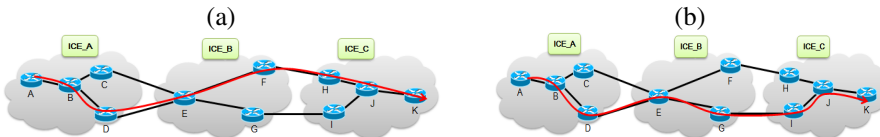


Fig. 8. Path Selection by cost (a) and delay (b)

5 Performance Evaluation

5.1 Testbed Organization

We tried to evaluate the effectiveness of the proposed multi-domain network-based Global Federation Organizer/ICE architecture. We configured the test-bed to

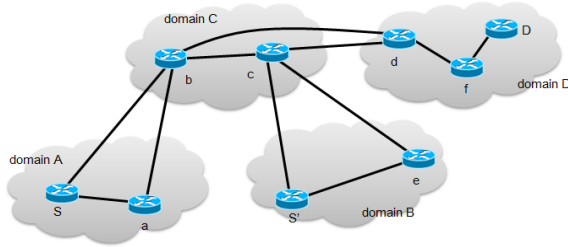


Fig. 9. Testbed for Global Federation Organizer/ICE

evaluate the model that can compute an optimal path by the representative transfer characteristics Fig. 9.

5.2 Configuration and Performance

We evaluated the architecture for transferring the large data which is restrictive transfer rate and bandwidth, and the over HD image which is sensitive to delay.

The test-bed consists of four domains with either 1 or 10 Gbps for a multi-domain like Fig.10. The S in domain A and S' in domain B are starting points, and the D in domain D is the destination. We evaluated the transfer rate and the delay time (delay, jitter) between the optimal paths based on the minimum number hops, and the proposed path selection model. The result is as follows:

Table 1. Comparison between available paths

	Throughput (Mbps)	Packet loos rate (%)	Delay (ms)
Path 1(min hops)	246	0.02	13.4
Path 2(throughput)	322	0.01	14.1
Path 3(delay)	243	0.01	5.6

For data transmission, the throughput of the path 2 considered the cost improves by 33% compared with the path 1 based on the minimum number of hop. For real-time HD video data transmission, the delay time of the path3 considered the delay time reduces 5.6 ms compared with the path 1.

6 Conclusion

On a multi-domain environment, the hierarchical path computation structure is beneficial for collecting and sharing the node information for path computation. Searching a perimeter quickly between the exit point of a domain and the adjacent domain is important. Therefore the architecture of ICE that controls the information of each domain, and the top-level path computation node that controls the ICE are necessarily.

Given that this paper proposed the architecture and the protocol of the Information Control Element that controls the information of each domain, and the Global Federation Organizer that controls the ICE. Since a network node does not have multi-domain processing or routing capabilities, we proposed the ICE and the Global Federation Organizer to perform optimal path computation. The proposed system can overcome the limited routing capability and manage the various information in multi-domain. Based on the information, the ICE of each domain computes available paths. The Global Federation Organizer then combines the result provided by the ICE and computes an optimal path. This process can reduce the processing overhead of a CPU by distributing excessive processes.

We proposed a method for effective use of resources, the optimal path selection among multi-domains, and the optimization of the end-to-end connection by designing the architecture of the ICE and the Global Federation Organizer considered a multi-domain environment, and applying the architecture to a multi-domain network. Additionally, we performed two case-studies, the 'large data transferring' and the 'real-time media transferring' by configuring the test-bed for the logical multi-domain environment. We proposed the optimal path computation model that uses resources effectively, and requires limited information between nodes for multi-domain information.

References

- [1] Bonerjee, A., Drake, J., Lang, J.P., Turner, B.: Generalized Multiprotocol Label Switching: an Overview of Routing and Management Enhancements. *IEEE Commun. Magazine* 39(1), 144–150 (2001)
- [2] Shaikh, A., Shin, K.: Destination-driven Routing for Low-cost Multicast. *IEEE Journal on Selected Areas in Communications* 15(3) (April 1997)
- [3] Cha, M., et al.: Path Protection Routing with SRLG Constraints to Support IPTV in WDM Mesh Networks. In: 25th IEEE International Conference on Computer Communications, INFOCOM 2006, pp. 1–5 (2006)
- [4] Vasseur, J., Ayyangar, A., Zhang, R.: A per-domain path computation method for establishing inter-domain traffic engineering (TE) label switched paths (LSPs). draft-ietf-ccamp-inter-domain-pd-path-comp-05 (work in progress) (2007)
- [5] Vasseur, J.P., et al.: A backward-recursive PCE-based computation (BRPC) procedure to compute shortest constrained inter-domain traffic engineering label switched paths. RFC 5441 (April)
- [6] López, V., et al.: Path computation element in telecom networks: Recent developments and standardization activities. In: Proc. of the 14th Conference on Optical Network Design and Modeling (ONDM), pp. 1–6 (2010)
- [7] Oki, E., Inoue, I., Shiimoto, K.: Path Computation Element (PCE)-based Traffic Engineering in MPLS and GMPLS networks. *IEEE Communications Magazine* (2008)
- [8] Ayyangar, A., Kompella, K., Vasseur, J.-P., Farrel, A.: Label Switched Path Stitching with Generalized MPLS Traffic Engineering. Internet draft (April 2007)
- [9] Seok, Y., Lee, Y., Choi, Y., Kim, C.: Explicit Multicast Routing Algorithms for Constrained Traffic Engineering. In: Proceedings of the 7th ISCC, pp. 455–461 (July 2002)
- [10] Aslam, F., Uzmi, Z.A., Farrel, A.: Interdomain Path Computation: Challenges and Solutions for Label Switched Networks. *IEEE Communication Magazine* (October 2007)

Robot Reinforcement Learning for Automatically Avoiding a Dynamic Obstacle in a Virtual Environment

Phuong Chu, Hoang Vu, Donghyeon Yeo, Byeonggwon Lee,
Kyhyun Um, and Kyungeun Cho*

Dept. of Multimedia Engineering, Graduate Schools of Dongguk University,
26, Pil-dong 3-ga, Jung-gu Seoul, 100-715, Republic of Korea
cke@dongguk.edu

Abstract. In a virtual environment, a robot can serve people by bringing things to them. However, when a robot moves within a house, it collides with a dynamic obstacle. These collisions make it difficult for a robot to complete its mission. We therefore apply reinforcement learning to the robot to make it more intelligent. Consequently, the robot can automatically move to avoid the dynamic obstacle in order to successfully complete its mission.

Keywords: Reinforcement learning, virtual environment, automatic, dynamic obstacle, virtual robot.

1 Introduction

Nowadays, humanoid robots can help people execute many tasks. However, robots have limited physical capabilities and ‘intelligence’; consequently, they are unable to perform many desirable tasks. Robots should therefore be improved in both physical and artificial intelligence aspects. In this study, we focus on how to make a robot more intelligent. We employ a humanoid robot (‘Nao’ by Aldebaran), a real house with furniture, and a program to control the robot’s behavior, such as moving, rotating, grasping an object, and releasing it. The robot is expected to move within the house and fetch the object for a human. When it moves, it should avoid colliding with the dynamic obstacle that randomly move. However, the robot does not know how to avoid the obstacles when navigating to the object and bringing it to the human.

If we use the A* [1] algorithm to find the path, the robot cannot avoid the obstacle because it may be situated in the shortest path chosen by the robot. Our approach to solving this problem is reinforcement learning. Nevertheless, another problem occurs. If we use a reinforcement learning algorithm, the robot must move innumerable times in as many periods; furthermore, the learning time is likewise very long. A real robot such as Nao cannot implement this immense task. We therefore must create a virtual house that emulates the real house.

Within our virtual house, a virtual human can stand in any position. A virtual object, such as a ball, is situated in the house; it can be placed in any position. In addition,

* Corresponding author.

a moveable object—a virtual dynamic obstacle—can randomly move around the house. The virtual robot should navigate to the object, grasp it, and bring it to the human. When the virtual robot moves, it can collide with the moveable object. It will therefore not successfully complete the task. To address this issue, we apply reinforcement learning [4-8] to the virtual robot so that it can automatically move to avoid the dynamic obstacle. We then apply the result of this virtual experiment to the real humanoid robot.

2 Related Work

In controlling a robot, path finding is very important. Many path finding studies exist [1-3]. We could employ the A* or Dijkstra algorithm [1], which focus on the shortest path. However, a problem occurs if either of these algorithms is used for mapping with dynamic obstacles. In many cases, the obstacle remains in the shortest path; therefore, the robot cannot move to its target without colliding with the obstacle.

The proposed algorithm was inspired by the work of [2,3]; however, rather than use a genetic algorithm, we provide a new approach. The reinforcement learning algorithm is very useful for learning human-robot interactions [4]. This algorithm can be used in a virtual environment for increasing the learning rate and saving time. The result is then applied in a real environment [4]. In this study, we combine reinforcement learning with path planning. We divide the robot movements into two parts. In the first part, the robot mission is to navigate to a ball and to grasp it. In the second part, the robot is to navigate to the human and give the ball to him/her.

3 Virtual Learning Approach

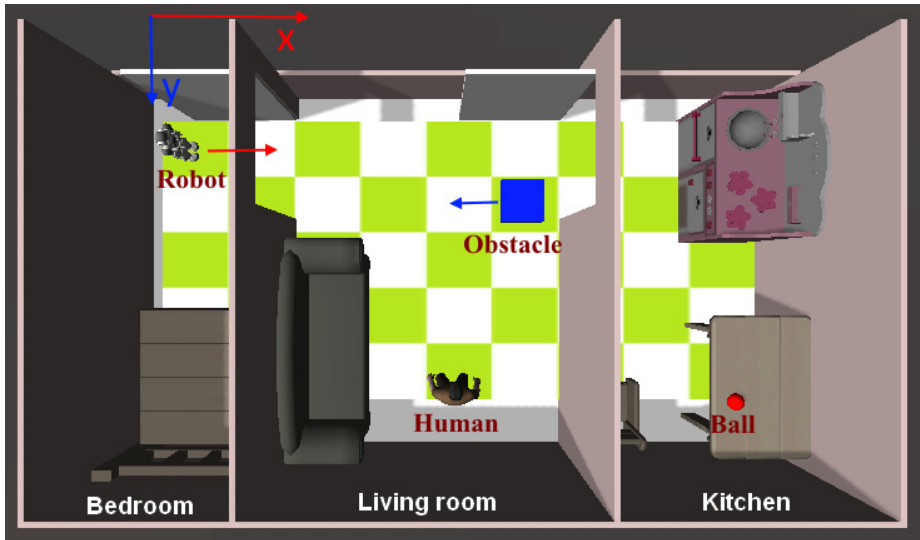
3.1 Virtual Environment Configuration

Our virtual environment was comprised of a virtual house with three rooms—a bedroom, living room, and kitchen—as shown in Figure 1. Each room contained various furniture objects, such as a bed, sofa, stove, chair, and table, which were all static. We used a 9 x 5 grid to calculate the respective positions of the robot, human, ball, and obstacle. We used a two-dimensional coordinate. We converted the positions of all objects from three-dimensional coordinates to two-dimensional ones. All positions of the same cell were converted to only one position in the two-dimensional coordinate. Therefore, when two objects remained in one cell, we recognized that they collided with each other.

All of the static objects had fixed positions in the house; therefore, some cells in the grid were permanently occupied. Not all moveable objects could move in these cells. The positions of all objects in this situation are shown in Table 1. In each period, the respective positions of the human and ball were static; however, they could be changed in a different period. The dynamic obstacle could randomly move within the house. Each time, it moved to the nearest empty cell in one of four directions: up, down, left, or right. The human stood in any position and waited for the robot to bring the object. The robot could move in any of the same four directions as the dynamic obstacle.

Table 1. Object positions within the virtual house

Object	Position
Robot	(0,0)
Human	(4,4)
Obstacle	(5,1)
Ball	(8,4)
Bed	(0,3), (0,4), (1,3), (1,4)
Sofa	(2,2), (2,3), (2,4)
Stove	(8,0), (8,1)
Table	(8,3), (8,4)
Chair	(6,4)


Fig. 1. Virtual house

3.2 Application of Reinforcement Learning to the Virtual Robot

Reinforcement learning was used to find an optimal action-selection policy for the robot. The reinforcement learning algorithm includes a function that calculates the quality, Q , of a state-action combination:

$$Q: S \times A \rightarrow R \quad (1)$$

where S is a set of states, A is a set of actions, and R is the result.

We used a Q table for saving the priorities of all states of the robot's actions (2). Each value in the table depended on the positions of the robot, human, obstacle, and ball, as well as on the moving direction of the robot. We initialized all values of the Q table as zero.

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha \times (r_{t+1} + \gamma \times \max Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)) \quad (2)$$

where s_t, a_t are the previous state and previous moving action, respectively, and s_{t+1}, a_{t+1} are the next state and next moving action, respectively. In addition, r_{t+1} is the reward observed after the robot implements moving action a_t in state s_t , $\max Q_t(s_{t+1}, a_{t+1})$ is an estimate of optimal future value, α is the learning rate, and γ is the discount factor.



Fig. 2. Robot grasps the ball

After any given step, the value in the Q table was updated. In any situation, the robot chose the moving direction with the highest priority in the Q table. If all values were equal, the robot randomly chose one of the four directions. If the robot moved to a position near the ball, the robot grasped the ball, as shown in Figure 2. The robot then continued moving.

If the robot and human were in the same position, and the robot was holding the ball, the robot gave the ball to the human, as shown in Figure 3. In this case, the robot received a positive reward. In the second scenario, the robot met the human but did not bring anything to him/her. This incurred a bad result and a negative reward for the robot. In the third scenario, the robot received a negative reward if it collided with the dynamic obstacle, as shown in Figure 4. The period was completed in all three cases. The next period randomly began with another situation.

When each period ended, the robot received a new experience. In the same situation, if the robot received a negative reward during the last period, it chose a different path in the subsequent period. This learning was repeated many times to provide the robot with many experiences. When the robot obtained a sufficiently large number of experiences, in each period, it chose the direction that would most effectively help it receive a positive reward for each situation. After a few million periods, the robot could automatically move to complete the mission and avoid the dynamic obstacle.

When it was evident that the virtual robot could intelligently move, we exported the Q table data into a database. We then imported this database into a module that we used to control the real robot. We situated a camera on the ceiling of the real house to obtain a map. In the real robot control module, we recognized the positions of all real objects and converted them into two-dimensional coordinates, as was done for the virtual environment. In each situation, we read the values from the Q table and provided the real robot with the optimal choice of the direction in which to move; i.e., up, down, left, or right.



Fig. 3. Robot gives the ball to the human

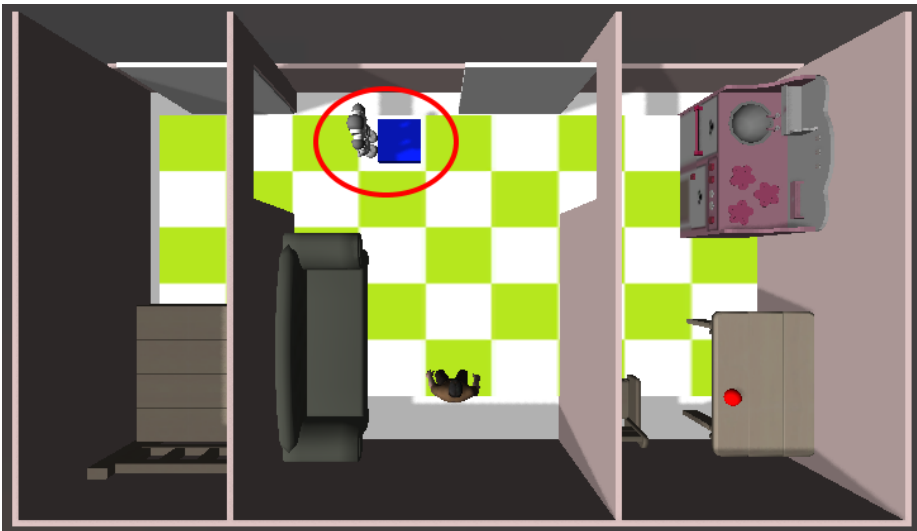


Fig. 4. Robot collides with the dynamic obstacle

4 Experiment

We conducted an experiment to evaluate the proposed approach. We counted the number of collisions that occurred between the virtual robot and dynamic obstacle in 10,000 respective periods, as shown in Figure 5. From the first to the ten-thousandth period, the number of collisions between the virtual robot and obstacle was very high. The number of periods had collision took 85% of the total of all periods. However, it very quickly decreased within the next one million periods. By approximately the one-millionth period, the number of periods that included a collision was only 30%. From that point, the percentage gradually decreased from 30% to less than 20%.

We additionally counted the number of successful missions of the virtual robot. A period was deemed successful if the virtual robot navigated to the ball's position, grasped it, and brought it to the human without colliding with the obstacle. A period was regarded as a failure in two cases: if the robot collided with the obstacle; and if the robot met the human but did not bring anything to him/her.

In our experiment, the second type of mission failure rarely occurred because the dynamic obstacle was always moving and the robot usually collided with it. In addition, after learning from a few thousand periods, the robot 'knew' that it should not meet the human without bringing anything to him/her. The numbers of successful periods are shown in Figure 6. When the virtual robot began learning, only 15% of the periods were successful. After a few million periods, the percentage of success was greater than 80%.

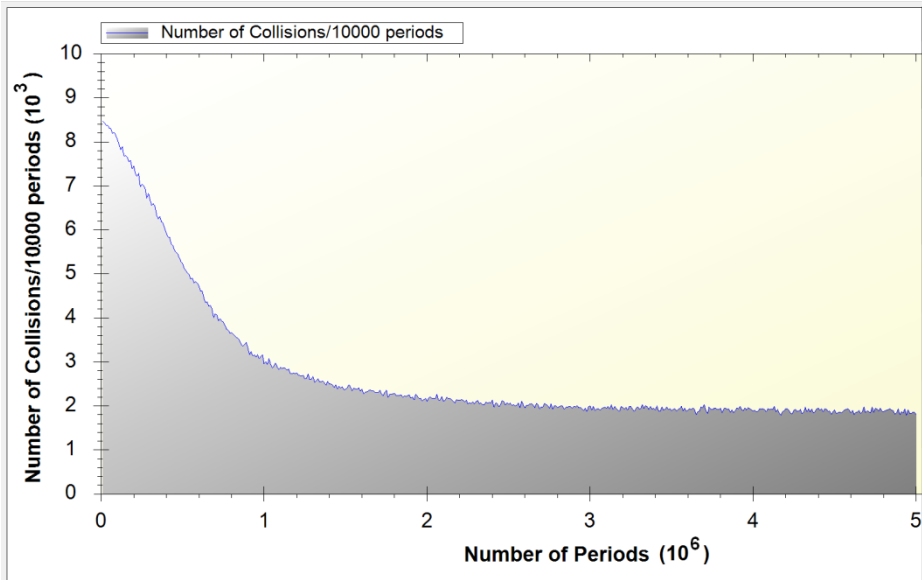


Fig. 5. Number of collisions between the virtual robot and dynamic obstacle per 10,000 periods

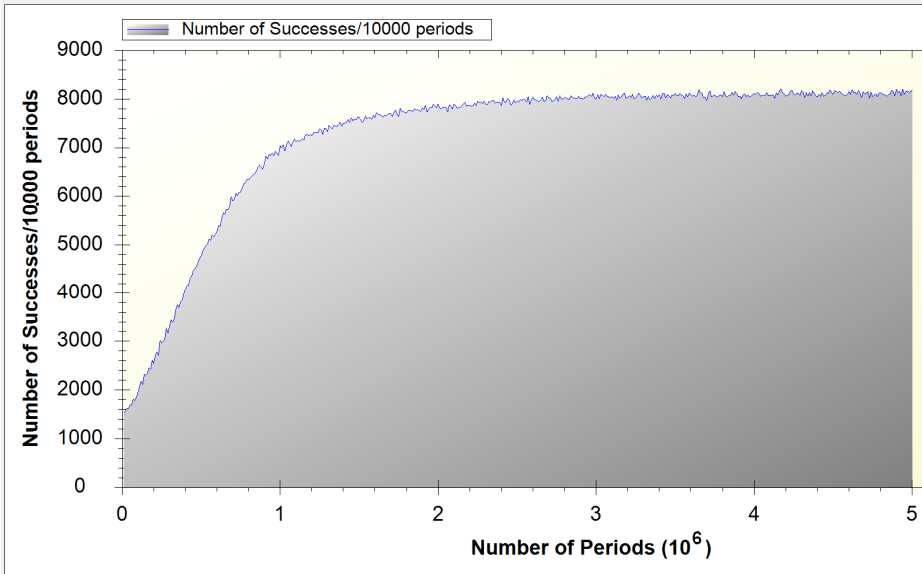


Fig. 6. Number of successes of the virtual robot per 10,000 periods

5 Conclusion

In this paper, we presented a novel approach to making a robot more intelligent. A robot is typically intended to provide some household function, such as fetching objects for a human. An effective robot must automatically avoid dynamic obstacles while moving. To achieve this objective, reinforcement learning cannot be directly applied to a real robot because it would be immensely time-intensive and the robot's physical ability is limited.

To address the above issues, we applied reinforcement learning to a robot in a virtual environment, which increased the learning speed. The virtual robot became more intelligent after each period. At the start of learning, the virtual robot could only randomly move because it had no experience. After learning through a few million periods, the virtual robot selected in each step the optimal path for avoiding the dynamic obstacle. When we change the grid for calculating the positions of all objects more finely. To this end, however, a larger number of learning periods would be required, and the learning time would considerably increase. Nevertheless, at times, the path chosen by the virtual robot was not the shortest one. We saved the data from the Q table into a database for application to a real environment. In future research, we intend to improve the learning to increase the percentage of successful periods and use more than one obstacle.

Acknowledgments. This work was also supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2009148).

References

- [1] Zhang, Z., Zhao, Z.: A Multiple Mobile Robots Path planning Algorithm Based on A-star and Dijkstra Algorithm. *International Journal of Smart Home* (2014)
- [2] Zou, X., Ge, B., Sun, P.: Improved Genetic Algorithm for Dynamic Path Planning. *International Journal of Information and Computer Science* (2012)
- [3] Achour, N., Chaalal, M.: Mobile Robots Path Planning using Genetic Algorithms. In: *ICAS 2011: The Seventh International Conference on Autonomic and Autonomous Systems* (2011)
- [4] Sung, Y., Cho, S., Um, K., Jeong, Y., Fong, S., Cho, K.: Human-Robot Interaction Learning using Demonstration-based Learning and Q-learning in a Pervasive Sensing Environment. *International Journal of Distributed Sensor Networks* (2014)
- [5] Sung, Y., Ahn, E., Cho, K.: Q-learning Reward Propagation Method for Reducing the Transmission Power of Sensor Nodes in Wireless Sensor Networks. *Wireless Personal Communications* (2013)
- [6] Even-Dar, E., Mansour, Y.: Learning Rates for Q-learning. *Journal of Machine Learning Research* 5 (2003)
- [7] Smart, W.D., Kaelbling, L.P.: Practical reinforcement learning in continuous spaces. In: *Proceedings of ICML 2000*, pp. 903–910 (2000)
- [8] Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research* 4 (1996)

Smart Virtual Lab Using Hand Gestures

Warda Ikram, Yoonji Jeong, Byeonggwon Lee, Kyhyun Um, and Kyungeun Cho*

Dept. of Multimedia Engineering, Graduate Schools of Dongguk University,
26, Pil-dong 3-ga, Jung-gu Seoul, 100-715, Republic of Korea
cke@dongguk.edu

Abstract. This study discusses the potential of motion sensors combined with natural user interfaces for an education system. In this research, we included a real-time virtual lab for pre-lab sessions for students. A pre-lab session conserves resources and allows students to prepare for the actual lab session, helping to avoid certain incidents. It can also be an effective set-up for practicing experiments they have already been performed. In this paper, we discuss the development of a virtual lab environment and gestures to interact with the lab. We designed a prototype chemistry lab with 3D models and effective gestures to interact with the virtual lab.

Keywords: Pre-lab, NUI, Effective Gestures, Virtual Lab.

1 Introduction

In conventional science labs, it takes a long time to set up or perform an experiment. Sometimes chemicals are so reactive and strong that students need a high level of supervision to perform those experiments. Along with the experiments, there are other issues of danger and lack of resources. In this study, we attempted to simulate a virtual lab using 3D tool. We chose an NUI (natural user interface) because we wanted the simulation to be simple and natural.

A user can perform a vast variety of hand gestures. We studied all the categories from previous studies and extracted the important points; there included various studies about hand gestures and human computer interaction using them [1], [2]. We gathered natural hand movements by mixing some sub-divisions of gestures from previous work. For hand gestures to work with a 3D environment we wanted them to be natural and effective for the application. There are many applications on the table that are using hand gestures, such as touch screens and modules and also gestures that are serving the robot industry [3], [4].

The gestures were selected after investigating previous studies. To use those gestures with an application, the main idea was to come up with a virtual lab. We selected various lab equipment like 3ds Max-based models to include in a 3D virtual environment. We made a virtual room using the Unity3D tool, and then we incorporated our

* Corresponding author.

models into that world space and used gestures to manipulate them. To recognize hand gestures we used Leap motion, a motion-sensing device that works very precisely and accurately. We used Leap motion for our prototype because it is accurate for hand and finger movements and it is also very cost effective.

2 Related Work

We wanted to introduce our work as an improved educational application. We searched for a better approach after reviewing previous work. The simplest way for a human being to interact with a machine is the way that they can most easily understand. The only possible way was to be natural. Studies have shown that visual information and pictorial data helps more in learning than ordinary theory only methods. Pictures and models make more sense to young minds than just words and text [5]. In the education system we don't usually find learning applications using actual hand movements and manipulating object across a virtual room. Even if there are advanced methods available, educational authorities hesitate to use them or include them in their curriculum. NUI is simple and reliable.

There are systems that have tried using models for mathematics and physics teaching [6], but models used only by teachers do not provide enough benefits to the education sector. There are applications that became very useful for children with impairments and other mental inabilities such as autism [7], [8], and outside the health sector teachers tried using material like 2D and 3D models to show students the deep meaning behind theoretical concepts. This sort of study method can enhance teaching quality [9]. These natural user interfaces and environments have included not just one motion sensor but also other devices, for instance, a wearable device such as a, 3D virtual reality device for better 3D world concepts [10]. Devices like these can be used for actual virtual world space. Game applications and entertainment systems have always been a part of this type of research but stating that they are whole educational systems is not appropriate. For an application to be specifically education-based and beneficial as well, it should be a perfect balance of fun and education.

There are many studies of entertainment packages that can be used actively and are good for physical purposes but are not that useful for education purposes [11], [12]. To include that essence of education, a proper approach for developing a learning application is to add equal proportions of both to keep the balance. Simple applications that just use some gestures for applications and focus mainly on gestures [13] are not enough; they cannot fulfill the requirements of an education-based application. Also, if the system is highly education-based it reduces the interest factor and, keeping in mind that we are dealing with young students, it will not be a better solution. There are other serious applications available for adult students [14].

Digital media have been very prominent in educational level applications because the potential of learning and extracting information in a proper fashion is high if

digital media is used [15]. All these studies lack a balanced proportion of education and entertainment features. Our approach is to use both in parallel to achieve our required result of the virtual lab application.

3 Application Hierarchy

Our focus of the application was to deliver an application that could serve the education sector as a pre-lab session for an actual lab experiment. We created an interface for a virtual chemistry lab and included 3D models. Students can grab these models like they can in daily life and manipulate them according to the given instructions. Moreover, they can perform experiments with as many resources as they want and without any limitations. That means that the system is not only cost-effective but also harm free. Introduction of a simple interface was always a priority due to the young audience. Practice with this virtual lab can not only boost their confidence but also enhance their learning experience. The application hierarchy is given below.

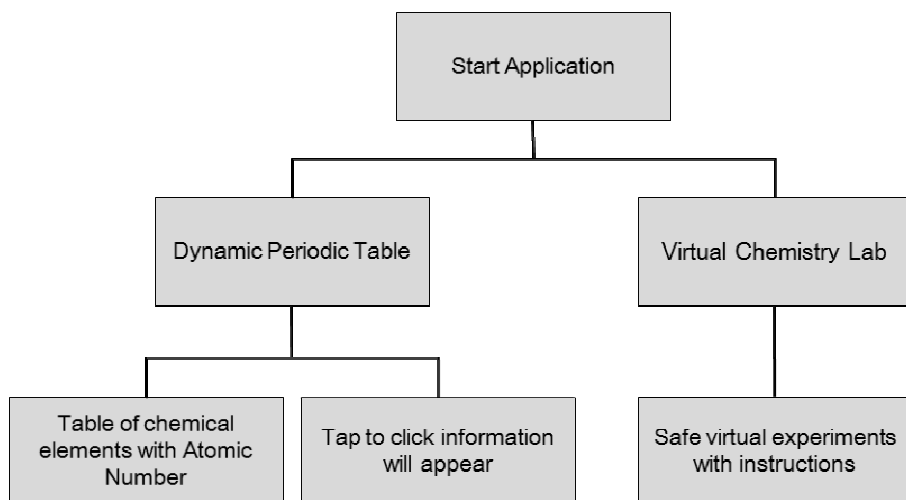


Fig. 1. Application flow diagram

We also set up some gestures to work with this application, like grabbing an object and moving it in the scene, rotating a model of a flask to pour liquid into a dish, etc. The gestures synched very well with the application. Most of them were used in the lab application. These gestures made the whole environment look more real.

4 Efficiency of Gestures

The use of gestures was very important in our application, but coping with gestures in a 3D space was difficult because of the world space co-ordinates. As mentioned earlier,

we used a mix of gesture categories. Given below is the pseudo code example of our gesture, the gesture indicating a wrong option or cancellation of selection of an object.

```

Cross_Gesture_Recognition(leaphand)

checkHandCount<- GetLeapHand(leaphand)

IF checkHandCount = FALSE THEN RETURN FALSE
checkFingerCount <- CheckFinger(leaphand)

IF checkFingerCount = FALSE THEN RETURN FALSE
handDistance <- GetDistance(leaphand)
fingerDistance <- GetDistance(leaphand, leapfinger)
CrossGesture<- CheckCrossDistance(handDistance, fingerDistance)
RETURN CrossGesture
END

```

This pseudo code for a cross gesture is a wrong indicator. function checkHandCount checks user's hand whether more than one or not. And then function checkFingerCount checks the finger that we need. If the wrong object has been selected it can be cancelled using this gesture. The gesture worked very accurately and was tested many times within the application.



Fig. 2. Using a gesture to perform an experiment



Fig. 3. The reaction of an experiment as shown in Unity3D

5 Analysis and Conclusions

The virtual environment for a pre-lab session was developed to provide the features that a real environment lacks. A real environment needs plenty of lab apparatus to fulfill the needs of experiments, but a virtual lab needs only 3D models, and the supply of models is unlimited the breakage issues in a real lab can be prevented and children can warm up with the pre-lab session before their actual experiments.

Students can consume as much time in the virtual lab as they want. A real lab provides only a limited time to perform an experiment, and the setup takes time, so a student cannot repeat the experiment many times.

Supplies of expensive chemicals are limited for the real lab so here, again, the virtual lab can serve better.

The virtual lab is portable and can be used anywhere, whereas a real lab needs a proper place; with heavy lab equipment you cannot practice or perform experiments outside the lab.

Of course, a virtual lab cannot replace an actual lab, but it can provide its services as a pre-lab session for young students. This pre-lab session can build confidence in young students and can be used as an additional piece of entertainment in the curriculum for better results.

We used this application with a number of students and asked them to give us their feed-back. The provided feed-back indicated that they were satisfied and they liked the gestures and the Leap motion device. New experiences like these can enhance the learning experiences of students.

Acknowledgments. This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (NIPA-2014-H0301-14-1021) supervised by the NIPA(National IT Industry Promotion Agency).

References

1. Karam, M., Schraefel, M.C.: A taxonomy of gestures in human computer interactions. *ACM Transactions on Computer-Human Interactions* (2005)
2. Varga, E., Verlinden, J., Klass, O., Langenhoff, L., Van Der Steen, D., Verhagen, J.: A Study on Intuitive Gestures to Control Multimedia Applications. In: *Iadis International Conference Interfaces and Human Computer Interaction* (2008)
3. Kim, E.C.H., Chung, M.K.: A taxonomy and notation Method for three-dimensional hand gestures. Elsevier (January 2014)
4. Bay, S., Brauner, P., Gossler, T., Ziefle, M.: Intuitive Gestures on Multi-touch Displays for Reading Radiological Images. In: Yamamoto, S. (ed.) *HCI 2013, Part II. LNCS*, vol. 8017, pp. 22–31. Springer, Heidelberg (2013)
5. Gleeson, B., MacLean, K., Haddadi, A., Crof, E., Alcazar, J.: Gestures for Industry: Intuitive Human-Robot Communication from Human Observation. In: *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, HRI 2013*, pp. 349–356 (2013)
6. Park, W., Han, S.H.: Intuitive Multi-Touch Gestures for Mobile Web Browsers. *Interact. Comput.* (2013)
7. Echeverría, M.A.M., Santana-Mancilla, P.C., Carrillo, H.F.Q., Enciso, E.A.F.: Natural User Interfaces to Teach Math on Higher Education. In: *4th International Conference on New Horizons in Education*, vol. 106, pp. 1883–1889 (2013)
8. Ringland, K.E., Escobedo, L., Zalapa, R., Tentori, M., Neal, M., Hayes, G.R.: Sensory-Paint: A Natural User Interface Supporting Sensory Integration in Children with Neurodevelopmental Disorders. In: *CHI* (2014)
9. Sampath, H., Agarwal, R., Indurkha, B.: Assistive Technology for Children with Autism – Lessons for Interaction Design. In: *APCHI*, pp. 325–333 (2013)
10. Chao, C.-H., Chen, Y.-C., Yang, T.-J., Yu, P.-L.: Intelligent Classroom with Motion Sensor and 3D Vision for Virtual Reality e-Learning. In: *The 2nd International Workshop on Learning Technology for Education in Cloud. Springer Proceedings in Complexity*, pp. 27–33 (2014)
11. Detlefsen, J.: The Cosmic Perspective: Teaching Middle-School Children Astronomy Using Ego-Centric Virtual Reality (May 2014)
12. Jung, H.Y., Kim, H.: Interface Design of Game Content for Children Using Motion Cognition Technology. *Advanced Science and Technology (Games and Graphics)* 54, 69–72 (2014)
13. Wigdor, D., Wixon, D.: *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Elsevier, USA (2011)
14. Girbacia, F., Butnariu, S.: Development of a Natural User Interface for Intuitive Presentations in Educational Process. In: *Proceedings of the E Learning and Software for Education*, pp. 74–79 (April 2012)
15. Yannier, N., Koedinger, K.R., Hudson, S.E.: Tangible Collaborative Learning with a Mixed-Reality Game: EarthShake. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) *AIED 2013. LNCS*, vol. 7926, pp. 131–140. Springer, Heidelberg (2013)

Internal Topology Based Flexible Shortest Path Planning Method for Indoor Navigation

Yan Li and Byeong-Seok Shin

Department of Computer Information and Engineering, Inha University, Korea
leeyeon622@gmail.com, bsshin@inha.ac.kr

Abstract. We proposed a navigation method based on topology analysis for a dynamically changing indoor navigation path. To support the dynamically changing characteristic, we pre-construct indoor route when the route is required. And we dynamically update its internal path information when it changes. The proposed method is suitable for both simple and complex large-scale indoor spaces, even when related indoor maps are difficult to be navigated. We conduct a performance evaluation to compare the proposed method with current research approaches. The results show that our method provides improved performance for indoor navigation.

Keywords: Topological analyzer, indoor navigation, shortest path.

1 Introduction

A location-based service(LBS) system[11] is typically combined with digital maps to provide user location data and related routing information. It is widely used in many fields, such as routing navigation[1], social networking[2], public safety, and mobile commerce. With the continuous development of digital cities and gradual popularization of mobile terminals, the demand is increasing for location services in large complex indoor environments[12] such as shopping centers, car parks, air ports and museums. Although the scales of outdoor spaces are typically larger than those of indoor spaces, outdoor structures are usually less complex than internal structures, which tend to have more elements, such as doors, corridors, rooms, and partitions. These elements are often not a part of outdoor spaces. At times, even a person who is familiar with a complicated indoor space may have difficulty for navigating it; moreover, if the person is unfamiliar with the space, it can be very easy to become lost. Therefore, the demand for effective directional guidance [2] in complex indoor environments is becoming ever more apparent. This increasing demand is resulting in the rapid development of indoor routing services. Currently, many LBS systems are primarily designed for routing in outdoor, not indoor, spaces. The development of a fast, high-performance, continuously updating indoor navigation system is needed.

Shortest path problem is the main purpose of path planning for road networks. Some renowned algorithms exist, such as the Dijkstra, A*, elastic [2], iNav [3,4], and context-aware [5] algorithms. Those models consider only the topological relationships of indoor geometries [6], internal path networks of the geometries, or link-node relationships [7,8] to obtain the navigation route.

The topology analysis is first proposed in Elastic algorithm. Based on indoor space object topology relations, this algorithm obtains a navigation route according to the coordinate information of spatial geometries. This algorithm uses only spatial geometries; however, it does not check if there is an internal road in the spatial geometry. Therefore, if the internal geometry has a common road, a better navigation route can be obtained by using it.

The iNav and context-aware algorithms can automatically calculate the shortest paths according to the geometric structure of the indoor space. They construct a means of determining an indoor network that is based on the geometry of the indoor space and supports length-dependent optimal routing. In [9], the model is extended by decomposing concave-shaped objects to ensure that users can view the next hop indicated in each directive during the navigation. However, the routes generated by this model are not the shortest ones.

When a significant amount of geometry information exists for a complex indoor space, such as a large shopping mall, many channels exist in the inner part of the geometry or among the geometries.

To solve the above problems, we propose a flexible path-planning algorithm that combines indoor route information with the given internal information, such as rooms, internal room path networks, corridors, and so on, within the indoor space. We propose a model to support length-dependent optimal routing based on the geometry of indoor structures. This model fully employs the relationships of rooms, internal room path networks, and lobbies. It applies spatial topology analysis [10] and internal room path networks to produce a navigation path according to the user's location and destination. By using the proposed method, the shortest route represents a finite sequence of path segments obtained by applying the shortest path algorithm in an indoor space. Detailed routing information, such as turns and the route length, can be obtained from the path segment.

The left of this paper is organized as follows: In chapter 2, an indoor network is built and discussed about its benefits of navigation. In chapter 3, an experiment of the proposed method is shown. Finally, we conclude our work.

2 FPP-Algorithm Based-on Internal Path Networks of a Room

We call our flexible path planning method as FPP-algorithm. When some rooms have similar shapes, they may serve different internal room path networks. Furthermore, even though some of them are totally different in shapes, they might provide the same function during routing selection. We therefore consider this room as one in which people can only enter but not pass through. When a room has two or more doors, more than one path through the room exists. We can use those paths to obtain an optimal route. A room is one of three base unit types; i.e., a room with one door, multiple doors, or multiple open doors.

Generally, a route includes many path segments situated between the start and end locations. Indoor space typically does not have a fixed path segment; nevertheless, users tend to choose a straight line as the path segment to their destination. Accordingly, we analyze the shortest path segment between any pair of access points

in different cells. In an indoor space, if a room has only one access point, it will not have a passage. Therefore, we must consider only the path segments that exist in a multiple-door room unit.

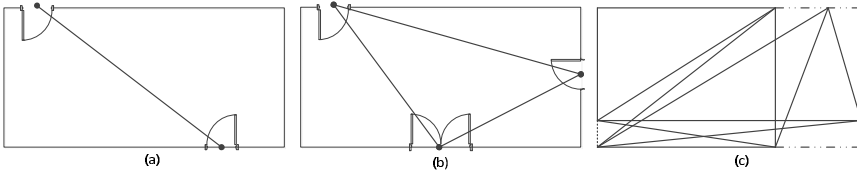


Fig. 1. (a)Path segment in multiple doors unit with two access points, (b)path segments in a unit with multiple access points,(c)path segments in open door unit with multiple access points

For a multiple-door room unit, the approach to determining implicit path segments is based on the unit shapes and locations of access points. The simplest case is a unit with only two access points that can be reached through a straight line. The shortest path from one access point to the other is a straight line between them (Fig. 1(a)). If a unit has more than two access points that can be directly reached from each other, we obtain multiple implicit path segments in the cell. The shortest path segments in this unit are straight lines connecting any two access points. For example, in Fig. 1(b), we find three implicit path segments in a cell with three access points. As shown in Fig 1(c), there are three open doors; we find eleven implicit shortest path segments in this unit with six access points.

Indoor graphic data include the room spatial data, MBR, and ‘DoorPoints.’ The MBR is an expression of the maximum extents of a 2D object (e.g.a-room) within its 2D coordinate system; we denote the dotted line rectangle as min(x), max(x), min(y), max(y). We set the DoorPoints as being comprised of two types of structures. When the DoorPoints unit has no opened door, we set the first data item as false. If the door is an opened one, we change the first data item to true; the next data item includes two points (the open door’s start and end points). We store all of the data in clockwise order.

In complex indoor environments, internal rooms exist within a given indoor room. We can use these internal room channels to obtain the shortest path. For example, several rooms exist on the same level in the room hierarchy.

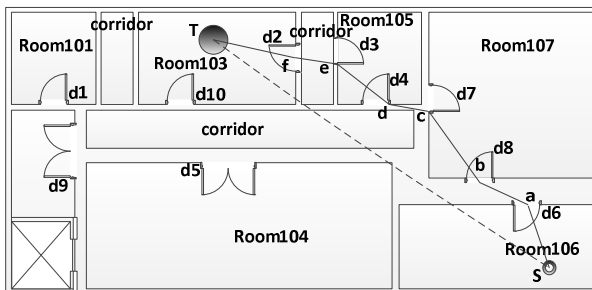


Fig. 2. Indoor space navigation by using FPP-algorithm

According to the design concepts of the proposed algorithm, we suggest the following steps for obtaining an indoor navigation path.

Algorithm 1. shortest path planning algorithm for FPP-algorithm

<p>Input: StartPoint: The user's location point or user selects point. TargetPoint: The user selects the target point.</p> <p>Output: pPaths: The navigation Path points</p> <hr/> <p>Algorithm FPP-Indoornavi(StartPoint,TargetPoint)</p> <p>Begin</p> 01: pA = StartPoint; 02: pB = TargetPoint; 03: pPaths.Add(pStartPoint);//get pA the nearest room 04: pIndoor_Room = getNearestRoom(pA,pB); 05: if (pIndoor_Room.GID != NULL)//get the room after cutting the //recent all point information 06: pPnts= LineInRPnts(pA,pB,pIndoor_Room); //get the small part of the room points and return the path Points 07: RMSMIPartPnts = RoomSmallPart(pPnts); 08: pPaths.Add(RMSMallPartPnts); //recursice FT-Indoornavi and get all the navigation path. 09: FT-Indoornavi(pA, pB); 10: endif 11: pPaths.Add(pB); 12: return pPaths END

A navigation algorithm that considers only a single floor is insufficient because a real-world complex indoor space typically has multiple floors. Navigation between floors is very important. We consider the specificity of building floors as special points—i.e., nodes that link different floors—such as the points of elevators, stairs, and escalators, which link one floor to another. In addition, we set the distance of an object shuttling between floors, such as an elevator, as a fixed value; the distance of other objects, such as stairs or escalators, are likewise a fixed value. Therefore, we can use these fixed-distance values to calculate the optimal path for navigation.

3 Experimental Results

We describe the flash memory and buffer cache used to conduct our experiments. We then compare our proposed method to others that do not refer to buffer contents during garbage collection.

Our FPP algorithm was implemented on the Eclipse platform, Java programming language, MySQL database, and Java database connectivity (JDBC). The indoor map data source was from WanDa Plaza, Chongqing, China which was created in our project. The data included a map of four floors and contains 150 rooms. There were approximately 37 records for each floor (floor room information). The experimental system had an Intel(R) Core 2-GHz CPU, 2 GB of RAM, and 500-G hard disk.

We compare the FPP-algorithm and other link-node algorithms.

Update: The link-node network updates the geometry (such as rooms, corridors, lifts, lobbies, and stairways) and link-node network. However, FPP algorithm updates only the geometry.

Fault Tolerance: If the network link-node connection relations between the link and node have errors, the link-node algorithm cannot be run, and the system will shut down. On the other hand, if any errors exist for indoor space geometries, the FPP-algorithm may obtain a path that may not be the nearest path, but it is nevertheless a navigation path.

In our experiment, the FPP-algorithm used topology analysis and room internal paths in the indoor space. The elastic algorithm used topology analysis but not internal room paths. The iNav and context-aware algorithm used internal room paths but not topology analysis. The FPP-algorithm model did not involve the link-node relationship; rather, it combined the relationships among geometries with the inner channel of geometry and therefore quickly obtained the navigation route. It provided improvements in two aspects: updating data and fault tolerance. This model greatly increased the data utilization rate, which included that of the indoor room, door, and corridor; it did not use the node-link data. Based on these results, the proposed algorithm is more effective than the others.

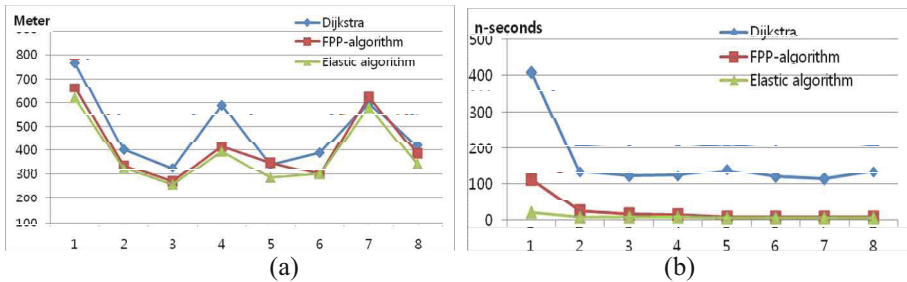


Fig. 3. The evaluation result of (a)path distance and (b)path computation time

There are eight groups of data in Fig. 3(a) and (b); each group represents the navigation path distance acquired by the three algorithms with the same start and end points. Although FPP-algorithm is slightly more than the Elastic algorithm, in terms of computing time it can provide users with shorter than Dijkstra and Elastic shortest path selection, which has application value. It is evident that the distance of acquired navigation path using the proposed algorithm is shorter than the one using Dijkstra and Elastic.

4 Conclusions

We proposed an indoor topology based indoor navigation method called FPP-algorithm. The proposed method effectively leverages indoor facilities, such as

rooms, corridors, doors and so on, in complex and indoor spaces. This algorithm does not depend on traditional road network structure. Even for a complex internal room our method can provide the shortest path in a short time, because it pre-computes the internal path and constructs the road network. In future work, we intend to make this algorithm more widely applicable to actual circumstances, including to indoor and outdoor connected environments.

Acknowledgments. This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HI14C0765).

References

1. Guha, R.K., Chen, W.: A distributed traffic navigation system using vehicular communication. In: Proc. of 1st. IEEE Vehicular Network Conference (2009)
2. Chen, Y., Wu, K.-Q., Wang, D., Chen, G.: Elastic Algorithm: A New Path Planning Algorithm About Auto-navigation in 3D Virtual Scene. In: Pan, Z., Cheok, D.A.D., Haller, M., Lau, R., Saito, H., Liang, R. (eds.) ICAT 2006. LNCS, vol. 4282, pp. 1156–1165. Springer, Heidelberg (2006)
3. Yuan, W., Schneider, M.: iNav: An Indoor Navigation Model Supporting Length-Dependent Optimal Routing. In: 13th AGILE Int. Conf. on Geographic Information Science (2010)
4. Goetz, M., Zipf, A.: Formal Definition of a User-Adaptive and Length-Optimal Routing Graph for Complex Indoor Environments. *Geo-Spatial Information Science* 14(2), 119–128 (2011)
5. Dao, et al.: Spatial models for context-aware indoor navigation systems: A survey. *International Journal of Geographical Information Science* 26(3), 469–494 (2012)
6. Shen, J., Wu, M., Lv, G., Wen, Y., Wang, X., Zhang, H.: Topological relationships calculation for 3D curves data set based on monotone chains. In: 2010 18th International Conference on Geoinformatics, pp. 1–5. IEEE (2010)
7. Park, I., Jang, G.U., Park, S., Lee, J.: Time-Dependent Optimal Routing in Micro-scale Emergency Situation. In: 10th Int. Conf. on Mobile Data Management: Systems, Services and Middleware, pp. 714–719 (2009)
8. Yuan, W., Schneider, M.: Supporting 3d route planning in indoor space based on the lego representation. In: 2nd ACM SIGSPATIAL Int. Workshop on Indoor Spatial Awareness (ISA 2010), pp. 16–23. Springer, Heidelberg (2010)
9. Stoffel, E.-P., Lorenz, B., Ohlbach, H.J.: Towards a Semantic Spatial Model for Pedestrian Indoor Navigation. In: Hainaut, J.-L., et al. (eds.) ER Workshops 2007. LNCS, vol. 4802, pp. 328–337. Springer, Heidelberg (2007)
10. Xu, T., Zhou, Y.: Topology Analysis of Rural Power Distribution Network Based on Spatial Database. In: 2012 Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE (2012)
11. Lee, K.-J., Lee, J.: A Geocoding Method on Character Matching In Indoor Space. *Journal of Korea Spatial Information Society* 21(1), 87–100 (2013)
12. Lee, S., Lee, J.: Navigable Space-Relation Model for Indoor Space Analysis. *Journal of Korea Spatial Information Society* 19(5), 75–86 (2011)

Result Verification Scheme with Resource Clustering in Desktop Grids

Joon-Min Gil¹, Yong-Hyun Cho¹, and Sung-Hwa Hong²

¹ School of Information Technology Engineering, Catholic University of Daegu,
13-13 Hayang-ro, Hayang-eup, Gyeongsan, Gyeongbuk 712-702, S. Korea

² Dept. of Maritime Infom. and Comm. Engineering, Mokpo National Maritime University,
91 Haeyangdaehak-ro, Chukkyo-dong, Mokpo, Jeonnam 530-729, S. Korea
{jmgil, yhcho}@cu.ac.kr, shhong@mmu.ac.kr

Abstract. In desktop grids, it is important to guarantee the verification of task results in the presence of dynamic properties, such as volatility and heterogeneity. To achieve this objective, we propose the result-verification scheme with resource clustering that can tolerate the errors of task results. The scheme also determines the number of replications per task needed for result-verification on the basis of the resources classified by trusty degree and result-return probability. Simulation results show that our result-verification scheme can reduce the turnaround time of entire tasks even if few resources are used.

Keywords: Replication, Result-verification, Resource clustering, Replication, Desktop grids.

1 Introduction

Result-verification in desktop grids is required to guarantee the correctness of task results executed by any unspecified resources [1]. Voting-based scheme [2, 3] and trust-based scheme [4, 5] have been used for this objective, but they have two fundamental problems: waste of resources, due to redundant replications of each task, and increased turnaround time, due to the inability to deal with a dynamic execution environment. Such the dynamic execution can give rise to unexpected errors, because it relies substantially on unspecified and unstable resources. The errors arising in desktop grids can be grouped into two causes: internal and external. Internal causes include I/O errors occurring in a local disk, as storage for task results, precision problems due to heterogeneous resources from differing types of CPU and OS, and CPU miscalculations due to overclocking [6]. A typical external cause is a result alteration intentionally created by malicious resources [7]. As a result, result-verification is required to guarantee the correctness of task results obtained from resources in the dynamic execution environment of desktop grids.

To aggressively cope with these errors, we propose a result-verification scheme with resource clustering that can determine the number of replications per task needed for result-verification. The scheme is based on the classification of resources according to the trusty degree and the result return probability. Based on these two factors, the scheme performs the result-verification that allows correct task results to

be returned by a task deadline for current desktop grid environments. Therefore, we believe this approach provides a more efficient and flexible result-verification in the presence of the dynamic properties of desktop grids. Moreover, simulation results indicate that our scheme is superior to others in terms of turnaround time and quantity of resources consumed.

The remainder of this paper is organized as follows. In Section 2, we present our result-verification scheme based on resource clustering. The k-means clustering technique is briefly described and a detailed algorithm for our resource-verification scheme is presented in this section. Performance evaluations via simulations are described in Section 3. Finally, Section 4 concludes the paper.

2 Resource-Verification with Resource Clustering

2.1 Resource Clustering

To classify resources according to their dynamic properties in desktop grids, we first consider the definitions of trusty degree and result return probability as follows [8, 9].

Definition 1: Trusty degree – the factor to determine the correctness of task results executed by a resource.

Definition 2: Result Return Probability – the probability that a resource can complete a task by a given deadline, even if failures occur.

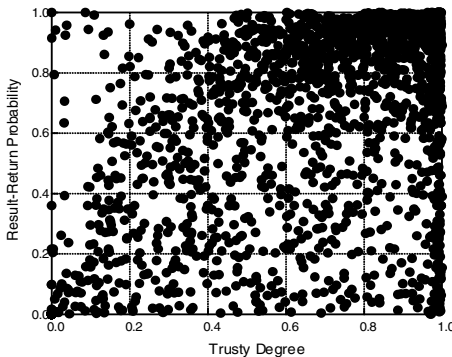


Fig. 1. Resource distribution [9]

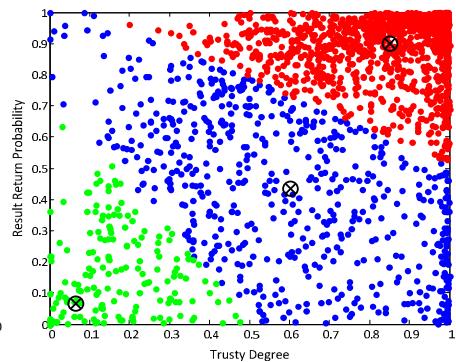


Fig. 2. Resource clustering [9]

The two factors presented in Definitions 1 and 2 are basically extracted using actual log data obtained from the Korea@Home desktop grid system [10]. Fig. 1 shows the resource distribution classified by trusty degree and result return probability, based on the actual log data that is accumulated during a period of one month.

Typically, clustering techniques are used to classify a set of data into classes of similar data [11]. It has been applied to various applications in many fields such as marketing, biology, pattern recognition, Web mining, and analysis of social networks.

Among the various clustering techniques, we choose the k-means clustering algorithm, which is an unsupervised learning algorithm, because of its effectiveness and simplicity. Mathematically, the k-means clustering algorithm can be described as follows:

$$E = \sum_{i=1}^k \sum_{j \in C_i} \|x_j - c_i\|^2$$

where k is the number of clusters, x_j is the j th data point in the i th cluster C_i , and c_i is the centroid of C_i . The notation $\|x_j - c_i\|^2$ stands for the distance between x_j and c_i , and Euclidean distance is commonly used as a distance measure. To achieve a representative clustering, the sum of the squared error function, E , should be as small as possible.

On the basis of the resource distribution presented in Fig. 1, we classify resources according to the two different kinds of resource features: trusty degree and result return probability. Fig. 2 shows the clustering result obtained when the number of classes (k) is 3.

2.2 Result-Verification

This section describes task replication and result-verification algorithms based on our resource clustering model. Fig. 3 and Fig. 4 show a task replication algorithm and result-verification algorithm, respectively.

```

1: TASKS = { $t_1, t_2, \dots, t_T$ }
2: RESOURCES = { $r_1, r_2, \dots, r_R$ }
3: while TASKS  $\neq \emptyset$  do
4:   waitForEvent()
5:   If (Event( $r_i$ , 'taskResult')) then
6:     select a task  $t_k$  TASKS
7:     ALLOCATION = { $r_i$ }
8:      $D$  = index of the closest cluster center to  $r_i$ 
9:     while (|ALLOCATION|  $\leq D$ ) do
10:      select a resource  $r_j \in$  RESOURCES with greatest contribution
11:      ALLOCATION = ALLOCATION  $\cup$  { $r_j$ }
12:    end while
13:    for all  $r_i \in$  ALLOCATION do
14:      distribute  $t_k$  to  $r_i$ 
15:      RESOURCES = RESOURCES - { $r_i$ }
16:    end for
17:    TASKS = TASKS - { $t_k$ }
18:  end if
19: end while

```

Fig. 3. Task replication algorithm

```

1: RESULTS =  $\emptyset$ 
2: while ( $|\text{RESULTS}| \leq D$ ) do
3:   waitForEvent()
4:   if (Event( $r_j$ , 'resultReturn')) then
5:     RESULTS = RESULTS  $\cup$  {  $res_j$  of  $r_j$  }
6:     RESOURCES = RESOURCES  $\cup$  {  $r_j$  }
7:   end if
8: end while
9: find a result  $res' \in \text{RESULTS}$  with majority
10: set  $S$  to a number of  $res'$  in RESULTS
11: if ( $S \geq \lceil D/2 \rceil$ ) then
12:   send  $res'$  to a management server
13: else
14:   TASKS = TASKS  $\cup$  {  $t_k$  }
15: end if

```

Fig. 4. Result-verification algorithm

3 Performance Evaluation

We evaluate the performance of our result-verification scheme based on resource clustering through simulations, in terms of turnaround time and quantity of resources consumed. We also present the performance comparison of our result-verification scheme with the existing scheme such as majority voting-based scheme.

Our simulations were conducted using the log data obtained from the Korea@Home desktop grid system [10] that were accumulated during a period of one month (March 2008). The following two metrics were used to measure the performance of the two schemes.

- Turnaround time: The total time taken between the submission of the first task for execution and the return of the complete result of the last task received by the server.
- Quantity of resources consumed: The total number of resources consumed for all tasks to be completed even if any failures.

Fig. 5 and Fig. 6 show the performance results of two result-verification schemes in terms of turnaround time and quantity of resources consumed, respectively. As we can see in these figures, our scheme has the fastest turnaround time, regardless of the number of tasks involved. Our scheme also has the smaller quantity of resources consumed as compared with the existing scheme. The majority voting-based scheme randomly selects the same number of resources as a fixed one of replications without considering trusty degree and task deadline. Whereas, our scheme can properly select

the resources needed in the procedure of result-verification in accordance with the trusty degree and result return probability of resources. In other words, our scheme can extract less number of replications for the resources with high trusty degree and result return probability by resource clustering.

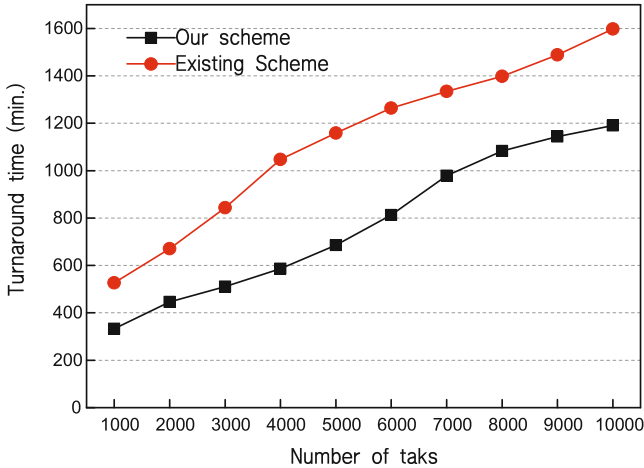


Fig. 5. Performance comparison (turnaround time)

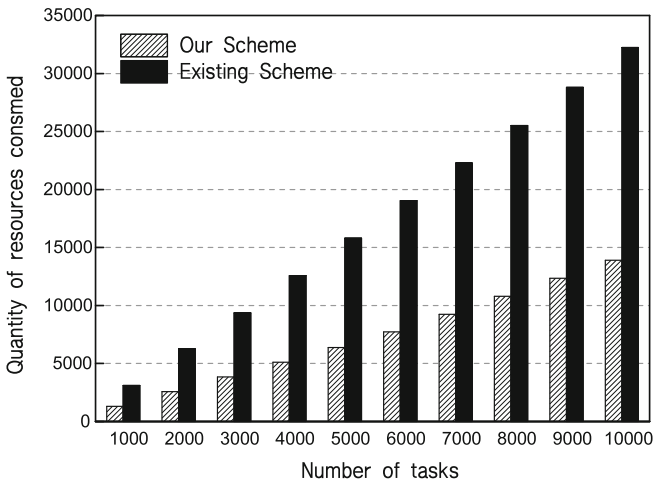


Fig. 6. Performance comparison (quantity of resources consumed)

4 Conclusion

We have presented a result-verification scheme based on resource clustering to efficiently support the result-verification, which is essential to ensure the correctness of task results from desktop grids. Our scheme incorporates trusty degree and result

return probability to determine the number of replications appropriate to the current resource-providing situation. It can thereby achieve correct task results by a given deadline. For performance evaluation, turnaround time and quantity of resources consumed have been considered to compare our scheme with the majority voting-based scheme. The results of the performance evaluation showed that our scheme had a faster turnaround time than the others, despite using fewer resources.

Acknowledgement. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2014R1A1A2055463).

References

- [1] Cerin, C., Fedak, G.: Desktop Grid Computing. Chapman and Hall/CRC (2012)
- [2] Kacsuk, P., Kovacs, J., Farkas, Z., Marosi, A., Balaton, Z.: Towards a powerful European DCI based on desktop grids. *Journal of Grid Computing* 9, 219–239 (2011)
- [3] Taufer, M., Anderson, D.P., Cicotti, P., Brooks III, C.L.: Homogeneous redundancy: a technique to ensure integrity of molecular simulation results using public computing. In: *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium*, pp. 4–8 (2005)
- [4] Zhao, S., Lo, V., GauthierDickey, C.: Result verification and trust-based scheduling in peer-to-peer grids. In: *Proceeding of the 5th IEEE International Conference on Peer-to-Peer Computing*, pp. 31–38 (2005)
- [5] Sonnek, J., Lino, L., Chandra, A., Weissman, J.B.: Adaptive reputation-based scheduling on unreliable distributed infrastructures. *IEEE Transactions on Parallel and Distributed Systems* 18, 1551–1564 (2007)
- [6] Kondo, D., Araujo, F., Malecot, P., Domingues, P., Silva, L.M., Fedak, G., Cappello, F.: Characterizing result errors in internet desktop grids. In: Kermarrec, A.-M., Bougé, L., Priol, T. (eds.) *Euro-Par 2007*. LNCS, vol. 4641, pp. 361–371. Springer, Heidelberg (2007)
- [7] Zhao, H., Li, X.: H-trust: A group trust management system for peer-to-peer desktop grid. *Journal of Computer Science and Technology* 24, 833–843 (2009)
- [8] Kim, H.-S., Hwang, C.-S., Lee, S.-K., Choi, S.-J., Gil, J.-M.: Priority based list scheduling for sabotage-tolerance with deadline tasks in desktop grids. *International Journal of Computer Systems Science and Engineering* 23, 121–131 (2008)
- [9] Gil, J.-M., Kim, S., Lee, J.: Task scheduling scheme based on resource clustering in desktop grids. *International Journal of Communication Systems* 27, 918–930 (2014)
- [10] Korea@Home, <http://koreaathome.org/eng/>
- [11] Xu, R., Wunsch, D.: *Clustering*. John Wiley & Sons, Hoboken (2008)

RDEA: A Novel Video Encryption Algorithm

Zhiyong Li, Xingjun Wang^{*}, Yuxi Lin, and Chao Cheng

Department of Electronic Engineering/Graduate School at ShenZhen,
Tsinghua University, China
{kismet_v,wangxingjun2015,linyuxide}@163.com,
chao3080103993@126.com

Abstract. More and more video encryption algorithms have been proposed with the popularity of multimedia applications and the necessity of content protection. Selective encryption algorithms are widely used because they can improve encryption efficiency meanwhile preserve the level of security. Visual degradation, security, encryption efficiency, compression friendliness, format compliance and codec compliance are used to evaluate a selective encryption algorithm. In this paper a novel video encryption algorithm called Random Data Encryption Algorithm (RDEA) is proposed. RDEA takes advantage of the fact that I-blocks are more important than P- and B-blocks, and selects data in I-blocks of I-, P- and B-frames to encrypt. The data is selected according to a pseudo-random sequence generated by a random sequence generator using a key as seed. The encryption ratio of RDEA is about 20%-35%, while it is almost as safe as naive encryption algorithm when it comes to ciphertext-only attack and known-plaintext attack. Also, RDEA keeps format compliant and has no impact on compression efficiency.

Keywords: Video encryption, selective encryption, security, encryption efficiency.

1 Introduction

With the improvement of the network technique and CPU performance, video is more and more widely used. Video on demand, Internet television, video telephony, and video conferences are typical examples [1]. At the same time, video encryption algorithms have been studied. The straightforward way is to encrypt the whole video with traditional encryption algorithm, such as AES [2]. This approach is called naive algorithm approach [3][4]. However, the naive algorithm is rarely used to encrypt a video because of the peculiarities of the video. For example, videos usually have large data size and high real-time demands, and use specific coding standard, such as H.264. Therefore, selective encryption algorithms are widely used to encrypt videos. Selective encryption algorithms select some key information of the video in order to reduce encrypted data and ensure the security at the same time [5].

^{*} Corresponding author.

The paper is organized as follows. A classification method and typical algorithms are discussed in section 2. Next in section 3, first we propose a new selective encryption algorithm called Random Data Encryption Algorithm (RDEA), then we introduce the criteria to judge encryption algorithms and analyze the performance of RDEA according to the criteria. In section 4, we conclude the paper.

2 Classification of Video Encryption Algorithms

In this section, we differentiate encryption algorithms into two classifications: joint compression and encryption algorithms and compression-independent encryption algorithms. This classification is based on algorithms' association with video codec.

2.1 Joint Compression and Encryption Algorithms

The main idea of joint compression and encryption algorithms is that encryption is applied in a certain step of the compression algorithm so that the output is significantly different from a video stream using a standardized compression algorithm. Many encryption algorithms are performed in this phase, such as the zigzag permutation algorithm proposed by Tang [10], MHT scheme proposed by Wu and Kuo [12], and REC/RPB proposed by Xie and Kuo [13].

Each joint compression and encryption algorithm has its own advantages and disadvantages. But the algorithms share some common features because they are all performed within video codec. Most of them have high encryption efficiency and keep format compliant. However, almost all of them will change standardized video codecs, and some have significant impact on compression efficiency.

2.2 Compression-Independent Encryption Algorithms

For compression-independent encryption algorithms, encryption and compression are performed separately. Therefore, they share some strengths and weaknesses. For example, they don't change video codecs. Compression-independent encryption algorithms can be further classified into two categories: algorithms before video codec and algorithms after video codec. Generally speaking, algorithms before video codec are rarely used because they destroy temporal correlation and spatial correlation of raw video data leading to low compression efficiency.

Algorithms after video codec include a selective algorithm proposed by Qihua Wang and Xingjun Wang [14], VEA proposed by Qiao and Nahrstedt [11], puzzle algorithm proposed by Liu and Koenig [3], and so on. The new algorithm proposed in this paper is also performed after compression.

3 The RDEA Algorithm

In this section, we first give an overview on the basic idea of RDEA. Then we introduce the criteria to judge encryption algorithms and analyze the performance of RDEA according to the criteria.

3.1 Basic Idea

The new video encryption algorithm is called Random Data Encryption Algorithm (RDEA). RDEA is performed after compression. The basic idea of RDEA is that we only encrypt data that we select randomly.

In this paper, we describe our algorithm assuming that the video to encrypt uses H.264 as the codec. However, RDEA can also be implemented in the video using other codecs.

In H.264 standard, prediction is carried out to remove redundancy and reduce the data size. Prediction comprises intra-frame prediction and inter-frame prediction. The former is to remove temporal redundancy while the latter can remove spatial redundancy. A frame is composed of multiple blocks. I-blocks just perform intra-frame prediction while P- and B-blocks perform inter-frame prediction. Although I-frames only include I-blocks, P-frames and B-frames comprise both I-blocks and blocks predicted by data in other frames.

Based on the fact that P- and B-frames cannot be decoded without I-frames, data to be encrypted should be in I-frames. However, some scene contents are still discernible by directly playing back the selectively encrypted video stream on a conventional decoder because the unencrypted I-blocks in the B- and P-frames can be fully decoded without any information from the I-frames. Therefore we select data to encrypt in I-blocks of I-, P- and B-frames.

To select data randomly, RDEA use a random sequence generator to generate a pseudo-random sequence S labeled by '0' or '1' using a key as seed. The generator can be a stream key generator, such as Rabbit [9].

We classified data in I-blocks into two parts: part E and part UE . If one bit corresponds to '1' in S , it will be added into E ; otherwise, it will be added into UE . Data in E will be encrypted using traditional encryption algorithm, such as AES. Data in UE will not be encrypted. In Fig.1, the above is the pseudo-random sequence S , and the below is data in I-blocks. According to S , d1, d5, d6, d8, d9, d12, d13 are put into E , while the other bits are put into UE .

The basic algorithm can be described in the following four steps.

- Use a random sequence generator to generate a pseudo-random sequence S .
- Record positions of I-blocks in I-, P- and B-frames.
- Select bits in I-blocks according to S and encrypt them.
- Replace the original bits in selected positions with encrypted bits.

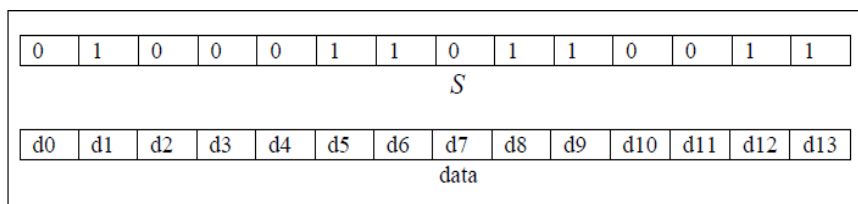


Fig. 1. The pseudo-random sequence S and data in I-blocks

3.2 Experiments and Analysis

In this part, first we define a set of evaluation criteria to measure an encryption algorithm. For every criterion, we analyze the performance of RDEA according to the principal and experiment results. The experiments are carried out on the following videos: paris, foreman, mobile, silent, and tempete. At the end of this part, we summarize the advantages and disadvantages of RDEA.

The performance of REDA and other algorithms are shown in Table.1 [1].

1) Visual degradation

Visual degradation measures the perceptual distortion of the cipher image (or video) with respect to the plain image (or video). It assumes that the cipher image (or video) can be decoded and viewed without decryption [6].

In Fig.2, the above are pictures before encryption, and the below are pictures after encryption. As we can see, pictures cannot be discerned after encryption.

2) Security

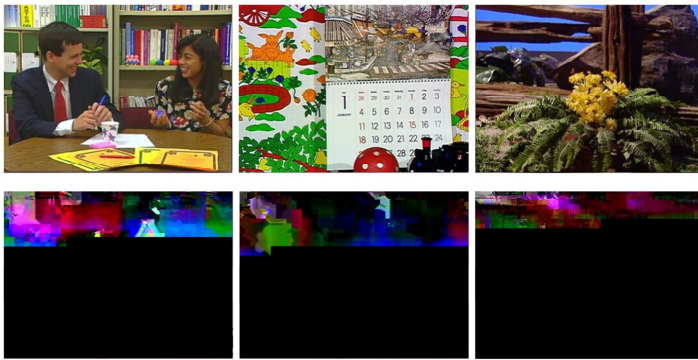


Fig. 2. Pictures before and after encryption

Table 1. Performance of RDEA and other algorithms

Algorithm	Security Weaknesses	Encryption Ratio	Decrease of Compression Efficiency	Format Compliance	Codec Compliance
RDEA	Chosen-plaintext attack	20%-35%	No	Yes	Yes
VEA	No attack reported	53%	No	No	Yes
REC/RPB	No attack reported	50%	None for Huffman, 5% for arithmetic	Yes	No
Zigzag Permutation	Ciphertext-only attack	1.56%	45% for MPEG 55% for H.263	Yes	No
MHT	Known-plaintext attack	5%	None for Huffman, 5% for arithmetic	Yes	No

The purpose of encryption is to ensure the security of the video content. Selective encryption is to reduce the encryption complexity meanwhile preserve the level of security. The security of the video has two factors to decide. One is the encryption key and the other is the unpredictability of the encryption data [6][7]. So security can be expressed with the following formula [14]:

$$CS = \min(S_{key}, S_{ciphertext}) \quad (1)$$

In the formula, S_{key} represents the space of the key, while $S_{ciphertext}$ represents the space of the ciphertext.

In our experiments, we use a key of 256 bits as seed to generate S . And the length of the key used in AES is also 256 bits. For a piece of ciphertext of n bits, the space of the ciphertext is shown in formula (2). However, if attackers try to decrypt the ciphertext through the key, they have to find out the data that has been encrypted first, which means the space of the key is larger than the complexity of finding out the encrypted data, as shown in formula (3). Therefore, we can get the security through formula (4).

$$S_{ciphertext} = 2^n \quad (2)$$

$$S_{key} \geq 2^{256} \quad (3)$$

$$CS = \min(S_{key}, S_{ciphertext}) \geq 2^{256} \quad (4)$$

From formula (4), we know that RDEA is at least as safe as the naive algorithm when it comes to ciphertext-only attack. Also, RDEA can resist known-plaintext attack. However, RDEA is vulnerable to chosen-plaintext attack because attackers can find out the data that has been encrypted with ease.

3) Encryption efficiency

This measurement is especially important in real-time applications, such as online conferences and film online. We usually use encryption ratio to measure encryption efficiency. Encryption ratio measures the ratio between size of encrypted data and size of the whole video. The equation of encryption ratio is shown in formula (5).

$$ER = \frac{N_{bits_encryption}}{N_{bits_all}} \quad (5)$$

Experiments on the five videos show that the encryption ratio of RDEA is about 20%-35%.

4) Compression friendliness

Compression friendliness measures the size change after adding the encryption operation. It can be represented by the decrease of compression efficiency as shown in formula (6).

$$R_{compression_friendliness} = \frac{N_{bits_encrypted_video} - N_{bits_unencrypted_video}}{N_{bits_unencrypted_video}} \quad (6)$$

After encrypting selected data, RDEA will write encrypted data back. So RDEA will not change the size of the video which means it is compression friendly.

5) Format compliance

Format compliance represents that the encryption algorithms do not change the compression standard format and the encrypted video can be decoded like the unencrypted video [1][8].

RDEA only encrypts media data in I-blocks, so it will not change the format of the video. Therefore, RDEA can keep format compliant.

6) Codec compliance

Codec compliance measures whether the encryption algorithm changes the video codec. In some occasions, the video codec is implemented by hardware and therefore cannot be changed. An algorithm is considered codec compliant when it doesn't change the video codec.

RDEA is performed after compression, so it will not change the video codec, which means that it can keep codec compliant.

From Table.1 and the analysis above, we can see that RDEA can resist ciphertext-only attack and known-plaintext attack while its encryption ratio is just 20%-35%. Furthermore, it keeps format and codec compliant and has no impact on compression efficiency. However, RDEA is vulnerable to chosen-plaintext attack.

4 Conclusion

In this paper, we propose a new video encryption algorithm called Random Data Encryption Algorithm (RDEA). RDEA has high efficiency while it is almost as safe as naive encryption algorithm considering ciphertext-only attack and known-plaintext attack. Besides, RDEA has no impact on compression efficiency and keeps format compliant and codec compliant. However, RDEA cannot withstand chosen-plaintext attack. In the future, we will improve RDEA to be secure enough to withstand chosen-plaintext attack.

References

- [1] Liu, F., Koenig, H.: A survey of video encryption algorithms. *Computers & Security* 29(1), 3–15 (2010)
- [2] NIST: Advanced Encryption Standard. FIPS 197 (2001)
- [3] Liu, F., Koenig, H.: A novel encryption algorithm for high resolution video. In: *Proceeding of ACM NOSSDAV 2005*, Stevenson, WA, USA, pp. 69–74. ACM Press, New York (2005)
- [4] Agi, I., Gong, L.: An Empirical Study of MPEG Video Transmission. In: *Proceedings of the Internet Society Symposium on Network and Distributed System Security*, pp. 137–144 (1996)
- [5] Wang, Q.H., Wang, X.J.: A New Selective Video Encryption Algorithm for the H. 264 Standard [C]

- [6] Massoudi, A., Lefebvre, F., De Vleeschouwer, C., et al.: Overview on selective encryption of image and video: challenges and perspectives. *EURASIP Journal on Information Security* 2008, 5 (2008)
- [7] Wang, Q.H., Wang, X.J.: Study of Selective Video Encryption for the H. 264 Standard. *Applied Mechanics and Materials* 397, 2097–2103 (2013)
- [8] Stutz, T., Uhl, A.: A Survey of H.264 AVC/SVC Encryption. *IEEE Transactions on Circuits and Systems for Video Technology* 22(3) (2012)
- [9] Boesgaard, M., Vesterager, M., Pedersen, T., Christiansen, J., Scavenius, O.: Rabbit: A New High-Performance Stream Cipher. In: Johansson, T. (ed.) *FSE 2003*. LNCS, vol. 2887, pp. 307–329. Springer, Heidelberg (2003)
- [10] Tang, L.: Methods for encrypting and decrypting MPEG video data efficiently. In: *ACM International Conference on Multimedia*, Boston, MA, USA, pp. 219–229 (November 1996)
- [11] Qiao, L., Nahrstedt, K.: Comparison of MPEG encryption algorithms. *Computer and Graphics* 22(4), 437–448 (1998)
- [12] Wu, C.-P., Kuo, C.-C.J.: Design of integrated multimedia compression and encryption systems. *IEEE Transactions on Multimedia* 7(5), 828–839 (2005)
- [13] Xie, D., Kuo, C.-C.J.: Multimedia encryption with joint randomized entropy coding and rotation in partitioned bitstream. *EURASIP Journal on Information Security* 2007(1) (January 2007)
- [14] Wang, Q.H.: *Researches of Selective Video Encryption Algorithms*. Tsinghua University, Beijing (2014)

Embedding Data into Audio Signal by Combining Sliding Window Technique with a Novel Marking Bit Method

Vu Van Tam¹, Tran Duc-Tan^{2*}, Nguyen Thanh Thuy¹, and Phan Trong Hanh³

¹ University of Technology and Logistics, Vietnam
tamt36bca@gmail.com

² VNU University of Engineering and Technology, Vietnam
tantd@vnu.edu.vn

³ Le Quy Don Technical University, Vietnam
tronghanhmai@yahoo.com

Abstract. Embedding data into audio signal is an interesting field and attracted by many scientists. It allows taking advantage of available transmission lines to transmit data with safety and high security. However, the biggest challenge while embedding data in the audio signal is to achieve a small embedding error (EE) and large signal-to-noise ratio (SNR). If EE is large and SNR is small, it is vulnerable to be tracked and attacked by enemies during transmitting embedded audio signals. In this paper, we have proposed to combine the sliding window technique with a novel marking bit method. The sliding window technique is used to find out the bit segment of the original audio signal coinciding with the bit segment of data needed to be embedded. After that, we mark the location of windows found by the coded bit, and overwrite the coded bit on the original audio signal at the appropriate locations. The experimental results show that the proposed method provides a better quality than using other embedded methods such as the LSB (Least Significant Bit), ELS (Embedding Large Sample) and without using sliding windows.

Keywords: Sliding window, Embedding error, Audio signal, Signal-to-noise ratio, Embedding signal.

1 Introduction

Embedding signal is a technique to hide certain information or data in the original signal to preserve confidentiality or authenticate information [1] [6]. Embedded data and the original signal can be in the form of text, audio, image, video ...Original signal which brings in its embedding data is called embedding signal. When the original signal is audio, it will have the following advantages: (i) The audio signal is one of the popular signals, widely transmitted on the systems such as: radio and television broadcasting, internet.... Therefore, the opponents will be hard to control, block, catch and attack; (ii) Audio signal usually has a high sample rate (SR) from 2 kHz to 48 kHz, bits per sample (BP) from 8 bit to 24 bit. Thus, it is possible to embed high capacity data [2] [3]; (iii) The audio signal has parameters: amplitude, frequency,

* Corresponding author.

phase. So the parameters above may be impacted to embed data [3]. Despite the advantages mentioned above, the system of human audition (hearing) is sensitive to changes in the audio signal. Therefore, EE and SNR must be particularly interested when embedding data in the audio signal.

According to the authors' survey, there are now some embedding data into audio signal methods such as: (i) LSB method [2] [3] [4] [5]; (ii) Parity method [4]; (iii) Phase method [3] [4] [5]; (iv) Echo method [4]; (v) Spectrum spreading method [1] [2] [7]. (vi) ELS method (Embedding Large Sample) [9].

In this paper, we propose sliding window (SW) method combining with bit marking (MB) to embed data in audio signals. In particular, SW will increase the capacity of embedded data and MB allows reducing EE, increasing the SNR of the embedded signal. The results of analysis, calculations are verified by experiment and compared with LSB [2] [3] [4] [5] embedding method, ELS [9], and without using sliding windows.

2 System Model

- **Model Structure:** Fig.1 is embedded model, de-embedding according to proposed method, the symbols are described as follows: C is the original signal (1), C_j is audio samples, n is the number of audio samples of C , C_∂ is the segment of the original audio signal (2), ζ is the number of audio samples of C_∂ , p is the number of segments C_∂ of C , and $p = \frac{n}{\zeta}$.

$$C = \{C_j; j = 1, 2, \dots, n\} \tag{1}$$

$$C = \{C_\partial; \partial = 1, 2, \dots, p\} \quad \text{with} \quad C_\partial = \{C_j; j = 1, 2, \dots, \zeta\} \tag{2}$$

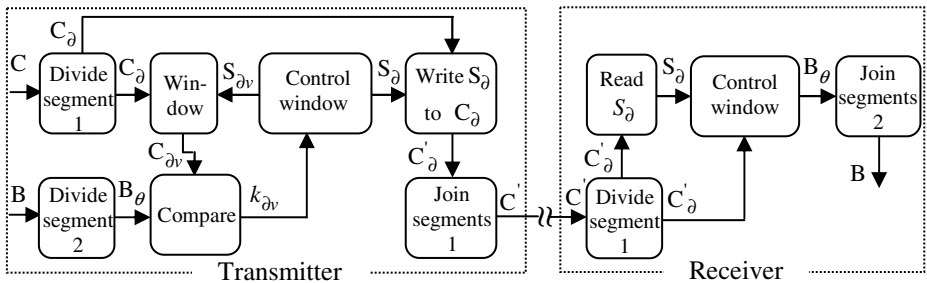


Fig. 1. The proposed procedures of embedding and de-embedding data

B is the data needed embedding into C , $B = m[\text{bit}]$. B_θ is the segment of data needed embedding (3), $B_\theta = \Omega[\text{bit}]$, with q is the number of segments B_θ of B and $q = \frac{m}{\Omega}$.

$$B = \{B_\theta; \theta = 1, 2, \dots, q\} \tag{3}$$

$S_{\partial v}$ is the bit string marking the location of the v window in C_{∂} , $S_{\partial v} = u$ [bit] with $v = 1, 2, \dots, w$. w is the number of windows created in C_{∂} and $w = 2^u$. $C_{\partial v}$ is the bit string in v window, $C_{\partial v} = \Omega$ [bit]. $k_{\partial v}$ is comparing bit: (i) If $C_{\partial v} \equiv B_{\theta}$ then $k_{\partial v} = 1$; (ii) If $C_{\partial v} \neq B_{\theta}$ then $k_{\partial v} = 0$. S_{∂} is the sequence of bit marking. There would be two cases: (i) If one of w windows in C_{∂} gives the result $k_{\partial v} = 1$ then $S_{\partial} = \overline{0}1$ [bit], $\overline{0}1$ [bit] = $1 + u$, in which the first bit is equal to 1 (with note that the receiver in segment C_{∂} has B_{θ}) the rest bit is equal to $S_{\partial v}$; (ii) If all w windows in C_{∂} have the same result $k_{\partial v} = 0$ then $S_{\partial} = \overline{0}0$ [bit] = 1 [bit] and $S_{\partial} = 0$. C'_{∂} is the segment of embedded audio signals (4) with C'_j is embedded audio sample. C' is the embedded audio signal (5).

$$C'_{\partial} = \{C'_j; j = 1, 2, \dots, \zeta\} \tag{4}$$

$$C' = \{C'_{\partial}; \partial = 1, 2, \dots, p\} \tag{5}$$

- **Sliding Window:** Fig. 2 describes sliding window according to proposed method. Distance of sliding step is 1 bit. The number of windows in 1 C_j is $k = BP - \Omega$. The number of windows in 1 C_{∂} is $l = k \times \zeta$

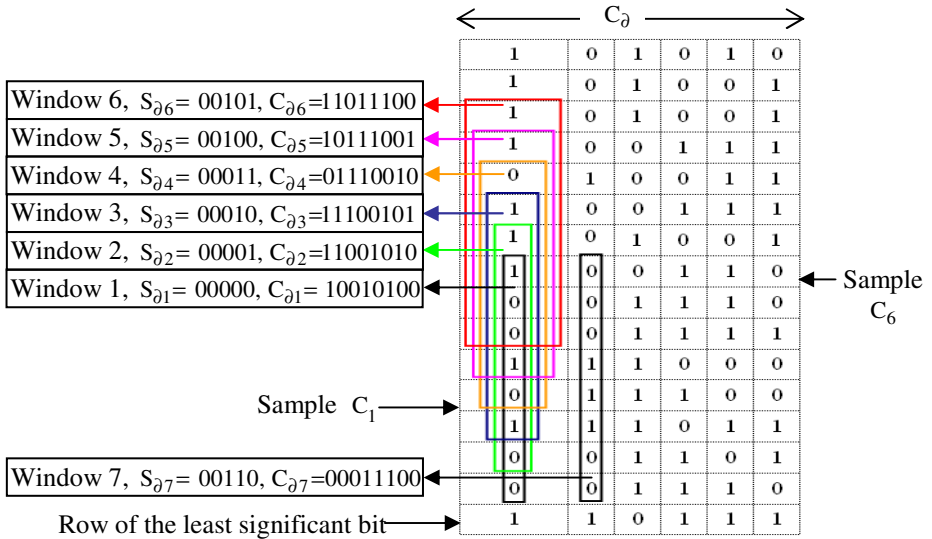


Fig. 2. Description of sliding window, the values $S_{\partial v}$ and $C_{\partial v}$ corresponding in segment C_{∂}

- **The Process of Embedding:** Firstly, C is divided into p segments C_{∂} as shown in Equ. (2), and B is divided into q segments B_{θ} as shown in Equ. (3). Secondly, embedding B_{θ} into C_{∂} as follows: $S_{\partial v}$ locates the windows in C_{∂} ; $C_{\partial v}$ is

compared to B_θ , if $k_{\partial v} = 0$ then the window will move to a new location ($v = v + 1$) and continue to compare $C_{\partial v}$ with B_θ . If $k_{\partial v} = 1$ or $v = w$ and $\{k_{\partial v}; v = 1, 2, \dots, w\} = 0$, S_∂ will be overwritten on the location of the least significant bit of C_∂ to form C'_∂ (4). Finally, merging all C'_∂ into C' as shown in Equ. (5) to transmit to the receiver. In the case, the segment B_θ will be embedded into the next.

- **The Process of De-embedding:** C' is handled by reverse processes comparing to the transmitter to extract S_∂ in each C'_∂ . If the first bit of S_∂ is equal to 1, then the corresponding window is regenerated to read out B_θ . Connect q segments B_θ into B data (3).

3 Analysis and Calculations of Parameters

- **Embedding Error (EE):** According to [1] [4] EE is calculated by the expression (6). However, with the proposed method, EE is calculated by S (7), with p_0 is the number of segments C_∂ which is overwritten 1 bit. Some bits of S may coincide with the bit of C in the location overwritten, so $EE [\text{bit}] \leq S [\text{bit}]$. Therefore, EE is calculated by (8). Because the bits of S is overwritten on the location of least significant bit of C, EE is calculated by amplitudes is equivalent to computing EE by bit (9).

$$EE = |C' - C| = \sum_{j=1}^n |C'_j - C_j| = \sum_{j=1}^n |EE_j| \quad (6)$$

$$S[\text{bit}] = (m \times \bar{\mathcal{U}}_1) + (p_0 \times \bar{\mathcal{U}}_0) = (m \times \bar{\mathcal{U}}_1) + p_0 \quad (7)$$

$$EE[\text{bit}] \leq (m \times \bar{\mathcal{U}}_1) + p_0 \quad (8)$$

$$EE = EE[\text{bit}] \leq (m \times \bar{\mathcal{U}}_1) + p_0 \quad (9)$$

- **Signal-to-Noise Ratio (SNR):** According to [1] [4], SNR is calculated by:

$$SNR = \frac{\sum_{j=1}^n |C_j|^2}{\sum_{j=1}^n |EE_j|^2} = \frac{\sum_{j=1}^n |C_j|^2}{\sum_{j=1}^n |C'_j - C_j|^2} = 10 \log_{10} \left[\frac{\sum_{j=1}^n |C_j|^2}{\sum_{j=1}^n |C'_j - C_j|^2} \right] [\text{dB}] \quad (10)$$

From (8) and (9), it can be seen that there is no larger than $(m \times \bar{\mathcal{U}}_1) + p_0$ audio samples would change their least significance bit ($C'_j - C_j = 1$), so

$$SNR [\text{dB}] \geq 10 \log_{10} \left[\frac{\sum_{j=1}^n |C_j|^2}{(m \times \bar{\mathcal{U}}_1) + p_0 \sum_{j=1}^n |C'_j - C_j|^2} \right] = 10 \log_{10} \left[\frac{\sum_{j=1}^n |C_j|^2}{(m \times \bar{\mathcal{U}}_1) + p_0} \right] \quad (11)$$

- **Maximum Embedded Capacity** (D_{max}): Assuming all C_0 are embedded, then

$$D_{max}[\text{bit}] = \frac{n}{\zeta} \times \Omega = p \times \Omega \tag{12}$$

- **Parameter Ω** : The windows slide vertically (as shown in Fig. 2), so $2 \leq \Omega \leq BP$. If Ω is small, according to (12), D_{max} will decrease. On the contrary, if Ω is large, p_0 will increase, and according to (9) (10), EE will increase and SNR decrease. It can be seen that Ω is proportional to the amplitude fluctuations of C . Therefore, if C has a slow tempo (voice, light music...), $5 \leq \Omega \leq 7$ and if C has a fast tempo (rock, rap...), $8 \leq \Omega \leq 10$. The experimental method can be used to find the best value.

4 Experiment Results

- **Testing scenario**: To test the effectiveness of the proposed method, we have installed the embedded and de-embedding algorithms according to the proposed model in order to compare the results with LSB [2] [3] [4] [5], ELS [9] and not sliding window. The testing scenario is as follows: C is 50 different audio files $\{C_\omega; \omega = 1, 2, \dots, 50\}$, B is 10 text segments $\{B_\mu; \mu = 1, 2, \dots, 10\}$, which have a capacity of 100 bits to 18000 bits. Each C_ω is embedded into 10 text segments using all methods mentioned above. Get the average of parameters: C , B , EE, SNR of each method to make comparison charts.

- **Analysis of Results**: 1) The test result shows that the proposed method has the better quality of EE, SNR and EE amplitude than the LSB, ELS methods (Fig. 3, Fig. 4, Fig. 6), and the capacity of embedded data is 5 times larger than the case of without using the sliding window (Fig. 5, Fig. 6c). There are above advantages due to: (i) Using S_δ with $S_\delta = \mathcal{U}_1[\text{bit}]$ to encode the location of B_θ with $B_\theta = \Omega[\text{bit}]$ and $\mathcal{U}_1 < \Omega$. (ii) By using the sliding window, so ζ and p_0 decrease. (iii) Overwriting bits of S on the least significant bits of C so $EE = EE[\text{bit}]$.

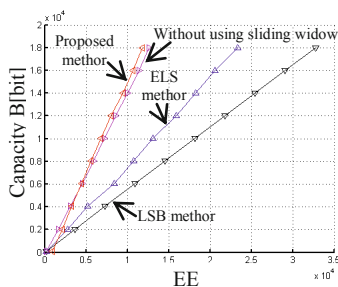


Fig. 3. EE performance

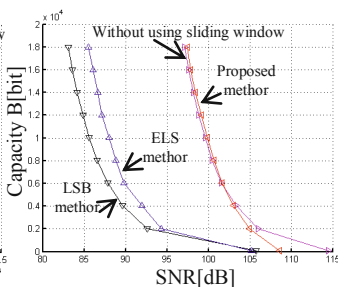


Fig. 4. SNR performance

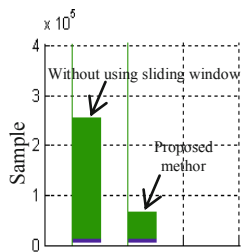


Fig. 5. Embedding capacity

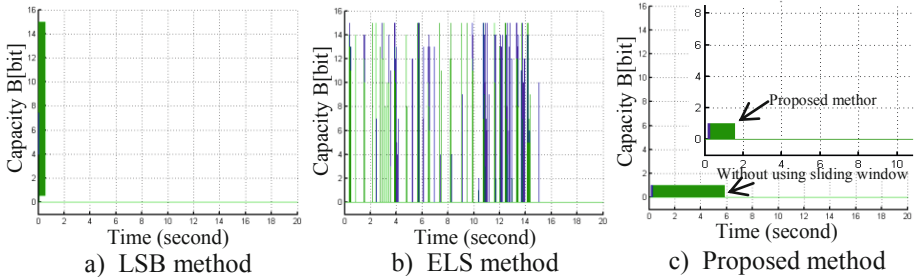


Fig. 6. The amplitude and distribution of EE

2) However, when B_θ is small ($B_\theta \leq 6 \times 10^4[\text{bit}]$), the proposed method has the lower quality of EE, and SNR than the case without the sliding window (see Fig. 3, and Fig. 4). The reason is that: At the beginning and the end of audio file, there is usually white space, when ζ decreases, the probability $C_\theta = \{C_j; j = 1, 2, \dots, \zeta\} = 0$ increases, so p_0 increases, too. According to (7) (8) (9), the increase in EE leads to the decrease in SNR. In addition, Fig. 5 and Fig. 6b shows that the proposed method is not as good as the ELS method in the EE distribution. Two limitations above will motivate the authors to continue to research further in this field.

5 Conclusion

In this paper, we propose to use the sliding window technique combining with the bit marking method to embed effectively data in audio signals. In particular, the sliding window method creates many bit segments in order to compare with the bit segments of the embedded data; the bit marking method reduces the number of bits overwritten on the original audio signal. The results show that the quality of EE, SNR, the amplitude embedding error of the proposed method is better than the LSB method [2] [3] [4] [5], ELS [9] and the capacity of embedded data has significantly improved compared with the case without the sliding window. Although there are a few minor drawbacks, but the proposed method has helped to improve the quality of the embedded audio signal.

References

[1] Baluja, Y., Mishra, S., Saini, T.S., Patil, M.V.: Frequency domain based data hiding technique for audio signal. International Journal of Innovative Research in Science, Engineering and Technology 2(5), 1564–1569 (2003)
 [2] Babu, L., Jais John, S., Parameshachari, B.D., Muruganantham, C., Divakaramurthy, H.S.: Steganography method for data hiding in audio signal with LSB & DCT. International Journal of Computer Science and Mobile Computing 2(8), 54–62 (2013)

- [3] Kumar, S., Barnali, B., Banik, G.: LSB modification and phase encoding technique of audio steganography revisited. *International Journal of Advanced Research in Computer and Communication Engineering* 1(4), 1–4 (2012)
- [4] Malviya, S., Nayak, K.K., Saxena, M., Khare, A.: Audio steganography in a nutshell. *International Journal of Electronics Communication and Computer Technology (IJECCCT)* 2(5), 219–222 (2012)
- [5] Saroha, K., Singh, P.K.: A variant of LSB steganography for hiding images in audio. *International Journal of Computer Applications* (0975-8887) 11(6), 12–16 (2010)
- [6] Bhattacharyya, D., Dutta, P., Kim, T.-H.: Secure data transfer through audio signal. *Journal of Security Engineering* 6, 187–194 (2009)
- [7] Geiger, R., Yokotani, Y., Schuller, G.: Audio data hiding with high data rates based on INTMDCT. In: *IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*, pp. V-205–V- 208
- [8] Petrovic, R., Winograd, J.M., Jemili, K., Metois, E.: Data hiding within audio signals. *Facta Universitatis Series: Electronics and Energetics* 12(2), 103–122 (1999)
- [9] Van Tam, V., Hanh, P.T.: A new method to embed data into audio signals. *Journal of Army Science and Technology Research* 33, 13–19 (2014) (in Vietnamese)

A Study on the Hierarchical Structure of Conditional Replenishment Algorithm for SC-MMH 3DTV System*

Kyeong-Hoon Jung¹, Sung-Hoon Kim², and Dong-Wook Kang¹

¹ School of EE, Kookmin University, Republic of Korea

² Electronics and Telecommunications Research Institute, Republic of Korea
{khjung, dwkang}@kookmin.ac.kr, steve-kim@etri.re.kr

Abstract. CRA (Conditional Replenishment Algorithm) has been proposed to improve the visual quality of asymmetrical 3DTV where the resolutions of two views are not matched. CRA adopts variable-sized PU (Processing Unit) and quad-tree is employed to represent its hierarchical structure. In this paper, we determine the largest and smallest sizes of PU which reduce the complexity while keeping the quality degradation as small as possible. In case of HD resolution images, the processing time of CRA can be reduced to 60 % but the decrease of BD-PSNR is only 0.034 dB by limiting the size of PU from 4x4 to 128x128.

Keywords: SC-MMH, CRA (Conditional Replenishment Algorithm), PU (Processing Unit), quad-tree, disparity vector.

1 Introduction

SC-MMH (Service Compatible 3DTV using Main and Mobile Hybrid) system [1], which fully guarantees backward compatibility with existing main and mobile ATSC-M/H services [2], was proposed as one of prominent candidates for terrestrial 3D broadcasting service. Quality mismatch might occur in SC-MMH system, that is, one view encoded by the MPEG-2 video codec is transmitted via ATSC main channel and the other view encoded by the AVC is transmitted via ATSC mobile channel. This mismatch problem can be considerably settled by BSE (Binocular Suppression Effect) of human visual system [3]. However, when the difference between two views is quite large, it is desirable to enhance the lower quality view and CRA (Conditional Replenishment Algorithm) has been proposed for this purpose [4].

The basic idea of CRA is to use the correlation between two stereoscopic views. In order to compose 3D, the view via ATSC main service are compensated by disparity vector and view via ATSC mobile service are simply interpolated. After that, the best substitute to generate enhanced view is decided as mode information. The optimum disparity vector and mode information are estimated during encoding process and need to be transmitted as additional data to be utilized at 3DTV receiver.

* This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program.

Since there is much spatial correlation in disparity vectors and modes, the pixels which have the same disparity vectors and modes can be merged into a single PU (Processing Unit) in order to keep the amount of additional data as small as possible. The appropriate size of PU depends on the spatial characteristics of images and quad-tree structure with variable block-size is employed to represent the overall distribution of PUs. There is also much temporal correlation in disparity vectors and modes of consecutive frames and the video via ATSC main service which is compensated by previous disparity vectors and modes can be a third candidate. Therefore there are two types in mode decision of CRA; one is INTRA type in which the above two substitutes are used and the other is INTER type in which the third temporal candidate is also considered. It was showed that just a small amount of additional data could greatly improve the visual quality of video via ATSC mobile service by using CRA [4]. However, the computational load of CRA is considerable since the disparity vectors are estimated at each layer in hierarchical quad-tree structure.

In this paper, we investigate the effect of hierarchical structure of quad-tree on the performance and complexity of CRA. The largest and smallest sizes of PU, which reduce the processing time while keep the quality degradation as small as possible, are determined by simulation on several sequences with various spatial and temporal characteristics.

2 Hierarchical Structure of CRA

When the bitrate for additional information is given, the hierarchical structure of CRA is adaptively determined through RD (Rate Distortion) optimization. In order to determine the optimum mode in CRA, the cost function of Eq. (1), where D is MAD(Mean Absolute Difference) and R is the amount of required bits, was introduced.

$$J = D + \lambda R \quad (1)$$

At a given layer, the cost functions of every possible mode are calculated for each PU. And the mode which has the smallest cost is selected as the candidate mode of that PU. After every PU at that level are examined, the process moves to the upper level and determines whether four sub-PUs need to be merged or not by comparing the cost of each PU at the upper level and the sum of costs of its four sub-PUs at the lower level.

The Lagrange multiplier λ , which used for RD optimization, plays an important role to regulate the bitrate of additional data. By controlling λ , the additional bitrate for CRA can be kept as constant as possible.

Typical examples of quad-tree structure of CRA are shown in Fig. 1, where black empty PU denotes the simply interpolated part and colored PU denotes the disparity compensated part. As expected, the higher PSNR gain was obtained when more bitrate was used for CRA.

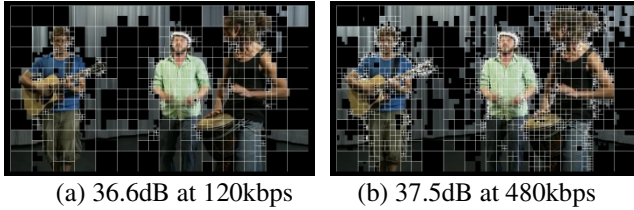


Fig. 1. The examples of CRA (PSNR at additional bitrate)

As the size of PU becomes smaller, the details are likely to be well recovered and the PSNR gain becomes higher. On the contrary, we can expect more coding gain when the size of PU becomes larger because just a single mode and disparity vector is assigned to each PU. As a result, it is better to keep the depth of hierarchical structure, that is, the number of levels as large as possible. However, when the depth is too large, the complexity can be a critical problem since the estimation of disparity vector requires quite substantial computational load and the disparity vector needs to be calculated at every level of CRA. Therefore, it is greatly desirable to limit the size of PU in order to reduce the complexity while keeping the quality degradation as small as possible.

3 Simulation Results

3.1 Experimental Environment

Total 10 test sequences, which are obtained from real video and animation, were used in experiment as shown in Fig. 2. As you can see, the spatial and temporal characteristics of test sequences are diversely varied. According to the service scenario of SC-MMH, the vertical resolutions of two view are 1080 and 480, respectively. The left view is encoded by MPEG-2 with 12Mbps and the right view is encoded by AVC with 480 kbps. And the additional bitrate for CRA is changed from 120 kbps to 480 kbps with a step size of 120 kbps.

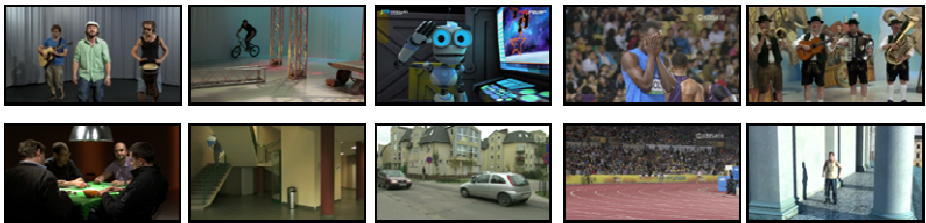


Fig. 2. Test sequences

In order to compare the complexity, we calculated the consumed processing times for every test sequences and averaged them at each additional bitrate as changing the upper and lower limit of hierarchical level. And TR (Time Ratio), which is given in

(2), was introduced as a relative measure of complexity [5]. T_c is the processing time of comparison target and T_r is that of reference target. As the complexity of comparison target becomes smaller than that of reference target, TR shows smaller value.

$$\text{TR} = T_c/T_r \times 100 (\%) \quad (2)$$

And the performances are compared by BD-PSNR. Similarly, we calculated the PSNRs of every test sequences and averaged them at each bitrate as varying the comparison target. The BR-PSNR of comparison target is always smaller than that of reference. As the difference between BD-PSNRs becomes smaller, the performance of comparison target is better.

3.2 Decision of the Largest and Smallest Size of PU

The first experiment was to decide the LPU (Largest Processing Unit). We fixed the smallest size of processing unit to 1x1 (level 0). And the sizes of LPU were changed from 32x32 (level 5) to 256x256 (level 8). The reference target was taken as the case when the LPU was 256x256, and the remaining three cases became comparison targets.

It can be found from Figure 3 that TR decreased in almost linear fashion as the size of LPU decreased. If the level of LPU was reduced by one, then the decrease rate of TR was approximately 7%. Meanwhile, Fig. 3 says that BD-PSNR decreased exponentially as LPU became smaller. When LPU was 128x128, the difference of BD-PSNR is quite small. But it increased to 0.16 dB when LPU was 64x64 and there was a considerable gap of 1.03 dB when LPU was further reduced to 32x32. Therefore, it is desirable to adopt the size of LPU not as small as 64x64.

Next experiment was to decide the SPU (Smallest Processing Unit). For this purpose, the largest size of processing unit was fixed to 256x256 (level 8) and the sizes of SPU were changed from 1x1 (level 0) to 8x8 (level 3). The size of SPU of reference target was 1x1, and the remaining three cases became comparison targets.

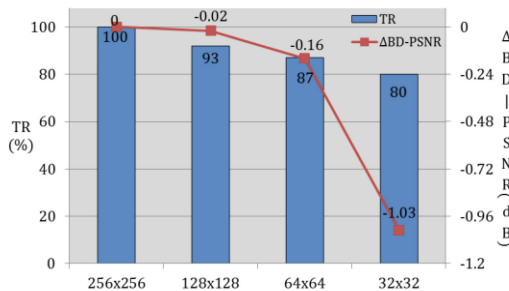


Fig. 3. TR and Δ BD-PSNR according to LPU

Fig. 4 shows the variation pattern of TR. As the level of SPU increased, TR decreased in almost linear fashion as before. But the decrease rate was nearly doubled. Also Fig. 4 gives the decrease of BD-PSNR according to SPU. It can be noticed that the difference of BD-PSNR was negligible until the size of SPU was 4x4, but there was an abrupt decrease in next level. Therefore, it is desirable to adopt the size of SPU not as large as 8x8 in order to avoid noticeable degradation in performance. As a result, LPU and SPU were decided to 128x128 (level 7) and 4x4 (level 2), that is, the hierarchical depth of quad-tree was six in our experiment.

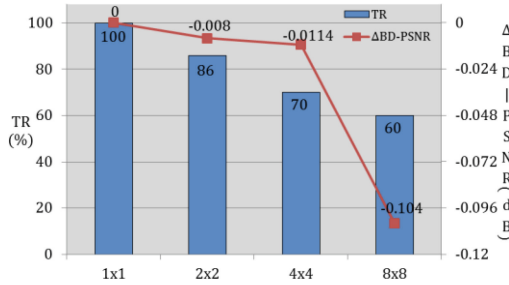


Fig. 4. TR and Δ BD-PSNR according to SPU

3.3 The Combined Result of LPU and SPU

Next, we observed the complexity and performance of CRA in case of the decided LPU and SPU were used. Thus, the size of PU of comparison target was from 4x4 to 128x128 and that of reference target was from 1x1 to 256x256. That is the numbers of level of comparison and reference target were six and nine, respectively. As before, 120 kbps, 240kbps, 360kbps and 480kbps were used as additional bitrate for CRA. And the processing times and PSNRs were calculated for each test sequence and averaged.

From Fig. 5, it could be noticed that the TR of comparison target was nearly constant and slightly less than 60% regardless of additional bitrate. Meanwhile, Fig. 6 showed that the RD curves of reference and comparison targets were very close. The PSNR gap between them was only 0.0034dB in terms of BD-PSNR, even though the relative complexity was reduced to about 60%.

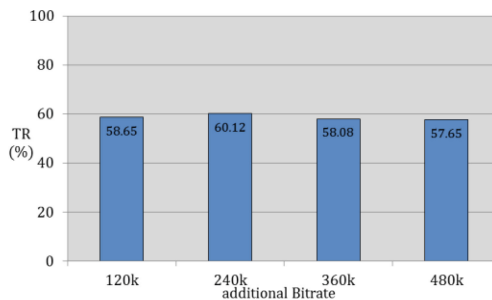


Fig. 5. TR according to additional bitrate

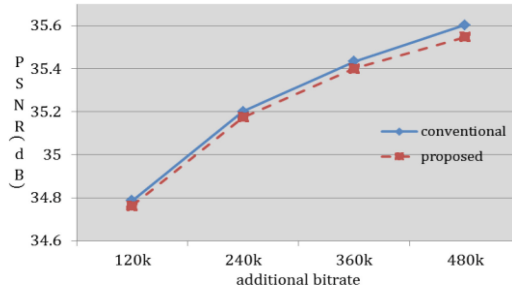


Fig. 6. RD curves according to additional bitrate

4 Conclusion

This research aims to make a framework for learning system on cloud computing environment. SC-MMH is one of the prominent candidates of terrestrial 3DTV system and CRA is the very efficient solution to solve the quality gap between two stereoscopic views of SC-MMH. And the hierarchical structure of CRA needs to be carefully determined, since the computational complexity of CRA can be a critical obstacle when the depth is too large.

In this paper, we examined the effect of the depth of quad-tree structured PUs on the complexity and performance of CRA. Total 10 test sequences of HD resolution were used for experiment. To decide the LPU, SPU was fixed to 1x1 and TRs and BD-PSNRs were calculated and averaged as the size of PU varied from 32x32 to 256x256. On the contrary, LPU was fixed to 256x256 and TRs and BD-PSNRs were calculated and averaged as the size of PU varied from 1x1 to 8x8 for decision of SPU. Simulation results show that the complexity and RD performances are in the trade-off relation and CRA with LPU of 128x128 and SPU of 4x4 reduces the processing time by 40% but decreases BD-PSNR by only 0.034dB compared with the original CRA with LPU of 256x256 and SPU of 1x1.

References

- [1] Kim, B., Bang, M., Kim, S., Choi, J., Kim, J., Kang, D., Jung, K.: A study on feasibility of dual-channel 3DTV service via ATSC-M/H. *ETRI Journal* 34(1), 17–23 (2012)
- [2] ATSC A/153, ATSC mobile/handheld digital television standard,” Advanced Television Systems Committee, Washington, D.C. (October 15, 2009)
- [3] Stelmach, L., et al.: Stereo image quality: effects of mixed spatio-temporal resolution. *IEEE Trans. on Circuits Syst. Video Technol.* 10(2), 188–193 (2000)
- [4] Jung, K., Bang, M., Kim, S., Choo, H., Kang, D.: Enhancement for hybrid 3DTV with mixed resolution using conditional replenishment algorithm. *ETRI Journal* 36(5), 752–760 (2014)
- [5] Goswami, K., et al.: An early termination algorithm for efficient CU splitting in HEVC. *Journal of Broadcast Engineering* 18(2), 271–282 (2013)

Performance Comparison for CFDP NAK Modes in Deep Space Communications

Yong Li¹, Kyungrak Lee¹, Yantao Kan¹, Inwhee Joe^{1,*}, and Yeonyi Choi²

¹ Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea

² Department of Fire Safety Management, Shinsung University, South Korea
{liyong0912, esilote82, kanyantao, iwjoe}@hanyang.ac.kr,
yychoi@shinsung.ac.kr

Abstract. Research about deep space communications has been studied consistently and briskly for deep space network. But, deep space environment has many problems such as long transmission delay time. To overcome these problems, CCSDS (Consultative Committee for Space Data Systems) standard provides a file transfer protocol, called CFDP), for reliable and efficient data transmission. In this paper, we analysis and compare the performance of four NAK transmission algorithms of CFDP protocol in terms of data transfer rate and the number of PDU. In the result, we can see that although the data delivery time in immediate NAK algorithm is lesser than other mechanisms, but the amount of traffic on the network will be increased due to many retransmissions. In the prompted NAK algorithm and asynchronous NAK algorithm, the data delivery time is increased, but the less retransmission can avoid congestion.

Keywords: CFDP, DTN, Deep Space Communication, NAK.

1 Introduction

Deep space communication is defined as a communication between a device in the earth and a device in other celestial body. Because of unique characteristic of space, deep space communication can't guarantee a stable connection. To overcome this problem, CCSDS (Consultative Committee for Space Data Systems) proposed a new file delivery protocol [1], called CFDP (CCSDS File Delivery Protocol). In CFDP, four retransmission algorithms are proposed for reliable file transmission. First, deferred NAK is retransmission algorithm that retransmits all lost packets after final packet is sent. In immediate NAK algorithm, it sends loss packet immediately when packet loss occur. Prompted NAK and asynchronous NAK are used if bandwidth is restricted [2] [3]. In this paper, we analyzed and compared pros and cons about four NAK algorithms with designed a numerical model.

* Corresponding author.

2 Deep Space Communications

2.1 Delay Tolerant Networking (DTN)

The research on DTN began with the IPN (Interplanetary internet) project of NASA Jet Propulsion Lab. in USA. It aims to deliver a data in harsh network because of long-term transmission delay and frequent disruption of connection [4].

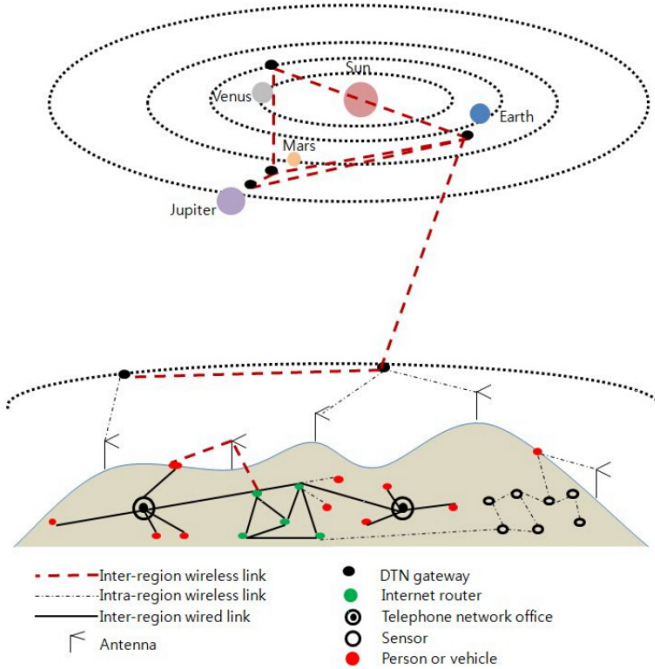


Fig. 1. DTN network

DTN (delay tolerant network) is originated from IPN and contains wireless sensor network, satellite network. This network has unique characteristic such as long-term and variable transmission delay, disruption and fragmentation, high bit error rate, and asynchronous data rate.

2.2 CCSDS File Delivery Protocol (CFDP)

CFDP is file delivery protocol in space environment and its default data unit is PDU, protocol data unit. When a meta PDU is sent, a file delivery is started. A meta PDU contains file name, file size, source ID and destination ID. If packet loss occurs,

destination transmits NAK message to source and requests retransmission. But, in contrast with conventional ARQ algorithm, CFDP uses NAK message not ACK message. There are four NAK algorithms in CFDP: deferred NAK, immediate NAK, prompted NAK, asynchronous NAK.

3 Analysis of CFDP NAK Modes

3.1 Deferred Negative Acknowledgment (DNA)

In deferred NAK, source node retransmits all missed packet after last packet is sent. Figure 2 shows the deferred NAK algorithm. As Figure 6 is shown, source node divides one file data into several PDUs and transmits all PDUs one by one. When a transmission of all PDUs is completed, source node sends EOF message. If packet loss occurs, destination requests retransmission through sending NAK message [5].

3.2 Immediate Negative Acknowledgment (INA)

In immediate NAK, destination node requests retransmission immediately if packet loss occurs. Figure 3 shows the immediate NAK algorithm. Because retransmission is requested immediately, it has an advantage when a transmission delay is short.

3.3 Prompt Negative Acknowledgment (PNA) and Asynchronous Negative Acknowledgment (ANA)

In prompt NAK and asynchronous NAK, an outside event is generated to check a packet loss and inform the result. These algorithms are shown in figure 4 and figure 5. In prompt NAK, source node generates outside event to check a packet loss. If packet loss occurs, destination send NAK message and source node performs retransmission. If packet loss does not occur, retransmission is skipped. In asynchronous NAK, destination node generates outside event message. The rest procedure is same to prompt NAK.

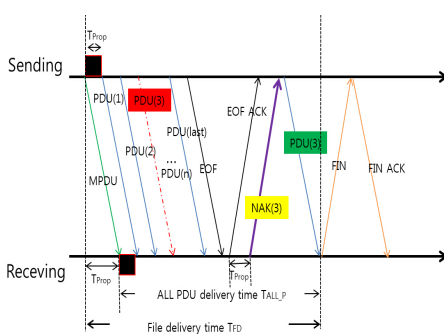


Fig. 2. Deferred NAK

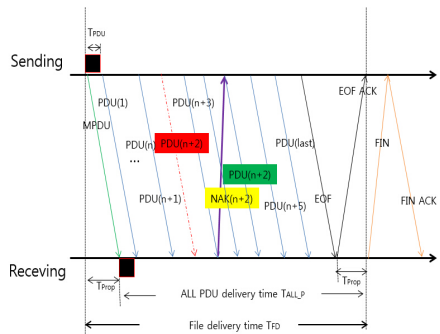


Fig. 3. Immediate NAK

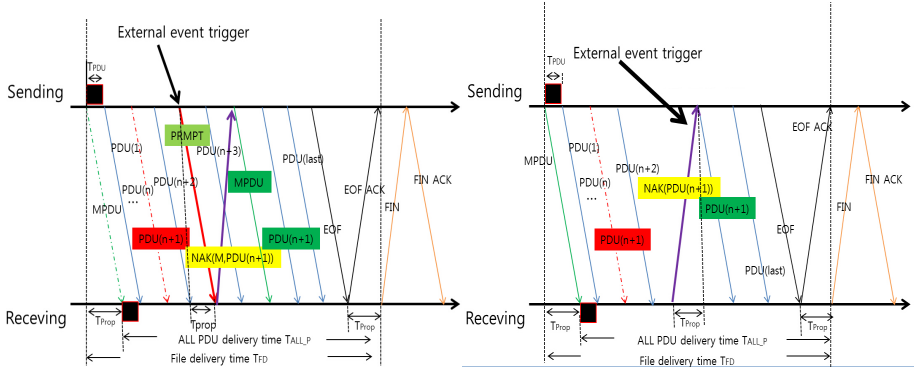


Fig. 4. Prompt NAK

Fig. 5. Asynchronous NAK

4 Performance Evaluation

4.1 Numerical Modeling of Transmission Delay for CFDP NAK Modes

Transmission delay time of EOF process can be obtained the following equation:

$$EOF = 2TPROP + TEOF + TEOF_ACK \tag{1}$$

The expected values of the transmission delay time T of Deferred NAK method $E(T1)$, Prompt NAK method $E(T2)$ and Asynchronous NAK method $E(T3)$:

$$E(T1) = E(NPDU) * TPDU + TPROP + EOF + TNAK + N * PePDU * TPDU \tag{2}$$

$$E(T2) = E(N) * TPDU + 3 * TPROP + TEOF_ACK + TEOF + TNAK + K * Tprompt + N * PePDU * TPDU \tag{3}$$

$$E(T3) = E(N) * TPDU + 3 * TPROP + TEOF_ACK + TEOF + K * TNAK + N * PePDU * TPDU \tag{4}$$

Table 1. Symbols and Definition

Symbol	Definition
LPDU	The length of PDU
LEOF_ACK	The length of EOF_ACK
LNAK	The length of NAK
TPDU	The transmission time of PDU
TEOF_ACK	The transmission time of EOF_ACK
LEOF	The length of EOF
LPROMPT	The length of Prompt
TPROP	One-way propagation delay
TEOF	The transmission time of EOF
R	The transmission rate

4.2 Performance Analysis and Comparison

Scenario 1:

The relationship between the file transfer delay time T of CFDP NAK method and the number of PDU (denoted by N) is shown in Figure 6. T becomes longer as N getting more, but, T starts to decrease when N reaches to 11000. We are able to get an expression $T_{INA} < T_{DNA} < T_{ANA} < T_{PNA}$ through figure 6. Therefore, when BER and PDU are stable, we can get a short transmission delay time by using Immediate NAK method.

Scenario 2:

The relationship between the file transfer delay time T of CFDP NAK method and transmission rate R is shown in Figure 7. T becomes shorter as R getting bigger, but, R has a smaller influence on T when increasing. We are able to get an expression $T_{INA} < T_{DNA} < T_{ANA} < T_{PNA}$ through figure 7. Therefore, when R and PDU are stable, we can get a short transmission delay time by using Immediate NAK method.

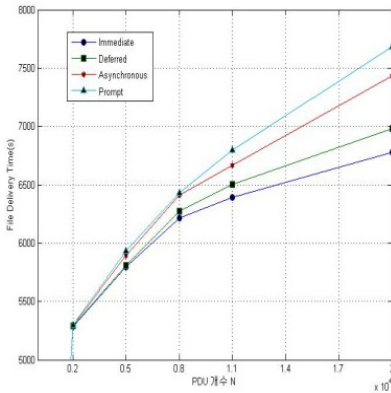


Fig. 6. Tx. delay according to the number of PDU

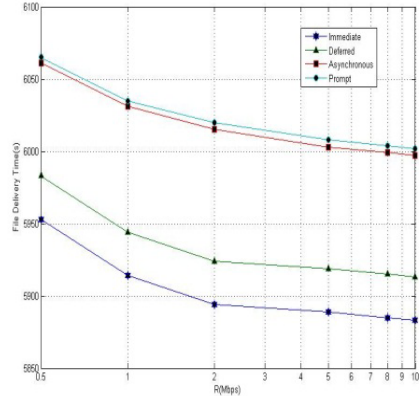


Fig. 7. Tx. delay according to transmission rate

5 Conclusions

In this paper, we have analyzed the four NAK modes of CFDP protocol (deferred NAK, immediate NAK, prompted NAK, asynchronous NAK). The BER of NAK, the data transfer rate, the number of data units, outside the Event was found to impact on the transmission delay time. The deferred NAK is an effective technique for long-range link in cutting off the connection and Immediate NAK scheme is a suitable technique for short-range link. The prompted NAK and asynchronous NAK are used especially when the bandwidth is limited. In the future, we are going to search for a new effective Immediate NAK method to reduce retransmission time.

Acknowledgments. 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 (No.2012R1A1A3012227).

References

1. Communications Operation Procedure 1, Recommendation for Space Data System Standards, CCSDS 232.1-b-1 (2003)
2. CCSDS File Delivery Protocol (CFDP), Part 1: Introduction and Overview, CCSDS 720.1-G-3 (2007)
3. CCSDS File Delivery Protocol (CFDP), Part 2: Implementers Guide, CCSDS 720.1-G-3 (2007)
4. Cerf, V., et al.: Interplanetary Internet (IPN): Architectural Definition (2001)
5. Lee, D.C., Baek, W.: Expected File-Delivery Time of Deferred NAK ARQ in CCSDS File-Delivery Protocol. *IEEE Transactions on Communications* 52(8) (2004)

The Design and Implementation of Mesh-Based DTN for Disaster Situations*

Jaewon Lee¹, Abdul Aziz², Suoning Ma¹, Inwhee Joe^{1**}, and Yeonyi Choi³

¹ Division of Computer Science & Engineering, Hanyang University, South Korea

² Department of Electronics & Computer Engineering, Hanyang University, South Korea

³ Department of Fire Safety Management, Shinsung University, South Korea
{kallontz, iwjoe}@hanyang.ac.kr, azizsheraz@yahoo.com,
msnllly@gmail.com, yychoi@shinsung.ac.kr

Abstract. Disaster is an unpredictable hazard which requires immediate buildup of emergency network to overcome any possible destruction to society or living beings. For such critical issue, yet no international standard is defined for emergency network. Therefore in this paper we have proposed an emergency network which is immune to network disconnection. Proposed network uses delay tolerant networking layer with CFDP, BP & LTP; moreover BATMAN is used as routing protocol and proposed emergency network is emulated using Dummynet. Satisfactory results are obtained in terms of throughput and data transfer time.

Keywords: DTN, emergency network, mesh network, BP, LTP, CFDP.

1 Introduction

In South Korea, emergency network is under consideration since 2003, yet currently there is not any specific emergency network that can be particularly used for disaster situation. After several modifications and adjustments Public security for long term evolution (PS-LTE) [1] is under consideration for emergency networks, whereas it is expected that 3GPP will complete PS-LTE as an international standard by the year 2016. Once the standard is internationally available then performance evaluation such as reliability, security, connectivity, etc. should be done and for better results and further improvements it may take long time to produce reliable emergency network. In addition it may also be possible that LTE and PS-LTE devices encounter communication problems.

Other than traffic accidents, most of the accidents such as fire, etc. occur inside buildings. According to South Korean national emergency management agency, 83%

* 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 (No.2012R1A1A3012227).

** Corresponding author.

of total deaths due to fire disaster has occurred indoors and most the times that fire disaster damages the communication system which makes it difficult for victim to communicate with outer world for help.

Therefore in this paper we are using Wi-Fi based DTN devices to build an emergency network that can be approached by victim easily to ask for help. In addition it is also considered that victim can be effectively located and found by using that emergency network.

2 Related Work

2.1 Delay Tolerant Networking (DTN)

DTN [2] is designed to provide communications in the most unstable and stressed environments, where the network would normally be subjected to frequent and long lasting disruptions and high bit error rates that could severely degrade normal communications. In general, longer delays in TCP/IP will cause communication failure. Conversely in that situation delay tolerant network will store data and wait and when the network is restored DTN completes the communication.

2.2 Bundle Protocol (BP)

BP [3] is one of the components of DTN. In DTN protocol data unit is called "bundle". BP retransmits the data by using a method which is called custody. In order to deal with intermittent network connection, BP uses endpoint identifier for dynamic binding of network address. By using this feature BP can preserve the data without any loss during the network disconnection situation.

2.3 Licklider Transmission Protocol (LTP)

The Licklider Transmission Protocol (LTP) [4] is designed to provide retransmission-based reliability over links characterized by extremely long message round-trip times and/or frequent interruptions in connectivity. Communication in interplanetary space is the most prominent example of this sort of environment, and LTP is principally aimed at supporting "long-haul" reliable transmission over deep-space RF links.

2.4 CCSDS File Delivery Protocol (CFDP)

CFDP [5] is a file transfer protocol that is used for space communication. In the deep space communication this protocol can ensure reliable file transfer. The CFDP has many unique characteristics compared to terrestrial file transfer protocols. CFDP uses FDU (File Delivery Unit) as a data unit for transferring the file. FDU consists of metadata that describes the attributes of the file to be transferred and file itself. FDU can be combined or broken into units called PDU (Protocol Data Unit).

2.5 Dummynet

Dummynet [6] is a program that emulates a network in real time at the MAC layer. It is designed to test networking protocols and it is used for a variety of applications including bandwidth management. It simulates/enforces queue and bandwidth limitations, delays, packet losses, and multipath effects. It also implements various scheduling algorithms. Dummynet can be used on the machine running the user's application, or on external boxes acting as routers or bridges. Dummynet runs on various operating systems, such as Windows, FreeBSD, OS X etc.

2.6 The Better Approach to Mobile Ad-Hoc Networking (B.A.T.M.A.N)

B.A.T.M.A.N [7] is a routing protocol to for multi-hop ad-hoc mesh networks. In this paper, we implement B.A.T.M.A.N routing protocol on Linux platform. Further B.A.T.M.A.N is used for Wi-Fi based emergency network for disaster situation.

3 System Design and Implementation

Requirements for emergency network in disaster situations are different than other networks. Yet network communication requirements in terms of bandwidth and file transfer speed are same. In disaster situation, it is difficult to pre-deploy emergency network, also it is difficult to install communication network immediately [8]. That's why system requirements can be given as follows.

- Network should be able to tackle connection failure
- Message should not be lost even in case communication failure
- Network communication shall be resumed as soon as links are up
- Transmitted message must be received by destination

Protocol stack used for emergency network is given in figure 01.

All data regardless the data type using CFDP should be transmitted as a file. These tasks are performed in CFDP Layer of ION-DTN program [9]. For data compatibility with lower layers, transmitted data is converted using the abstract technique. For lower layers abstract data is converted into suitable information using unit-data transport (UT) adapter. The transformed data is called File Delivery Unit (FDU). In this paper, we use the BP below CFDP layer and CFDP supports four Service classes. Our proposed network stack uses service class 1 for which communication flow is shown in figure 02.

DTN layer under CFDP layer is in charge of communication reliability. Reliable transfer is based on custody transfer mechanism of BP. LTP which is another component of DTN layer prevents loss of any messages even if connection is lost for few hours to number of days, and later LTP resumes message transfer as soon as communication links are up. This method supports the reliable communication method combined with the retransmissions. Finally BATMAN layer below DTN layer is responsible for mesh networking without router. Therefore this system is more suitable for disaster situations as compared to other networks where if communication router is damaged then communication is stopped.

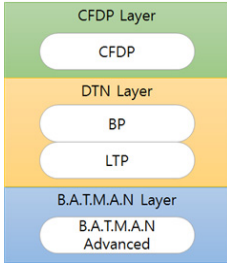


Fig. 1. Proposed network stack

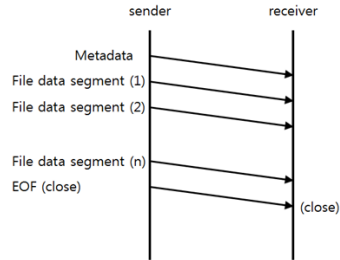


Fig. 2. CFDP unreliable mode (Class 1)

3.1 Network Configuration

Structure of our test-bed is multi hop ad hoc network that is shown in figure 03.

All nodes can communicate in both directions. If the destination node is neighboring node then data is directly sent otherwise data is sent in the direction of the destination node. This network communicates wirelessly and communication range is limited and fixed. In current scenario, we assume that node 4 is control center and node 1 to 3 are sending data to node 4. For a better understanding, the network configuration is shown in figure 04.

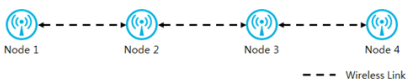


Fig. 3. Structure of test-bed network

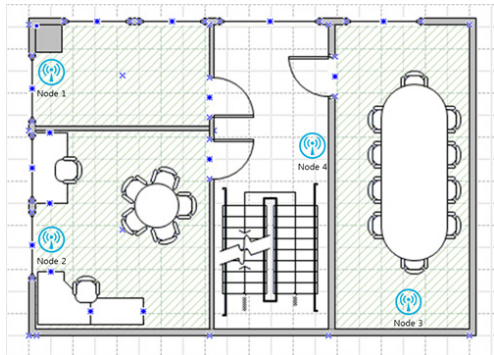


Fig. 4. Map of virtual building where nodes are installed

4 Performance Evaluation

In order to meet the requirements of emergency network, network performance is evaluated for three different file sizes and each experiment is repeated for five time to avoid any errors.

4.1 Transmission Speed Test

We have measured throughput of node 1 to 3 for three different file sizes that are 10MB, 15MB and 100MB. Experimental results show that as number of hops decreases, throughput is getting better as shown in table 01.

Table 1. Throughput result of each node

S.No.	File size	Throughput of node 1	Throughput of node 2	Throughput of node 3
1	10MB	344.83KB/s	454.55KB/s	476.19KB/s
2	10MB	333.33KB/s	476.19KB/s	476.19KB/s
3	10MB	344.83KB/s	454.55KB/s	476.19KB/s
4	10MB	344.83KB/s	454.55KB/s	476.19KB/s
5	10MB	344.83KB/s	434.78KB/s	476.19KB/s
6	15MB	335.57KB/s	471.70KB/s	490.20KB/s
7	15MB	297.62KB/s	387.60KB/s	490.20KB/s
8	15MB	285.71KB/s	409.84KB/s	480.77KB/s
9	15MB	279.33KB/s	480.77KB/s	490.20KB/s
10	15MB	274.73KB/s	467.29KB/s	485.44KB/s
11	100MB	465.12KB/s	483.09KB/s	485.44KB/s
12	100MB	448.43KB/s	480.77KB/s	485.44KB/s
13	100MB	456.62KB/s	483.09KB/s	487.80KB/s
14	100MB	420.17KB/s	478.47KB/s	485.44KB/s
15	100MB	462.96KB/s	480.77KB/s	483.09KB/s

4.2 Reliability Test

The purpose of this experiment was to determine whether the data can be reliably sent in a disaster situations because in disaster situation data loss rate is very high. In order to make a performance comparison of our network protocol stack, we have considered DTN layer with LTP and traditional TCP/IP as shown in figure 05.



Fig. 5. DTN layer with LTP (left side) and DTN layer with TCP/IP (right side)

With the increase of the packet loss ratio (PLR), LTP and TCP file transfer time is increased. Especially with PLR at 5% or less, the transmission speed of TCP is faster than LTP. LTP uses UDP internally. Initially LTP is slower than TCP but with the increase in file size LTP protocol transmission time is less than TCP. When PLR is more than 10%, the transmission speed of LTP is faster than TCP. In particular, the transmission time of TCP was increased very sharply as shown in table 02. Each result is repeated five times and averaged to improve results.

Table 2. File transfer size with LTP and TCP for varying PLR

S.No.	File size	PLR	Average transfer time of LTP	Average transfer time of TCP
1	10MB	1%	29.8s	21.0s
2	10MB	5%	32.2s	22.2s
3	10MB	10%	33.8s	62.6s
4	10MB	20%	64.0s	2183.6s
5	50MB	1%	186.6s	107.4s
6	50MB	5%	131.8s	109.2s
7	50MB	10%	134.8s	328.8s
8	50MB	20%	188.0s	13076.8s
9	100MB	1%	226.6s	216.0s
10	100MB	5%	236.6s	219.2s
11	100MB	10%	263.6s	679.0s
12	100MB	20%	390.4s	24429.2s

We also did network link disconnection tests. We experimented the proposed algorithm with changing link disruption time. Initially, the disruption time is 15 minutes and then keep doubling time until it is reached to 360 minutes, and performed test repeatedly. It was observed that with LTP, network was successful to transfer files after link is restored.

5 Conclusion

In this paper, we have designed and proposed a test bed used with DTN protocol which can be used for disaster situation where network communication status is not good. Generally, when the existing network connection is broken, data can't be transmitted but with our proposed network solution data can be sent successfully after link is restored, which is also confirmed experimentally with the help of test bed. Therefore proposed network architecture can be successfully used in a disaster environment.

References

1. LTE Standards for Public Safety – third generation partnership project (3GPP) view, http://www.3gpp.org/IMG/pdf/2013_05_3gpp_ccw.pdf
2. Delay tolerant networking Architecture, <https://tools.ietf.org/html/rfc4838>
3. Bundle protocol specification, <https://tools.ietf.org/html/rfc5050>
4. Licklider transmission protocol specification, <https://tools.ietf.org/html/rfc5326>
5. CCSDS file delivery protocol, <http://public.ccsds.org/publications/archive/727x0b4.pdf>
6. The Dummynet project, <http://info.iet.unipi.it/~luigi/dummynet/>
7. The better approach to mobile ad-hoc networking, <http://www.open-mesh.org/projects/open-mesh/wiki>
8. Aziz, A., Joe, I., Choi, Y.: A localization algorithm for disaster situation using flying sensor nodes. In: International Conference on Information Science and Applications, Ewha Womans University, Seoul, South Korea (May 2014)
9. ION DTN, <http://ion-dtn.sourceforge.net/>

A Fire Evacuation Guidance System Based on Ubiquitous Sensor Networks

Yeonyi Choi¹, Seokjoon Hong², and Inwhee Joe^{2*}

¹ Department of Fire Safety and Management, Shinsung University,
1 Daehak-ro, Jungmi-myon, Danjin, Chungnam, 343-861, Republic of Korea
yychoi@shinsung.ac.kr

² Department of Electronics and Computer Engineering, College of Engineering, Hanyang
University 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Republic of Korea
{daniel379, iwjoe}@hanyang.ac.kr

Abstract. The purpose of this paper is to investigate the fire analysis and to reduce the evacuation time in high-rise building fire by using the mesh routing protocol based on ubiquitous sensor networks (USN). In this work, in order to reduce the evacuation time for a fire in a high-rise building, an evacuation guidance system is proposed by applying the mesh routing protocol to the building fire. A USN-based evacuation system can quickly detect a fire's origin, optimal path of evacuation involved with the exits and evacuation lights, and the location of evacuees using information collected by the USN system. The results of the adaptive evacuation guidance system from the network simulation showed better performance than the conventional routing method with connected by tree structure.

Keywords: USN, mesh routing, fire evacuation, high-rise building.

1 Introduction

The number of super high-rise buildings is increasing due to their optimal use of limited land area. High-rise buildings have inherent fire risk because many people live in given areas and they are distant from the safety of the ground. The current systems dealing with human safety in evacuation situations are not optimal and may be improved in terms of reliability and evacuation time.

A ubiquitous sensor network (USN) is a system which can detect and manage environmental information, such as temperature, toxin concentration, and humidity, through a wireless sensor network. With developments in wireless communication technology, USN systems are able to provide information on traffic, the climate, and the environment, all in real time. A USN can change sensor network topology by inserting and eliminating sensor nodes [1] and can be ad hoc and self-organized [2, 3]. In order to use this type of system for fire protection in high-rise buildings, new network structures and routing protocols are required.

* Corresponding author.

A USN-based evacuation system can quickly detect a fire’s origin, optimal path of evacuation involved with the exits and evacuation lights, and the location of evacuees using information collected by the USN system.

The purpose of this paper is to investigate the fire danger analysis and the application of mesh routing based on ubiquitous sensor network for fire evacuation guidance system in a high-rise building. We evaluate the performance of proposed adaptive routing algorithm for fire evacuation guidance system through simulation using the OPNET network simulator.

2 USN-Based Fire Evacuation Guidance System

2.1 Structure of a USN-Based Fire Evacuation System

In order for occupants to evacuate safely and efficiently, a USN system must effectively detect fire and smoke. The proposed system is a hierarchical hybrid system composed of an intelligence evacuation guidance system and an urgent evacuation system. If the system is damaged by fire, a hierarchical cluster tree network is changed to a temporary sensor network structure for evacuation-guidance and rescue. Sensor nodes are connected to the evacuation exit lights and exits. After confirming the fire, a disaster sensor network is formed to automatically create a mesh routing structure. The system structure includes a cluster sensor network for each floor of the building, consisting of sensor nodes and a cluster head (CH). The cluster head collects the fire information from the sensor node and then transfers it to the sink. The whole building is designed as a set of separate sensor network structures which can recover from part damage in care of fire heat by selecting an available node. In addition, the disaster prevention center (DPC) is able to connect with outside networks and provide real-time transfer of relevant information not to dispatch the response forces but also to make safe evacuation. Figure 1 shows a schematic of the structure of a fire evacuation system using USN.

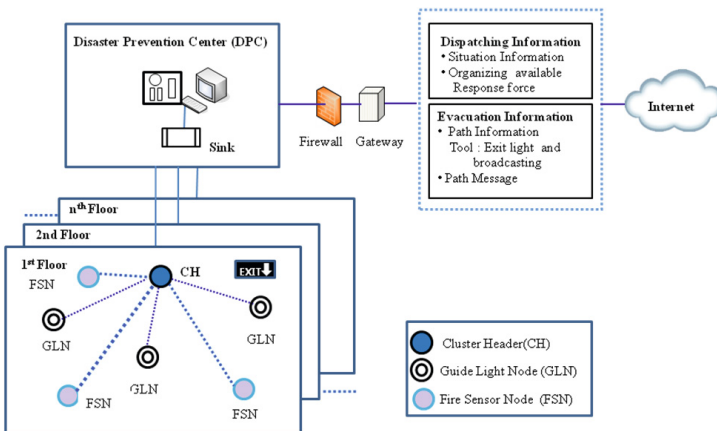


Fig. 1. Structure of a fire evacuation system for high-rise buildings

2.2 Routing Protocol for a Fire Evacuation Guidance System

A routing protocol is a protocol that specifies how routers communicate with each other to disseminate information, allowing them to select routes between any two nodes in a network. Routing is a process of identifying a path from a source to a destination before or after a request for transmission. The routing protocol selected must be achievable and loop-free, and the chosen routes represent the shortest path between the source and the destination [4].

Routing protocols can be used to help efficiently evacuate a building through the use of exit lights, emergency lighting, exits, and disaster planning information. Two kinds of items are needed for fast and efficient evacuations: sensors and intelligent exit lights.

2.2.1 Evacuation Guidance Algorithm Using USN Mesh Routing

The proposed hierarchical mesh routing algorithm, which indicates a safe evacuation route by connecting fire emergency lights and exits in real-time, is more efficient than that of the existing conventional system. The USN-based evacuation routing algorithm is composed of three steps as follows.

(1) Pre-fire - Tree-based structure: the system monitors every disaster-related variable including fire, temperature, flame, smoke density, and gas through each ID possessing sensor networked by wireless RF. In this step, an alarm is raised after a fire is sensed.

(2) Fire occurrence - Mesh structure: when a fire is detected, the hierarchical tree structure is changed to a mesh network structure that is composed of a sensor connected with to evacuation exit lights, emergency lighting and exits.

(3) Optimal evacuation route guidance – Mesh Routing: the system determines the nearest evacuation exit and identifies the origin of the fire. Real-time fire information about the building is applied to determine the optimal evacuation route, and signals, such as emergency and exit lights, TTS (Text to Speech) broadcasting controlled and provided by DPC, are used to guide the evacuation.

2.2.2 Evacuation Route Selection in DPC

In the proposed algorithm, DPC selects optimal route using fire hazard information which received from sensor nodes. And for measuring fire hazard in each node, we suggested that each node has its own hazard value and send packet which include hazard value to DPC through cluster head node periodically. The hazard value can be determined by the degree of danger of fire and it can be one of value from 0 to 9. Here, 0 means no harm and 9 means very dangerous. At first, DPC receives packets from sensor nodes through cluster head periodically. After that, it checks if there is any node has hazard value which is more than zero. If it is true, then make all candidates mesh routes consisted of nodes which have minimum average hazard value. Next, DPC selects shortest route among all candidates routes using location information of each node. We assumed that DPC already has exact location information of each sensor node in the building. Then, DPC send mesh routing information to all nodes through cluster head node so that each node can send packets using the optimal mesh routing path.

Finally, DPC decides the direction of guide light and turn on according to optimal mesh routes which DPC selected. Although DPC already selected optimal mesh routing path, if the average hazard value of the current mesh network is increasing, then it tries to find other optimized mesh routing path based on minimum average hazard value and shortest path. This is because optimal mesh routing path can be changed as fire range increases.

3 Performance Evaluation

In this section, we evaluate the performance of the proposed adaptive routing algorithm for fire evacuation guidance system through simulation using the OPNET network simulator[5]. Figure 2 shows the network model of one floor which includes 13 sensor nodes. Each sensor node uses a routing algorithm which is either static or adaptive routing algorithm. The OPNET simulation parameters for network simulation are as shown in Table 1.

In this simulation, we assumed the node can select a route by using only tree routing algorithm under the static routing algorithm. On the other hand, in the adaptive routing algorithm, if the networks are stable, the node uses tree routing algorithm. But if there is some node failure occurred in the network, the node uses the mesh routing algorithm.

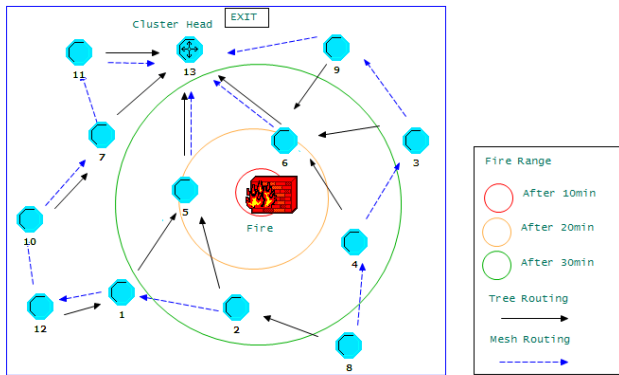


Fig. 2. OPNET network model

Table 1. OPNET simulation parameters

Parameters	Values
Total simulation time	1 hour
Cluster head node	Node 13
End node (Fire sensor node)	Node 1~12
Sensing period of the end node	1 second
Packet size/Data rate	30 bytes/250kbps

For evaluating adaptive routing algorithm variously, we also set that simulation has two scenarios. Scenario 1 is configured for measuring throughput of the cluster head node. We also assume the node fail occurs when nodes are in fire range. Whereas Scenario 2 is configured for measuring average hazard value at the cluster head node which is included in packet received from each sensor node. Each node send packet which includes hazard value. Hazard value can be from 0 to 9. In this scenario, there is no failure of nodes. The *throughput* and the average hazard value can be calculated by formula (1) and (2), respectively. In formula (2), *Hazard(i)* means hazard value which is included in packet received from sensor node *i*.

$$\text{throughput} = (\text{total received packet number} \times \text{packet size}) / \text{time} \quad (1)$$

$$\text{average hazard} = \frac{\sum_{i=1}^{12} \text{Hazard}(i)}{12} \quad (2)$$

Figure 3 shows the result of throughput of cluster head node when network scenario 1 is applied. After 20 minute of simulation time, it decreases more sharply by using static routing algorithm which uses only tree routing than an adaptive routing algorithm. This is because the adaptive routing algorithm select the optimal route for safety using a mesh routing algorithm after the fire occurred whereas static routing algorithm select only shortest path. Figure 4 shows result of the average hazard value in scenario 2. After 15 minute of simulation time, it increases rapidly to almost 7 when the static algorithm is used. While it increases slightly to 1.5 by using an adaptive routing algorithm. This result shows that the proposed adaptive routing algorithm selects more safe path than static routing algorithm as fire range expands.

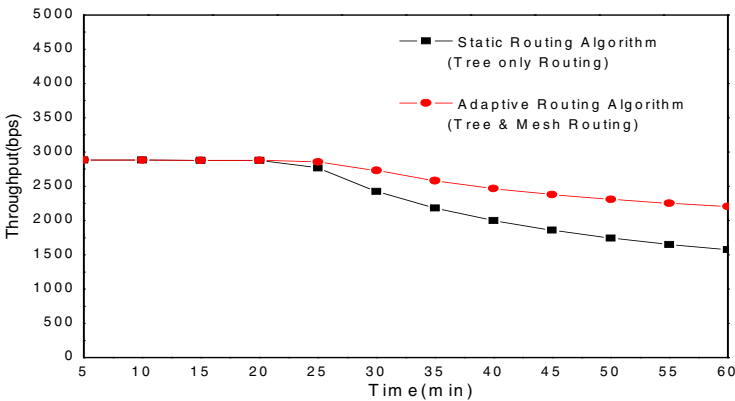


Fig. 3. Throughput of cluster head node in Scenario 1

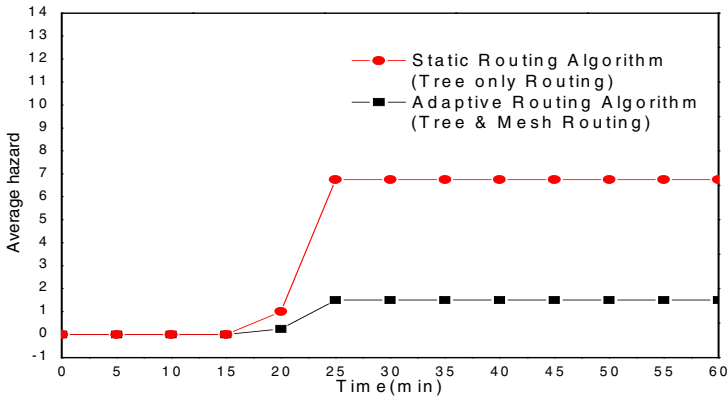


Fig. 4. Average Hazard at cluster head node in Scenario 2

4 Conclusions

In this paper, in order to reduce the evacuation time for a fire in a high-rise building, we propose an evacuation guidance system by applying the mesh routing protocol to the building fire.

In order to analyze the danger of fire and evacuation safety assessment, the fire simulation and evacuation simulation were performed by using fire dynamic simulation and USN based evacuation analysis program. From these results, the adaptive routing method can be used for the most effective method for safety evacuation and minimized fire disaster in the super-highrise building fire by application of the adaptive safety guided light in building of the main passage and rooms.

Acknowledgments. 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 (No.2012R1A1A3012227).

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, S.: A survey on sensor networks. *IEEE Communications Magazine* 40(8), 102–114 (2002)
2. Dressler, F.: A study of self-organization mechanisms in ad hoc and sensor networks. *Computer Communications* 31(13), 3018–3029 (2008)
3. Mills, K.L.: A brief survey of self-organization in wireless sensor networks. *Wireless Communications and Mobile Computing* 7(7), 823–834 (2007)
4. Latiff, L.A., Faisal, N.: Routing protocols in wireless mobile ad hoc network - a review. *IEEE Communications* 2, 600–604 (2003)
5. OPNET Modeler, <http://www.opnet.com/>

Stage Image Control System Using Visual Tracking

Sooyeon Lim¹ and Jaeha Lyu²

¹ Faculty of Liberal Arts, Dongyang University, Yeongju, Korea

² Department of Fine Art, Kyungpook National University, Daegu, Korea

syylim@dyu.ac.kr

Abstract. By the demand of the times, recent performing stages are attempting to change introducing images as a means of communication with the audience and free expression beyond the limits of time and space. It is a method of visualizing the flow of space beyond the physical limits of the stage by combining the motion of performers with space where images are projected and changing static space to dynamic space by making the audience create an illusion that the object is moving. For the implementation of realistic stage, this study is to propose a method of using the whole surrounding space surrounding performers as a removable screen not a non-removable camera. The location information of viewers received from the camera is used as an interaction element and the result of the interaction is used for the location control of the screen where contents are played as well as changes in the contents being played. The proposed method may take temporal and spatial, economic benefits by reducing the number of image playback equipment required to create 3D space and improve the quality of performances by inducing the audience to be immersed.

Keywords: Stage Image Control, Visual tracking, Motor Control.

1 Introduction

Performing arts refer to the art based on the characteristics of various genres. By the demand of the times, recent performing stages are attempting to change introducing images as a means of communication with the audience and free expression beyond the limits of time and space. Stage design is to select space for artistic activities and express the aesthetic formation in that space. By introducing images to the stage for the first time, Erwin Piscator played a leading role in the stage design. Image introduction to stage started in a way of simply projecting images on the wall became the basis for state-of-the-art imaging technique due to the development of equipment and the development of new techniques. And then, stage image played a key role in media based on the stage. Due to its limits of fixing, existing stage background systems have a limit in implementing space and time through stage space. However, using image media enables spatial and temporal changes of stage background according to the flow of the story. Images on the stage became a tool to transfer visual information as well as a story and freed the stage from the constraints of time and space.

Due to the advance in Multimedia, the public gradually began to want new stimuli as they encounter various art fields and producers are producing new expression ways of the field of image design of performing arts while starting to attempt to escape from a fixed perspective for newer visual effects. And many attempts have been made in order to increase the interaction with the audience in various performing arts.

Recent interaction trend is the emergence of a new appreciation scheme that visitors create interaction without any mobile device with the human body as a tool. It is a method of visualizing the flow of space beyond the physical limits of the stage by combining the motion of performers with space where images are projected and changing static space to dynamic space by making the audience create an illusion that the object is moving. The important element of a show is to effectively implement the motion of performers, virtual objects in the virtual space and interaction events.

For the implementation of realistic stage, this study is to propose a method of using the whole surrounding space surrounding performers as a removable screen not a non-removable camera. To this end, a performer tracking stage device composed of a camera, motor, beam projector is made and the proposed algorithm is verified through the experiment. The location information of viewers received from the camera is used as an interaction element and the result of the interaction is used for the location control of the screen where contents are played as well as changes in in the contents being played. The proposed method may take temporal and spatial, economic benefits by reducing the number of image playback equipment required to create 3D space and improve the quality of performances by inducing the audience to be immersed.

2 Backgrounds

2.1 Visual Tracking

Research on computer vision is to build a system to take the place of the human visual ability and especially, the real-time visual tracking system is recently rising as the important research field [1]. Object tracking refers to tracking the motion of the object present in a given image and is a technology applied to various fields such as augmented reality, HCI, navigation, 3D Display, robot control etc. Initial systems recognizing the behavior of performers for interaction recognized the motion of the human hand by attaching the marker to the tip of a finger and using a model-based approach [2] or used optical tracking interface [3], IR tracking [4] for IR beacon attached to the body of a performer etc. On the other hand, recently advanced gesture recognition methods are using several computer vision technologies using cameras among methods using a variety of sensors obtain location information of the viewer. The vision-based system tracks the motion of the human hand or body by using 2D / 3D spatial information obtained from images entered in the camera.

Camera-based tracking system is divided into two kinds depending on the relationship between the camera and the object. One is about a fixed camera and moving object and is mainly used in the field of gesture recognition (surveillance camera, product quality inspection, etc.).The other is research on a moving camera

and moving object and suitable for expressing general natural phenomena but it is difficult to overcome changes in the background due to the motion of the camera.

Of studies for tracking the whole motion of the user and effectively controlling the physical system based on recognized information, the method of recognizing motions [5][6] by extracting a particular color from the image entered through the connected camera has the disadvantages that colors for detecting viewers are limited and it is highly influenced by light. Therefore, it is very important to set a reasonably structured environment because many problems arise due to sensitively changing recognition depending on the ambient environment such as lighting.

If a recent motion recognition sensor is used, tracking can be more free from the influence of light [7][8]. Motion recognition sensors such as Kinect of Microsoft have the advantage of high recognition rate because they can track the information on human joints. However, the Kinect sensor is not suitable for performances using large space such as stage due to its field of view of 57.5° horizontally, 43.5° vertically and the recognition distance of 40cm up to 4m. Therefore, this study uses a camera with wide field of view for object tracking in large space such as a concert hall.

2.2 Step Motor

With the recent increase in demand for 3D printers and robotic devices, Step Motor is studied and developed more actively because of its advantages of lower price and more accurate angle control compared to servo motor. Step Motor can control the spin angle of the motor relatively easily only with simple control circuit and control relatively fixed location without having a separate expensive location sensor. Step Motor can control the spin angle and spin speed by controlling an electrical pulse signal because it is a motor rotating at a predetermined angle by an applied pulse signal. Spin angle is characterized by being proportional to the number of input pulse signals and spin speed being proportional to the frequency of the input pulse signal.

The selection of a motor is determined by startup speed, acceleration and deceleration time, operation speed, positioning time etc. of the motor. This study used shaft type geared mounted 5-phase Step Motor, 5-phase Step Motor driver and motion controller of Autonics [9] by considering the weight and spin speed of equipment to be used in the experiment. Step Motor used in this study is 5-phase hybrid type spinning by 0.72° per 1 pulse. Hybrid type Step Motor can maintain the stop position without relying on mechanical stop and control signal because it has large holding torque when stopped when the power is on. Also, its advantages are as follows; settling time that the motor shaft is stopped by the inertia of the load and inertia of the motor rotor while spinning normally and reversely when the motor reaches the stop location and stops after spinning is short and there is no hunting phenomenon that the motor shaft is stopped while finely spinning normally and reversely when maintaining the stop location. Settling time means the time required until the output signal is entered into the steady state range after the input signal is entered into the system and hunting refers to control vibration caused when not capable of exactly following the setpoint because of delay in the control process in the automatic control system.

Motor driving methods through programming include PWM (Pulse Width Modulation), PID (Proportional Integral Derivative) and so on. Step Motor selected in this study uses PWM control system. PWM is called as pulse width modulation and a modulation scheme of having the constant amplitude of sampling pulse and changing the pulse width depending on the signal to be transmitted and can control the spin angle and spin speed at the pulse rate and number of pulses. For the control of Step Motor by the PWM method, we used the driver, the dedicated drive unit provided by Autonics and dedicated controller controlling this. A motor driver supplies power to the motor according to the phase order of the motor and the dedicated controller controls the spin angle and spin speed etc. of Step Motor.

3 Performer Tracking Stage Background Image System

3.1 System Structure

The purpose of this study is to implement a system of tracking a single moving object by using a fixed camera. The tracking system is largely made up of three steps: Preprocessing process of images received from the camera, process of acquiring the location information by estimating the motion based on difference image technique and contour extraction and process of controlling the motor and image making up the beam projector support based on the estimated coordinate. Figure 1 shows the entire diagram of the stage image control system of using performer tracking proposed in this study.

Space influenced by the motion of the performer is information space where interaction with a virtual object is generated. We control the motor by recognizing and tracking a moving object by decoding the images of a camera received from the webcam. This enables to implement the stage background using the motion information of performers by effectively splitting, analyzing the entire space where the performance takes place.

3.2 Motor Control Using Real-Time Tracking

3.2.1 Performer Tracking Algorithm Using a Camera

By analyzing data through camera sensor scanning, this study obtains the location information of the viewer and uses it as the input information of motor image driving system. The optimization of related algorithm is the required condition because high-performance of processing an image of hundreds of megabytes in real time in the tracking using the camera image processing. We use OpenCV, the library suitable for real-time image processing. OpenCV is the library in the form of C language-based C++ and must use the interface for interlock with processing. Figure 2 shows the control flow diagram of the performer tracking stage image system including image processing.

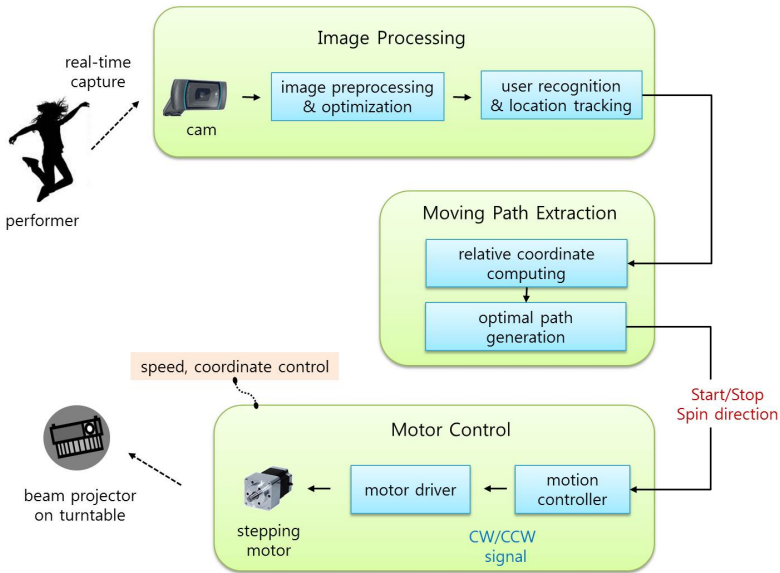


Fig. 1. Diagram of the stage image control system using performer tracking

If acquiring the data sent by a webcam, they go through the process of obtaining an image and processing the image in pc. Since the acquired image is a color image, it is converted to the black and white image for easy image processing and then goes through correction work. In order to remove background noise, the image is corrected and then the background is removed for the tracking of the moving object. In order to remove the background, calculate the difference between the pixel value of the image obtained in the previous frame and the pixel value of the image obtained in the current frame to obtain the difference image. Binarize the difference image and find the contour by using the contour extraction algorithm of OpenCV. If matching this with the current frame, the image recognizing the moving object can be obtained. If motion is detected from the obtained difference image and the size of the detected moving object is greater than threshold, it is determined as a performer and the location information obtained plays spin, stop and running functions in line with Step Motor. The difference image and contour technique are applied at the same time because if using only the difference image technique, it may respond to the subtle motion of different background not people. In addition, using only the contour extraction technique makes it difficult to accurately detect motion because it responds sensitively depending on the brightness of light. The moving lines of a moving object are found to be set to place the central point of the detected object in the center of the camera image. Placing a screen at the point where the viewer can meet contents face to face is done by moving or spinning the motor so that location of the central point of the detected object can be placed in the center of the camera image.

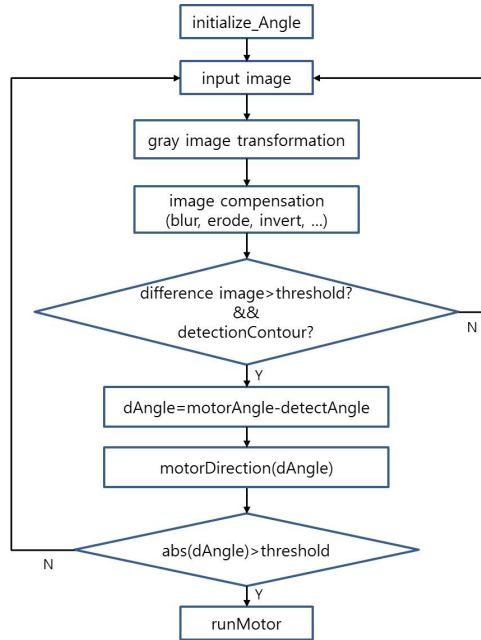


Fig. 2. Control flow diagram of real time tracking stage image system

3.2.2 Motor Control Using the Location Information

If a camera succeeded in tracking the motion information of a moving object, it transmits its location coordinate data to the computer. The computer analyzed the acquired location information and then transmitted it through RS-232 serial communication and 115200bps was used as communication speed. The spin direction and location of the actual motor are controlled by operating motor driver by sending a desired control signal to the motion controller and using the signal. By comparing the relative location of the performer obtained from the camera information and the current location of the motor, the angle difference depending on location difference is obtained and motor is moved. If the movement angle is greater than 180 degrees, it can be moved to the two directions of left and right and the movement angle is also converted to reduce the movement trajectory of the motor (See figure 3).

We could find that the vibration of the motor occurs as the location of Step Motor also responds sensitively according to the fine changes in the location of the moving object while testing the system. We were able to solve this problem by setting critical zone for detected objects. Repeated tests were carried out for motor vibration and as a result, the size of critical zone was set to 1.5 times of the size of the detected object tracked. The change in the location of the moving object within critical zone was ignored and set to be unrelated to change in the motor location (See figure 4).

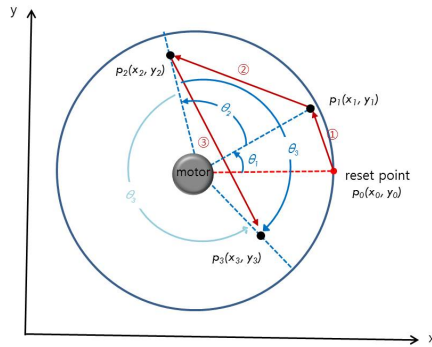


Fig. 3. Optimal trajectory of the motor

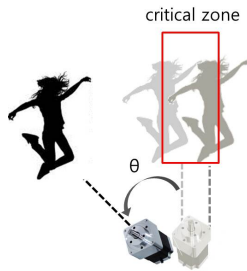


Fig. 4. Setting critical zone for preventing the vibration of the motor

Compared to the fact that the maximum driving speed of Step Motor we used is 8000pps, the speed of the motor was tested through interlock with the camera and as a result, if driving over 125pps, hunting was caused when Step Motor was stopped. This seems to be due to the fact that time delay to pass the location coordinate value of the moving object obtained from the camera image to Step Motor occurs.

4 Conclusion and Future Work

In this study, we were to use the location information of the viewer as an interaction element and implement the result of the interaction through the motion of the screen where contents are played as well as changes in contents. To this end, we proposed the location detection method of a moving object using camera tracking and control method of Step Motor and developed the stage image control system to which the proposed method was applied. As development tools of the proposed system, Processing 2.2.1 and Open CV 2.4.5 were used and image editing tools such as Aftereffect, Photoshop were used for production of contents. The test for the applicability of the developed stage image control system to the stage showed that the location of a performer within the visible range of the camera is tracked successfully and the movement of the screen is smooth. Tracked information was used in the

control of the motor controlling the screen but also used as the feedback information for the audience. We edited RGB images of the performer entered from the camera or made the contents responding to the performer's motion and played them again. As a result, viewers showed an active motion in the theater expressing an interest in the screen moving by being synchronized with their motion and real-time monitoring video receiving feedback. Unlike conventional interactive contents interacting with contents expressed on the wall or inside the screen of the limited equipment, implementation of dynamic image screen means expansion of the exhibition space and infiniteness of expression space. The dynamic image control system proposed in this study makes it possible to produce sensory and interesting space by inducing the performer's natural motion. It also showed that it can be a tool of new artistic media to express three-dimensional space through the two-dimensional plane.

Therefore, we are sure that this study is a useful element in the design of actual stage image for smooth interaction between the audience and the stage. Also, the proposed algorithm will be also used as the development algorithm for effective interaction of physical functional games.

References

1. Wechsler, R.: Artistic considerations in the use of motion tracking with live performers: a practical guide. In: *Performance and Technology: Practices of Virtual Embodiment and Interactivity*. Palgrave Macmillian (2006)
2. Davis, J., Shah, M.: Visual Gesture Recognition. *Vision, Image and Signal Processing* 141(2), 101–106 (1994)
3. Paradiso, J., Flavia, S.: Optical tracking for music and dance performance. *Optical 3-D Measurement Techniques IV*, 11–18 (1997)
4. Salman, A., Khalid, R., Yasir, R., Zahid, B., Usman, K.: On Stage Performer Tracking System. *Advances in Engineering & Technology* 3(2), 65–76 (2012)
5. Manigandan, M., Jackin, I.M.: Wireless vision based mobile robot control using hand gesture recognition through perceptual color space. In: *Advances in Computer Engineering (ACE)*, pp. 95–99 (2010)
6. Kao, S.-T., Yang, Z.-Y., Ho, M.-T.: Design and implementation of a color-based visual tracking control system. In: *2013 CACS International Automatic Control Conference*, pp. 371–376 (2013)
7. Fabian, J., Young, T., Jones, J.C.P., Clayton, G.M.: Integrating the microsoft kinect with simulink: Real-time object tracking example. *J. Mechatronics, IEEE/ASME Transactions on Mechatronics* 19(1), 249–257 (2012)
8. Parzych, M., Dabrowski, A., Cetnarowicz, D.: Aspects of Microsoft Kinect sensor application to servomotor control. *Bulletin of the Polish Academy of Sciences Technical Sciences* 62(3), 595–601 (2014)
9. Autonics Sensors and Controllers, <http://autonics.co.kr>

Real-Time Virtual Lego Brick Manipulation Based on Hand Gesture Recognition

Tran Van Thanh¹, Dongho Kim^{2,*}, and Young-Sik Jeong³

^{1,2} Department of Digital Media, Soongsil University, Seoul, Korea
thanhit08@magiclab.kr, cg@su.ac.kr

³ Department of Multimedia Engineering, Dongguk University, Seoul, Korea
ysjeong@dongguk.edu

Abstract. Recently, with the development of interactive virtual reality technology, many devices have been created to serve the purpose of helping people interact with virtual objects accurately and meticulously. Leap Motion, one of the new devices developed recently, not only supports some basic gestures, but also provides enough information so that developers can define their own gestures. In this paper, we present a 3D simulation which is quite similar to real Lego products using VR technology and Leap Motion. Players can perform simple actions like assembling, changing color, rotating Lego bricks, and zooming in and out. They can also control GUI easily through hand gestures that are analyzed and processed inside the system.

Keywords: virtual reality, gesture recognition, Leap Motion, NUI (natural user interface).

1 Introduction

Over the last a few years, in the field of interaction between people and computers, using natural human senses instead of computer hardware has become extremely popular. Besides of the introduction of new equipment, many new technologies were introduced such as body tracking, hand gesturing, sound recognition, eye tracking, facial recognition, etc. These technologies aim to a new user interface which enables users to employ gestures, sounds, or even eye movement to interact with video games, computers, or other machines. Gesture Interface, which is called Natural User Interface (NUI), is a concept used in a lot of science fiction films where human can use hands to manipulate on screen displays in the space. As examples of the equipment which supports NUI quite accurately, there are Kinect, Smart Gloves, Leap Motion, etc. By using the information obtained from these devices, developers can analyze gestures and give the corresponding commands.

Our simulation is based on the idea of interaction between player and computer, and a simple Lego game is created with the aim of making players feel like playing with real Lego bricks using Leap Motion. This game is aimed for young players, so

* Corresponding author.

all the gestures are defined very easy to do. Leap Motion provides developers with many ready-to-use gestures such as swipe, circle, etc. These gesture definitions are provided with Leap Motion SDK and can be implemented in any application rapidly. However, the build-in gestures are not sufficient to use in specific applications such as our system. In this research, we also define a set of new gestures, including fist gesture, push gesture, and zoom gesture.

This paper is structured as follows: After introduction, related study is presented. Section 3 explains on gesture based virtual Lego manipulation. In section 4, experimental results are presented. Finally, section 5 concludes the paper and suggests some further work.

2 Related Works

2.1 Lego Simulation

A real Lego game consists of colorful interlocking plastic bricks and accompanying arrays of gears, mini-figures, and various miscellaneous parts. Lego bricks can be assembled and connected in many ways, to construct such objects as vehicles, buildings, etc.

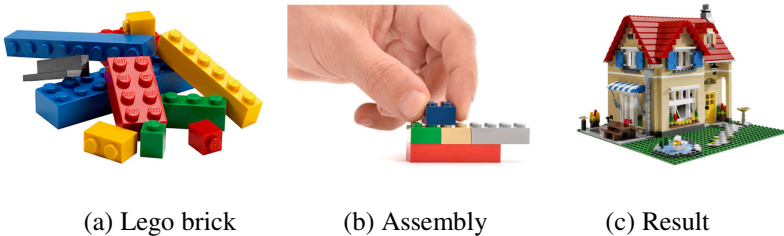


Fig. 1. A real Lego product

Nowadays, there are some applications created using the Lego modeling concept. The LDraw (1995) is free software for modeling Lego creations in 3D. The LDraw file format can be used to store any type of 3D model. Therefore, some virtual modeling programs can use the LDraw Lego part library as construction materials, such as Mike's Lego CAD (1999). Having the same idea with LDraw, the Lego Digital Designer, or LDD, is produced by the Lego Group (2004). The program allows users to build models using virtual Lego bricks, in a computer-aided design like manner.

2.2 Gesture Recognition

In Hongzbe Liu's paper (2013), they have proposed path modification based on NUI. They used Leap Motion as a device which can enable effective use of gestures to perform path operations instead of a mouse and proposed a method of detecting 3D points and modify 3D paths in real-time. D. Bassily (2014) presented a method for interacting with a robot based on gesture and position tracking system with









sub-millimeter accuracy using Leap Motion. In this paper, they discussed about the problem of different coordinate systems between the Leap Motion Controller and the Jaco arm. Similarly, we present a method to convert a position from Leap Motion coordinates to game coordinates when rotating the virtual camera. There have been a few related works in the context of hand gesture recognition using other devices. J. Wu (2013) explored multi-scenario gesture recognition using Kinect. In this research, gestures were recognized by passing hand information through three layers of classifiers: finger counting, finger name collecting, and vector matching. Guan-Feng He (2011) presented a method for detecting hands and fingertips using depth data acquired by PrimeSense 3D camera. In this paper, they could move the mouse and click it by recognizing the shape change of hand contour.

3 Gesture Based Virtual Lego Manipulation

3.1 Gesture Design

Our application primarily aims for children in elementary school or middle school. Therefore, these gestures should be designed easily to remember and perform. Based on experiment combination with other previous researches, we propose a set of new gestures used in our interactive simulation based on the Leap Motion device and the above concepts as presented in Table 1.

Table 1. Gesture List

	Hand	Gesture	Icon	Effect
Leap Motion Default Gestures	Right Hand	Point and move		Move brick
	Right Hand	Point and hold		Place brick
	Right Hand	Circle movement		Rotate brick
	Right Hand	Swipe		Switch menu item
New Gestures	Right Hand	Flat and Push		Select menu item
	Right Hand	Fist		Open/Close Menu
	Left Hand	Fist and move		Rotate camera
	Left Hand & Right Hand	Fist and move		Zoom

With the Point gesture, users can move the current brick by moving their finger. We use the Coordinate system Transformation Algorithm to re-calculate the position of the player's finger in the game coordinate system and apply Hint Brick Generation Algorithm to show a virtual brick as a projection of the current brick on the plane. With the Point and Hold gesture, the brick will be placed when the player remains their index finger for a moment. With the circle gesture, the current brick will rotate by 90 degrees about the y-axis. With the Fist gesture, the player can open or close an in-game menu. With the left hand, users can rotate the game's camera. And the user can use both hands to zoom in or zoom out the game's view. With the Swipe gesture and Flat and Push gesture, a user can switch and choose an item from the in-game menu.

3.2 Coordinate System Transformation Algorithm (CST Algorithm)

Leap Motion and the game camera operate on two different coordinate systems. By Leap Motion's gesture tracking, the object should move in a hand's direction. However, the camera's rotation causes object movement still to follow the original coordinate system direction that's not the same as the hand's moving direction.

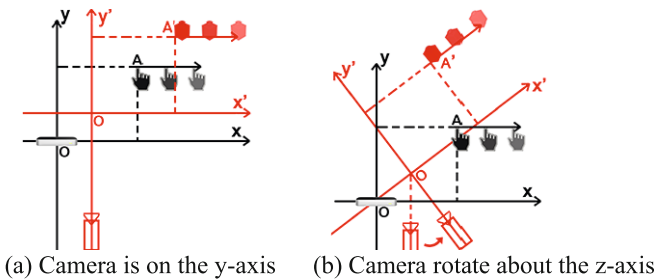


Fig. 2. Hand and object movement

In other Leap Motion applications created by Unity, we can attach Leap Motion Controller object to the main camera as a child object. And then, following the main camera's view direction, Leap Motion Controller's direction will be changed. However, this method has some limitations. First, it is used exclusively in Unity. Secondly, the problem is only solved when we rotate the virtual camera about the y-axis. While rotating the camera about another axis, the problem still occurs. To resolve the problem thoroughly, we propose a method to recalculate the in-game object's position based on the hand's position and the rotation angle of game camera.

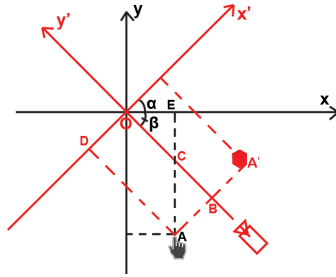


Fig. 3. Coordinate system transformation

From hand’s position A (x, y) and A’ (x’, y’) are in-game object position. Now, we must convert the position of the hand from Leap Motion coordinates to game coordinates. In other words, we have to calculate the length of OB and OD as the point coordinates of A on the game coordinate system. OB and OD are defined as follows:

$ \begin{aligned} OB &= OC + CB \\ &= \frac{OE}{\cos\beta} + CA * \cos\alpha \\ &= \frac{OE}{\cos\beta} + ((EA - EC) * \cos\alpha) \\ &= \frac{OE}{\cos\beta} + ((EA - (OC * \sin\beta)) * \cos\alpha) \\ &= \frac{OE}{\cos\beta} + \left(\left(EA - \left(\frac{OE}{\cos\beta} * \sin\beta \right) \right) * \cos\alpha \right) \\ &= \frac{x}{\cos\beta} + \left(\left(y - \left(\frac{x}{\cos\beta} * \sin\beta \right) \right) * \cos\alpha \right) \end{aligned} $	$OD = AB = CA * \sin\alpha$
--	-----------------------------

Fig. 4. Coordinate system transformation Algorithm

3.3 Hint Brick Generation Algorithm (HBG Algorithm)

In a 3D game, it is difficult to estimate the orthogonal projection of a point onto the floor plane. It is very important for easy assembly of Lego bricks. Therefore, we propose a method to determine the projected point of the current brick’s position on the plane and at which position is able to place the brick. Fig. 5 illustrates the hint of the current brick on the layer. Each layer is a plane that is separated by a distance which is equal to the height of a basic Lego brick.

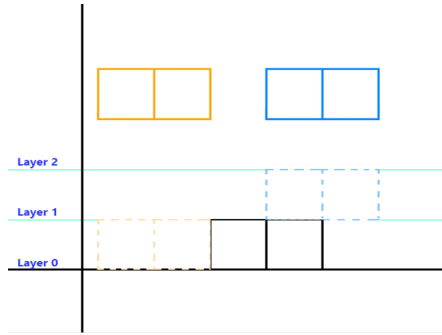


Fig. 5. Hint Generation

The algorithm is described by pseudo code below:

While **true**

 Calculate the **projected point** of the moving brick on the Layer **i** (initial **i** = 0)

 Check **collision** between the projected point and the **placed** bricks

 If it has **collided** → Return the **projected point**.

 Else → Increase Layer index **i**

Display a **virtual brick** at projected point

3.4 Gesture Processing

3.4.1 Pointing Gesture

This gesture is used for moving a brick. This process is implemented as follows:

For **each Fingers** in a hand

 If **Finger** is **index finger** and is **outstretched**

 Calculate **Tip Finger's position** using **CST Algorithm**

 Normalize Tip Finger's position and set **that position** to the current Brick

 If the current Brick or the Tip Finger **has collided** with a placed Brick → **hide** the current Brick and **highlight** the placed brick

3.4.2 Flat and Push gesture

This gesture is used for selecting a menu item. We have the pseudo code below:

Threshold θ , τ

If **Hand** is **left hand** and all **fingers** are **outstretched**

 If **Hand's** direction in the y-axis is **greater** than θ

 If **Hand** is moving in the **opposite direction** from that of the z-axis

 If the **distance** moved is **greater** than τ → **Select** the current menu item

4 Experimental Results

We developed a simple Lego game using Unity 4.6.1 and Leap Motion LP410. In the experiment, we performed a list of gestures to control the menu, the current Lego brick, and the camera with the parameters $\theta = 0.7f$ and $\tau = 60.0f$ that we gave good results.

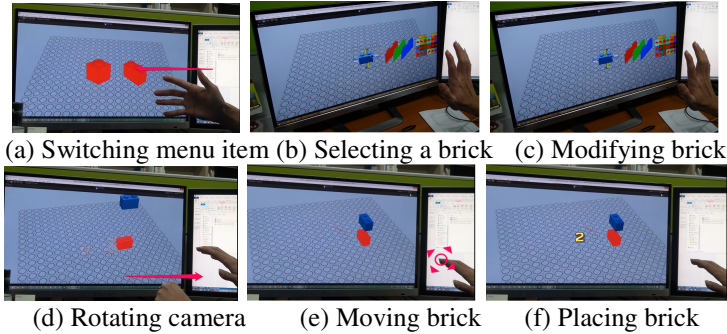


Fig. 6. Experiment result

Firstly, we put the hand in the Leap Motion operating range. When the hand gesture is a swipe gesture and a menu screen is displayed, we can switch between menu items, as shown in Fig 6. (a). Then, we can select a menu item by standing hand up and pushing toward the screen, as shown in Fig 6. (b). We can adjust the properties of a brick such as rotation and color, or switch to another brick, as shown in Fig 6. (c). We can rotate the camera to change the viewing angle by making a fist with the left hand, as shown in Fig 6. (d). Finally, we can move or place a brick according to the position of the index finger, as shown in Fig.6 (e) (f).

The experimental results show that our application can recognize valid gestures and convert them to the commands that will be applied to the menu, the game's camera, or the current brick, etc.

5 Conclusion and Future Work

In this paper, we present a 3D application that simulates a simple virtual Lego game using VR technology and Leap Motion. With our simulation, the players can use their hands to interact with all the game objects in real-time. With the proposed method – Coordinate System Transformation, any 3D application that uses Leap Motion can apply as a solution for converting positions from Leap Motion coordinate to the game coordinate. In the future, our gestures should be validated in not simple games by many people, because we cannot say for now that our gestures are best. We will continue to improve algorithms, remove noisy information, and enhance recognition sensitivity.

Acknowledgments. This work was supported by ICT R&D program of MSIP/IITP. [14-811-12-003, Development of the graphic engine and framework of 3D contents for the creative hands-on science education].

References

- [1] Sabir, K.: The Molecular Control Toolkit: Controlling 3D Molecular Graphics via Gesture and Voice (2013)
- [2] Liu, H.: Gesture-Based NUI Application for Real-Time Path Modification (2013)
- [3] Bassily, D.: Intuitive and Adaptive Robotic Arm Manipulation using the Leap Motion Controller (2014)
- [4] Shen, J.: GPU-Based Real-time Hand Gesture Interaction and Rendering for Volume Datasets using Leap Motion (2014)
- [5] Raheja, J.L.: Tracking of fingertips and centers of palm using Kinect (2011)
- [6] Li, Y.: Multi-scenario Gesture Recognition Using Kinect (2012)
- [7] Wu, J.: Dynamic time warping for gesture-based user identification and authentication with Kinect (2013)
- [8] Marin, G.: Hand gesture recognition with Leap Motion and Kinect devices (2014)
- [9] He, G.-F.: Real-time Gesture Recognition using 3D Depth Camera (2011)
- [10] Bacim, F.: Slice-n-Swipe: A Free-Hand Gesture User Interface for 3D Point Cloud Annotation (2014)
- [11] Leap Motion website, <https://www.leapmotion.com> (accessed on April 2014)

Object Detection Based on Exemplar Object Expression

Yuanyou Wang, Xiaoru Wang^{*}, Junping Du, and Tianming Du

Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing, China
{wangyuanyou3,wxr}@bupt.edu.cn

Abstract. Object detection is a hot research topic in the field of computer vision. The existing algorithms do not take full account of the diversity of features of objects in the same class and the similarity between objects in different classes. To solve the above problems, a method for object detection based on exemplar object expression is proposed in this study. In the proposed algorithm, the concept of multi-feature tree is introduced. By employing the differences and similarities between some features of the objects, the objects in the training set are partitioned into different clusters. Thus, the leaf nodes of multi-feature tree constitute exemplar objects. At last, all the generated exemplar objects are adopted for the expression of object. The information of both the diversity of objects in the same class and the difference of objects in different classes is encoded for the object expression. Thus, the object detector has a satisfactory integration capability for objects in the same class, and a good distinguishing ability for objects in different classes. In this study, the proposed algorithm is compared with the existing object detection algorithms through the experiment on datasets of both PASCAL VOC 2010 and PASCAL VOC 2012. The validity of the proposed approach is proved according to the experimental results.

Keywords: Multi-feature tree, Exemplar Object, Object detection.

1 Introduction

Object detection is a hot research topic in the field of computer vision. Its performance will influence the successive procedures of image classification and object identification. Currently, the factors influencing the performance of object detection are the diversity of objects in the same class and the ambiguity of classes in different classes. The diversity means that the features of the objects in the same class do not show the consistency as expected. Ambiguity reflects the similarity in features of objects in different semantic classes. Hence, the object detection algorithm should have the integration capability for the diversity of objects in the same class, and the distinguishing ability for the ambiguity of objects in different classes.

The existing algorithm models [1,2,3] usually use the features of each class to establish the unique object detector of the class. Because of the diversity of objects in the same class, a single detector is difficult to fit all the objects in the class, thus leading to a

^{*} Corresponding author.

poor detection performance. Literature [4] proposes an object detection method based on deformable parts. The diversity of objects in the same class in aspects such as posture and angle of view is expressed based on the variability of the relative location of the detected parts, and LSVM is used to train the object detector. Nevertheless, since the features that are most discriminative for each class are different, the method using single feature cannot yield satisfactory performance.

With the above analysis, a new approach for object detection based on exemplar object expression is proposed in this study. First, the concept of multi-feature tree is introduced. The differences and similarities between some features of objects are employed to partition the objects in training sets into different clusters in a hierarchical way. Thus, the objects form the exemplar objects on the leaf nodes of the multi-feature tree. These exemplar objects reflect the diversity of features of objects in the same class and the ambiguity of objects in different classes. At last, all the exemplar objects generated with the multi-feature tree are adopted for the object expression. Thus, the information of both the diversity of objects in the same class and the difference of objects in different classes is encoded for the object expression. In the following paragraphs, we train a linear classifier with structured sparsity constraints to select the exemplar with greater expression ability for the last expression instead of all the exemplars.

2 Exemplar Object

Exemplar object means a type of exemplar that can represent the object with independent feature. From the perspective of statistics, the objects in the same class have similar features, while that belong to different classes have different features. In the field of image understanding, due to the influence from the semantic gap, the objects in the same semantic class may have different features, that is, the diversity of objects in the same class. There may also be similar features between objects in different classes, which is the ambiguity of objects in different classes. Furthermore, from the view of multi-feature, the diversity of features of objects in the same class is manifested as the fact that a lot of low-level features of objects are similar, while a few of them are different. Likewise, the ambiguity of objects in different classes refers to that many low-level features are different, and only very few of them are similar.

As a result, the structure of semantic tree is adopted to organize a number of features, and the diversity and ambiguity of object classes are expressed in a hierarchical way. In this study, the tree structure is defined as the multi-feature tree of exemplar object. On the multi-feature tree, the root is the set of all objects in training set, while the leaf nodes are the exemplar objects. From the root of the tree, a low-level visual feature is used to perform clustering on the high-level nodes for each level downward. After the clustering on each level, the object set which belongs to the same class on the high level is partitioned into several subclasses. By conducting clustering on each level based on multi-feature, when reaching the leaf nodes, the object set on root is partitioned into several subclass sets. From the perspective of the multi-feature tree, although objects sharing the same semantics, i.e., in the same class, belong to different leaf nodes, there are more coincident intermediate paths on the path from root to leaf nodes. This indicates that the objects in the same class are similar in major

features, but show some diversity in a few features. On the contrary, the objects having different semantics, that is, in different classes, share fewer ancestor nodes, which means that the objects in different classes are different in most features.

3 Construction of the Multi-feature Tree of Exemplar Objects

There are N objects after the partitioning in the training set, $X = \{x_1, x_2, \dots, x_N\}$, and L different features, $V = \{v_1, \dots, v_L\}$, are selected to construct the multi-feature tree. The construction algorithm of multi-feature tree is shown in the following:

```

Input : Object set : X
        L different types of feature : V = {v1, ..., vL}
Output: The root of tree:root

1 root.set = X
2 Insert root into nowlayer
3 for i from 1 to L
4   for j from 1 to length(nowlayer)
5     cluster[1..K] ← Cluster_EM(nowlayer[j].set, vi)
6     for k from 1 to K
7       node.set ← cluster[k]
8       Insert node into nowlayer[j].son
9       Insert node into newlayer
10    End for
11  End for
12  nowlayer ← newlayer
13 End for

```

By using the algorithm, an L -level multi-feature tree is built. At each level of the multi-feature tree, several clustering results will be derived for the parent nodes at high-level after the clustering of some features, that is, the corresponding child nodes. The mean of class of each clustering is employed as the result of each clustering. Besides, the exemplar objects corresponding to the leaf nodes are expressed as an ordered combination of cluster means of all the nodes along the path from the root to leaf nodes (non-root nodes). Suppose $m(i, j)$ denotes the cluster mean of the j -th cluster node on the i -th level of the multi-feature tree. For a leaf node passing the path, $root \rightarrow l_1 \rightarrow l_2 \rightarrow \dots \rightarrow l_L$ (where $l_i (i=1, \dots, L)$ denotes that the path passes through the l_i -th node of level i), the corresponding exemplar object F is expressed as $(m(1, l_1), m(2, l_2), \dots, m(L, l_L))$. A 3-level multi-feature tree is shown in Fig .1.

4 Object Expression Based on Exemplar Object

Object detection is actually the process of classification on object expression. Therefore, object expression is the core problem for the framework construction of object detection. Currently, high-dimensional eigenvector is mainly used to describe objects in object expression. However, because of the ambiguity of objects in different classes, the minor differences between these objects are concealed in high-dimensional vectors. They are difficult to be discovered by classification algorithms. Therefore, the capability of object detection is unsatisfactory. In this study, a new approach for object expression based on exemplar object is proposed in order to solve this problem. The approach improves the object detection performance greatly.

The leaf nodes of multi-feature tree are regarded as exemplar objects in this study. Meanwhile, an ordered combination of the cluster means along the path from root to leaf node is employed to express the exemplar object. There are multiple exemplar objects in the bank of exemplar objects constructed in this study. In this study, the bank of exemplar objects is denoted as EOB. There are M exemplar objects in the bank, $EOB = \{F_1, F_2, \dots, F_M\}$, where $F_j (j = 1, \dots, M)$ denotes an exemplar object, which is an ordered combination of cluster means of L features, $F_j = (f_{j1}, f_{j2}, \dots, f_{jL})$. For object x_i , L features are extracted and denoted as the way of expression corresponding to the exemplar object, which is $x_i = (x_{i1}, x_{i2}, \dots, x_{iL})$. Then, the feature of each dimension in x_i and the Gaussian similarity of the feature for the dimension corresponding to each exemplar in the bank are calculated, so the expression of the exemplar object can be derived.

$$h_i = (G_s(x_{i1}, f_{11}), \dots, G_s(x_{iL}, f_{1L}), G_s(x_{i1}, f_{21}), \dots, G_s(x_{iL}, f_{2L}), \dots, G_s(x_{i1}, f_{M1}), \dots, G_s(x_{iL}, f_{ML})) \tag{1}$$

$$\text{Where } G_s(x_i, x_j) = \exp \left\{ \frac{-\|x_i - x_j\|^2}{\delta^2} \right\}.$$

As shown in Eq. 1, it is expressed by the similarity of the features between an object and all exemplar objects. In this type of expression, the similarity between

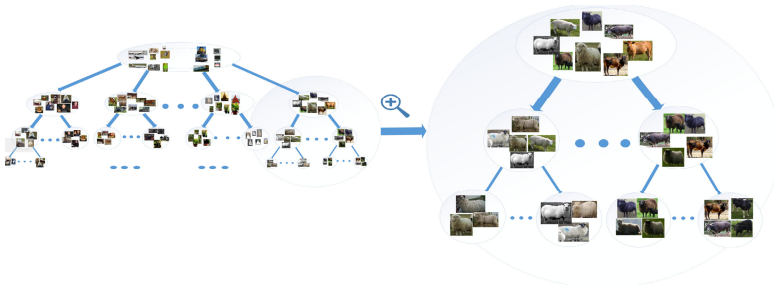


Fig. 1. A 3-level multi-feature tree, and HOG feature, HSV color feature and textural feature are used to construct this multi-feature tree

features of different exemplar objects in the same class and the target object is calculated, so the diversity information of objects in the same class is fully reflected in the object expression. Meanwhile, the difference between features of the target object and exemplar objects in other classes is also calculated. In this way, the information of the differences between the objects is encoded into the expression of objects more directly. Next, the object expression which can fully reflect the information of the diversity of objects in the same class and the difference of objects in different classes is adopted to train the object detector. Thus, the object detector possesses both the integration capability for objects in the same class and the distinguishing ability for objects in different classes.

5 Object Expression for Structure Regularized Learning Based on Exemplar Object

In this section, the above object expression is applied in the learning process of object detector. Suppose $E = \{e_1, \dots, e_N\} \in R^{D \times N}$ denotes the expressions of N training objects based on exemplar object, and the class label of training object is denoted by $Y = (y_1, y_2, \dots, y_N) \in \{1, \dots, K\}^N$. In this study, linear classifier is used to construct the object detector under the classification strategy of 1-versus-all.

$$y = \arg \max_{k \in \{1, \dots, K\}} \beta_k^T e \quad (2)$$

Where y denotes the prediction result of object detection. $\beta_k \in R^{D \times 1}, k = 1, \dots, K$ denotes the detector parameter obtained with the method of structure regularized learning.

In the process of object expression based on exemplar object, the Gaussian similarity between the target object and all exemplar objects is calculated. Then, with the increasing of object classes, the amount of exemplar will increase correspondingly. Meanwhile, the dimension of object expression, D , will keep increasing. To deal with this problem, effective sparsification is necessary to select the exemplar with greater expression ability for the last expression instead of all the exemplars. Hence, the method of structure regularized learning is used to train the detector parameter β_k

$$\min_{\beta_k} \lambda R(\beta_k) + \frac{1}{N} \sum_{i=1}^N L(\beta_k^T; e_i, y_i^k), \forall k \quad (3)$$

$$\text{Where } y_i^k = \begin{cases} 1, & y_i = k \\ -1, & \text{else} \end{cases}$$

λ denotes the balance parameter, which can be determined with cross validation. $L(\cdot)$ denotes the loss function, where the widely used Log loss is selected.

$$L(\beta_k^T; e_i, y_i^k) = \log(1 / (\frac{1}{Z} \exp(\frac{1}{2} y_i^k \times \beta_k^T e_i))) \quad (4)$$

$R(\beta_k)$ denotes the regularization terms, considering that multiple exemplar objects may belong to the same class, and there may be some consistency of prediction in these exemplar objects in the same class. Therefore, the positive form of $\ell_1 / \ell_2 + \ell_1$ is employed to realize the sparsification. Thus,

$$R(\beta_k) = \|\beta_k\|_{1,2} + \lambda' \|\beta_k\|_1 \quad (5)$$

Where $\|\beta_k\|_{1,2} = \sum_{j=1}^K \|\beta_k^j\|_2$, β_k^j denotes the parameters corresponding to the exemplar objects in the j -th class, and λ' denotes the balance parameter which can be determined by cross validation.

Especially, the coordinate descent algorithm [7] is used for the optimization issue in Eq. 3, and the parameter β_k is obtained. Finally, Eq. 2 is adopted for the detection of target object.

6 Experiment

6.1 Data Set and Evaluation Standard

The datasets of PASCAL VOC 2010 and 2012 are used for the object detection experiment. From the images of the two databases, 2000 objects are extracted, and two datasets are divided into two subsets for training and test, respectively. Meanwhile, the metrics of average precision (AP) and mean average precision (mAP) are adopted to evaluate the performance of the algorithm.

6.2 Experiment Configurations

EOB: the method for object detection based on multi-feature exemplar object proposed in this study is called EOB. In this experiment, HOG[5] feature, HSV color feature and textural feature are used to construct the multi-feature tree, respectively, from the root to leaf nodes in the tree. On each level of the multi-feature tree, the EM algorithm is employed for clustering. Some exemplar objects corresponding to boat, sheep and TV obtained from VOC2010 are shown in Fig. 2.

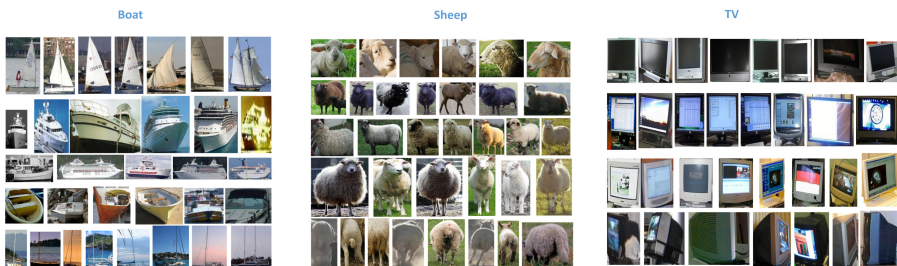


Fig. 2. Some exemplar objects corresponding to boat, sheep and TV obtained from VOC2010. Each row shows some objects of one exemplar object.

DPM: the object detection method based on parts proposed in literature [4] is called DPM. In the experiments, HOG [5] feature is used to extract the features of object, and LSVM[4] is used to train the object detector.

6.3 Experimental Results and Performance Analysis

In the experiment, the performances of object detection of EOB and DPM are compared. The object classification results on the datasets of VOC 2012 and VOC 2010 are shown in Fig. 3 and Fig. 4, respectively.

It can be seen from the experimental results that the performance of the proposed EOB method is obviously superior to that of DPM. The object detection method of DPM is based on parts. The method treats the target object for detection as being composed by a root and several parts. It can be seen from the figure that since only HOG features are adopted for the description of object, the detection is effective for the objects which have clear unique variations inside, such as person and bike. However, for other objects which have a uniform internal distribution and no regular variations, such as bird, sheep and cow, the classification performance decreases greatly. Performing object classification with single feature will not express all the feature of the object, so the performance is poor. However, the proposed EOB method employs the structure of multi-feature tree, so that multiple features are organized to express the diversity of objects in the same class and the ambiguity of objects in different classes. Therefore, the performance of EOB method based on multiple features is superior to that of DPM method based on a single feature. On VOC 2010 dataset, the average accuracy of classification with EOB is 42.99%, while that of DPM is 39.57%. On VOC 2012 dataset, the average accuracy of classification with EOB is 40.46%, while that of DPM is 33.85%. Moreover, for both the two datasets, EOB achieves satisfactory performance for all the 20 object classes.

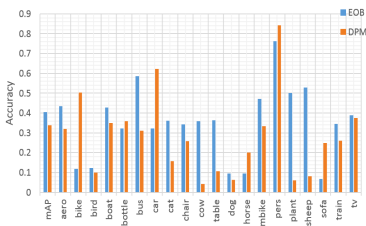


Fig. 3. Object classification results on the datasets of VOC 2012

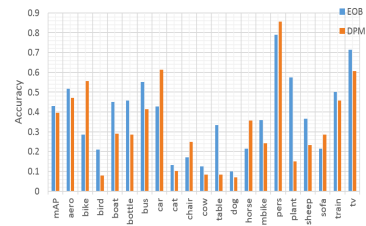


Fig. 4. Object classification results on the datasets of VOC 2010

And the object detection algorithm should be equipped with the ability to detect objects with local loss. However, DPM method is a model based on the scoring on parts. In this detection model, partial loss will decrease the final score greatly, and the ability to identify the objects will decrease sharply. In contrast, EOB takes the overall feature of objects into consideration, and employs multiple features for description.

Hence, the identification ability for objects with partial loss is assured to a great extent for EOB. In order to validate the superiority of EOB method under this circumstance, 100 objects with partial loss are picked up from VOC 2010 and VOC 2012. The approaches of EOB and DPM are adopted for the detection on these objects, and the mPA is calculated. On VOC 2010 dataset, the mPA of EOB is 40%, while that for DPM is 22%. On VOC 2012, it is 31% for EOB and 17% for DPM. It can be seen from the above comparison results that the performance of EOB algorithm is obviously better than that of DPM with respect to the detection ability for objects with partial loss.

7 Conclusion

In this study, an object detection method based on exemplar object expression is proposed. In this method, the multi-feature tree is employed to organize multiple features to form the exemplar object. It can well reflect the similarity of objects in the same class in multiple features and the difference of objects in a few features. In the last step, all the exemplar objects obtained are used for object expression, and the information of diversity of objects in the same class and difference of objects in different classes is successfully encoded into the object expression. Using this method, object detector possesses satisfactory integration capability for objects in the same class and distinguishing ability for objects in different classes. By comparing the proposed method with the existing algorithms on the datasets of PASCAL VOC 2010 and PASCAL VOC 2012, it is proved that the method has good performance for object detection.

In the future, more features will be selected to improve the capability of object detection effectively.

Acknowledgements. This research study was supported by the National Basic Research Program of China(973 Program) 2012CB821200 (2012CB821206), the National Natural Science Foundation of China (61320106006),the Fundamental Research Funds for the Central Universities(No.2013RC0306).

References

1. Song, Z., Chen, Q., Huang, Z., Hua, Y., Yan, S.: Contextualizing object detection and classification. In: CVPR (2011)
2. Chen, Q., Song, Z., Hua, Y., Huang, Z., Yan, S.: Hierarchical matching with side information for image classification. In: CVPR (2012)
3. Russakovsky, O., Lin, Y., Yu, K., Fei-Fei, L.: Object-centric spatial pooling for image classification. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part II. LNCS, vol. 7573, pp. 1–15. Springer, Heidelberg (2012)
4. Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D.: Object Detection with Discriminatively Trained Part-Based Models. TPAMI (2010)

5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR (2005)
6. Dai, J., Yan, S., Tang, X., Kwok, J.T.: Locally adaptive classification piloted by uncertainty. In: ICML (2006)
7. Li, L., Su, H., Xing, E., Fei-Fei, L.: ObjectBank: a high-level image representation for scene classification & semantic feature sparsification. In: Proceedings of the Neural Information Processing Systems, Vancouver, Canada (2010)

An Algorithm for Image Classification Based on Semantic Transfer Learning

Tianming Du, Xiaoru Wang^{*}, Junping Du, and Yuanyou Wang

Beijing Key Laboratory of Intelligent Telecommunication Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing, China
{mercedes1993, wxr}@bupt.edu.cn

Abstract. Images are a type of complex multimedia data, which contains a wealth of semantic information. Low-level visual feature of the image is usually applied to image expression and classifier learning in the existing algorithms of image classification. Classification performance with high quality cannot be obtained with this type of algorithms due to the semantic gap as the low-level visual feature does not contain semantics. In this study, a classification method based on transfer learning is proposed by synthesizing multiple characteristics in modal space to realize the image classification. Semantic features in the text space are transferred into image space with the proposed method, and then the text semantics relevant to the image are used to classify the images. The experimental result demonstrates the validity of the proposed method. The accuracy has been significantly promoted in comparison with other classification methods based on transfer learning.

Keywords: Transfer Learning, pLSA, Entropy, Image Classification.

1 Introduction

Image classification refers to that computer automatically assigns one or more than one class label to the image based on the content of the image. In the existing algorithms for image classification, low-level visual features of the image is usually first extracted. Then a mapping relationship between the low-level visual features and high-level semantic class is established with machine learning techniques. The classifiers are trained to realize the image classification with the established mapping relationship. Since the low-level visual features do not contain semantics, the association established extremely depends on the training set. When the test image is beyond the range of the training set, the classification performance sharply decreases. In recent years, researchers are focusing on how to classify the images using the features in multi-modal space. With transfer learning, knowledge can be transferred among different fields, tasks and distributions. Therefore, the idea of transfer learning can be used in the multi-modal image classification. The multi-modal information is comprehensively used in image classification based on the transfer learning. Examples

^{*} Corresponding author.

of multi-modal spaces can be synthesized to classify the images by sharing the association points of the same semantics via the multi-modal feature space.

Currently, the main algorithms for transfer learning are shown as follows. The first is the multitask learning [1]. The second is the transfer learning based on the distributions of different data. The problem of different data distributions between different data were resolved by transfer learning in literature [2-3]. In image classification, the two transfer learning methods introduced above can only be used to transfer the knowledge in the same feature space. The semantics in the text space cannot be transferred into the image space. The third is the transfer learning across different feature spaces. The difference between the above two methods is that the transfer is not confined to the same feature space any more. Dai [4] proposed the TL-Risk framework that could be used in the image classification. Then knowledge transfer between different feature spaces was realized. Though image and text modal are used in the TL-Risk, translator is simply used to associate the text feature with the low-level feature in the framework. Since the low-level visual feature does not contain semantics, the framework is still influenced by the semantic gap.

An image classification framework based on the semantic transfer learning is proposed in this study through a deep research on transfer learning and image classification. Features in multi-modal space are synthesized to realize the image classification. The semantic features in the text space are transferred into the image space with the proposed method. Images are classified with the text semantics relevant to them. In the transfer process, two co-occurrence probabilities are calculated. They are the co-occurrence probability between the text feature and image label and that between the former and low-level visual feature of the image. The semantic information transferred into the image space is closely related to the image via the two co-occurrence probabilities. Thus, the validity of the classification is proved. In particular, the entropy of the visual feature distribution of the test image is calculated during classification. It can adaptively adjust the parameters in the model with the described co-occurrence probabilities, and the accuracy of classification is further promoted. The problem in literature [4] that the visual features with the same distribution cannot be effectively classified in the framework has been resolved with the method, and the classification accuracy is promoted.

2 Basic Idea

2.1 TLR-Fusion Framework

The method proposed in this study is called TLR-Fusion. (Translated Learning via Risk Minimization and Semantic Fusion) In the TLR-Fusion framework, Markov chain and risk-minimizing model are used together with the framework of semantic image classification. The algorithm flow chart is shown as Figure 1, which is divided into training process and test process. In the training process, each class of training data contains images, image labels and the related texts. For a class, the low-level visual feature of the image, image label and text feature are extracted. With the co-occurrence relationship among them, the association of the probability between the features as well as the feature and class can be obtained, and the model of classification can be induced.

The training process will be discussed in detail in section 2.1.2. In the test process, for the test image, pLSA [5] (probability Latent Semantic Analysis) is used to extract them, which is used to establish the relationship between the image feature and image example. The model of the image can be established with the obtained relationship. Moreover, entropy of the distribution of the image visual feature is calculated to train the parameters of the class model dynamically. Finally, the trained class model which has the closest model divergence to the model of test image is found, and the class of the corresponding trained model is the classification result. The classification process will be discussed in detail in section 2.1.3.

Like the TL-Risk [4], Markov chain and risk-minimizing estimation model are used in the TLR-Fusion, which is briefly introduced as follows.

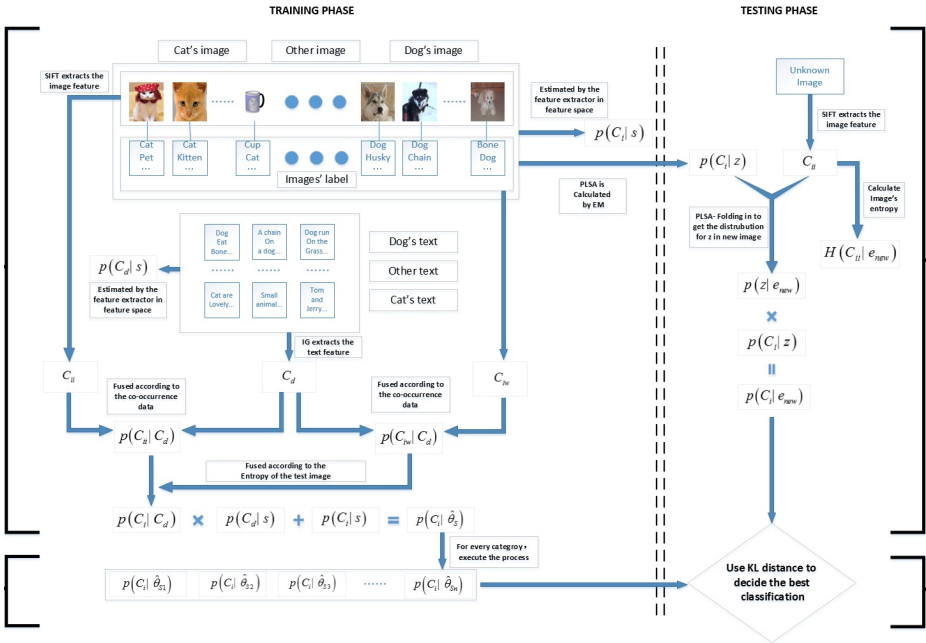


Fig. 1. Flow Diagram

2.1.1 For the TL-Risk Framework

In the TL-Risk framework, the process of image classification is based on two Markov chains, which are $\hat{\theta}_s \rightarrow s \rightarrow c_i \rightarrow e_i \rightarrow \hat{\theta}_{e_i}$ and $\hat{\theta}_s \rightarrow s \rightarrow c_d \rightarrow c_i \rightarrow e_i \rightarrow \hat{\theta}_{e_i}$, where d denotes the text space, and i is the image space. c denotes the feature, and e is the example. Moreover, $\hat{\theta}_s$ is the model of image class, and s is the class. c_i is the image feature, and c_d is the text feature. e_i represents the image model is denoted by $\hat{\theta}_{e_i}$. It should be noted that the parameters of model can be described by the

probability distribution of the image feature c_i with the model. The model of the image class $\hat{\theta}_s$ and that of the image example $\hat{\theta}_{e_i}$ are obtained with the TL-Risk framework. The risk is measured by the divergences between $\hat{\theta}_s$ and $\hat{\theta}_{e_i}$. For the text image, its model and all the trained class models with the TL-Risk is calculated. The trained model with the minimum risk is chosen as the class model, and the corresponding class is the classification result. The calculation of the class model is shown as follows.

$$p(c_i|\hat{\theta}_s) = \sum_{c_d} \sum_{s \in S} p(c_i|c_d) p(c_d|s) p(s|\hat{\theta}_s) + \lambda \sum_{s \in S} p(c_i|s) p(s|\hat{\theta}_s) \quad (1)$$

In the TL-Risk framework, the following formula is used to estimate the model of test image.

$$p(c_i|\hat{\theta}_{e_i}) = \sum_{E_i} p(c_i|e_i) p(e_i|\hat{\theta}_{e_i}) \quad (2)$$

The divergences between the models can be expressed as Kullback-Leibler divergence (KL divergence).

The problem of TL-Risk is that when establishing the class model, the extracted image feature c_i does not contain semantic feature. When the distributions of image feature of two classes are similar, but the semantics are different, the class models cannot be distinguished with TL-Risk.

What is different from TL-Risk is that three improvements have been made in the TLR-Fusion framework.

- a. Semantic information is added to the image feature c_i . $c_i = \{c_i, c_{i_w}\}$ and c_i are the visual features of the image, and c_{i_w} denotes the semantic information of the image, which can be obtained from the image label.
- b. The pLSA is used to obtain the theme z_i . The relationship between the image feature c_i and image example e_i is established in the trained image model.
- c. In the classification process, the entropy $H(c_i|e_i)$ of the visual feature of the test image is used to adaptively adjust the parameter $p(c_i|\hat{\theta}_s)$ of the class model

2.1.2 Process of Establishing the Class Model

First In the TLR-Fusion framework, formula (1) is used to calculate the class model. It should be noted that c_i contains the low-level feature c_{i_v} of the image and the image label c_{i_w} .

In the first part on the right side of formula (1), the probability distribution of the image feature c_i with the model $\hat{\theta}_s$ is calculated. c_i is transferred from the feature in text space c_d . In the second part, the probability distribution of the image feature c_i with the model $\hat{\theta}_s$ is directly calculated. c_i is extracted from the image space

instead of being transferred from the text space. Moreover, parameter λ is used to adjust the influence on the establishment of class model brought by the image feature c_i which is transferred from the text space and the one c_i directly extracted in the image space.

Where $p(c_d | s')$ can be estimated with the training sample in the text space. $p(c_i | s')$ can be estimated with the training sample in the image space. When $s = s'$, $p(s | \hat{\theta}_s)$ can be estimated as 1, otherwise 0. $p(c_i | c_d)$ is estimated with the co-occurrence data of the text and image feature. The forms of co-occurrence data are shown as follows: $p(c_i, c_d)$ is the co-occurrence of image feature and text feature; $p(c_i, e_d)$ is the co-occurrence of the image feature and text example; $p(e_i, c_d)$ is the co-occurrence of the image example and text feature; $p(e_i, e_d)$ is the co-occurrence of the image example and text example. First, the co-occurrence probability of the image feature and text feature needs to be estimated with the following three methods:

$$p(c_i, c_d) = \sum_{E_d} p(c_i, e_d) p(c_d | e_d) \tag{3}$$

$$p(c_i, c_d) = \sum_{E_i} p(c_d, e_i) p(c_i | e_i) \tag{4}$$

$$p(c_i, c_d) = \sum_{E_i} \sum_{E_d} p(c_i, c_d) p(y_d | x_d) p(y_i | x_i) \tag{5}$$

Where the image feature c_i contains the image feature c_{i_i} and label c_{i_w} . The co-occurrence probabilities $p(c_{i_w}, c_d)$ and $p(c_{i_i}, c_d)$ between them and the text feature are calculated, respectively. Then the conditional probability distributions of c_{i_i} and c_{i_w} under the text feature c_d are calculated through formula (6) and (7).

$$p(c_{i_i} | c_d) = p(c_{i_i}, c_d) / \sum_{c_{i_i}} p(c_{i_i}, c_d) \tag{6}$$

$$p(c_{i_w} | c_d) = p(c_{i_w}, c_d) / \sum_{c_{i_w}} p(c_{i_w}, c_d) \tag{7}$$

The results of formula (6) and (7) are combined using the following formula. That is to say, the probability distributions of c_{i_i} and c_{i_w} under c_d are used to calculate $p(c_i | c_d)$:

$$p(c_i | c_d) = \alpha_i p(c_{i_i} | c_d) + (1 - \alpha_i) \sum_{m=1}^k p(\text{fit}_m(c_{i_i}) | c_d) / k \tag{8}$$

In the formula, the fact that the quantities of c_{i_i} and c_{i_w} are different has been considered. c_{i_i} corresponds to c_i one by one. α_i in the first term on the right side of the formula is the weight, which is used to adjust the influence of the image feature and

text feature. The remaining part of the first term denotes the conditional probability distribution of the image feature c_{i_i} at the text feature c_d . In the second term, $fit_m(c_{i_i})$ denotes the m -th text feature word that is relevant to the visual feature word c_{i_i} . We choose the first k elements of c_{i_w} which is relevant to c_{i_i} . ($p(c_{i_w} | c_{i_i})$ can be used to measure the correlation), and the average probabilities of transferring them from the text feature space are calculated. The obtained result is treated the second term to show the influence of text word.

In particular, the experiment shows that the larger the entropy of the distribution of image visual feature, the worse the effect of image classification using the visual feature, and the better the effect of image classification using the image label will be. Considering that different test images have different distribution of visual feature word, the weight α_i is defined as follows:

$$\alpha_i = \begin{cases} 1 & H(c_{i_i} | e_i) \leq k_2 \\ \exp(k_2 - H(c_{i_i} | e_i)) & H(c_{i_i} | e_i) > k_2 \end{cases} \quad (9)$$

where k_2 is the empirical value. $H(c_{i_i} | e_i)$ is the entropy of the distribution of c_{i_i} under the test example e_i . For the test image, when the entropy is large, and the visual features are difficult to distinguish, the influence of c_{i_w} in the transfer process is enhanced. The images of different classes can be effectively distinguished with label c_{i_w} . Parameters of the model are adaptively adjusted with the entropy, and the classification accuracy is promoted.

2.1.3 Classification Process

Firstly, the model of test image is calculated with the following formula:

$$p(c_i | \hat{\theta}_{e_i}) = \sum_{E_i} \sum_{z_i \in Z} p(c_i | z_i) p(z_i | e_i) p(e_i | \hat{\theta}_{e_i}) \quad (10)$$

In the formula above, when $e_i = e_i$, $p(e_i | \hat{\theta}_{e_i})$ is 1, otherwise 0. $p(c_i | z_i)$ and $p(z_i | e_i)$ are learned with the pLSA model [5].

Then z_i of the pLSA and $p(c_i | z_i)$ are obtained with the EM algorithm. For the text image e_{new} , $p(z_i | e_{new})$ is obtained with the trained parameter $p(c_i | z_i)$, distribution of the visual feature of the test image e_{new} and the folding-in algorithm [6]. Then the parameters are substituted into the formula above, and the calculation of the model of test image is accomplished.

In the test process, the class model that has the minimum divergence to the model of test image is found. The model of test image is obtained with the method introduced in 2.1.3. Moreover, the class of the corresponding class model is taken as the classification result of the test image.

3 Experiment

3.1 Dataset

In this study, image dataset of Visual Synset [7] was used. The images in the Visual Synset were taken as the image set, and the corresponding label dataset as the text set.

In the image set, bag-of-words (BoW) model was used to extract the features and express the images. The SIFT descriptor [8] was taken as c_{i_i} in the model. Visual vocabulary was formed through K-means. Then the image expression was realized with the visual vocabulary. Similarly, for the text set, the text words are used to establish text vocabulary, with which the label text c_{i_w} of each image was annotated.

For the text in the training set, information gain (IG) algorithm [9] was used to extract several words with the maximum IG value, which were taken as the feature c_d of the text data.

For the co-occurrence data of the image and text features, since the Visual Synset is the dataset with labels, c_{i_i} and c_d can be directly extracted from it. Moreover, the co-occurrence data of c_{i_w} and c_d can be used to calculate the probability association between the image feature and text feature.

3.2 Experiment Configuration

In the experiment, the experimental result of the text was compared to that in literature [4]. Similarly, 10 groups of task were set up, which are bear vs kayak, binocular vs cake, coin vs horse, laptop vs sword, tomato vs llama, dog vs canoe, cd vs greyhound, guitar vs snake, cat vs dog and horse vs ox. Each of them was binary classification. There were obvious semantic gaps between the last two groups of the images to be classified. In each classification task, for the training set of the text data, 899 positive and negative examples were set up, respectively. Moreover, for each training set of image data, 120 positive and negative examples were set up, respectively. According to the result in literature [4], the value of λ in formula (1) was chosen as 1. k in formula (7) denotes the incidence number between the visual feature and text feature, and the empirical value was chosen as 3. k_2 in formula (8) denotes the average entropy of the images in the training set for this classification task.

3.3 Experiment Analysis

The experimental result is shown as Figure 2. The performances of TLR-Fusion are better than those of the TL-Risk for each classification. Different values of the

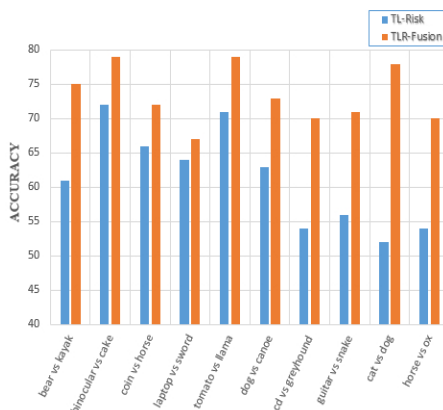


Fig. 2. Experiment Result

parameter k_2 in section 2.2.1 greatly influence the TL-Fusion model. It is usually effective to take the average entropy of all the training sets for the classification task as k_2 value.

For the two classifications with semantic gaps, cat vs dog and horse vs ox, the experimental result showed that with the method proposed in literature [4], they could not be effectively distinguished. However, high accuracy was maintained with the proposed TLR-Fusion framework in this study. Under the condition that the distributions of the image features were similar, the semantic labels became dominant, so the classification accuracy was greatly promoted.

4 Conclusion

For the problem that the semantic gap could not be resolved with the traditional image classification method based on a single feature space, a framework of image classification based on the semantic transfer learning is proposed in this study. Features in multi-modal space were comprehensively used to transfer the semantic feature in text space to the image space, and then the images were classified with the text semantics relevant to the image. In particular, during classification, the entropy of the distribution of the visual feature of test images was calculated. Using the entropy, the parameters in the model based on the co-occurrence probability could be adaptively adjusted. Hence the classification accuracy was further promoted. The experimental result proved the validity of the proposed method.

Acknowledgements. This research study was supported by the National Basic Research Program of China (973 Program) 2012CB821200(2012CB821206), the National Natural Science Foundation of China (61320106006), the Fundamental Research Funds for the Central Universities(No.2013RC0306).

References

1. Caruana, R.: Multitask learning. *Machine Learning* 28(1), 41–75 (1997)
2. Zadrozny, B.: Learning and evaluating classifiers under sample selection bias. In: *Proceedings of the Twenty-First International Conference on Machine Learning* (2004)
3. Heckman, J.J.: Sample selection bias as a specification error. *Econometrica* 47, 153–161 (1979)
4. Dai, W., Chen, Y., Xue, G.-R., Yang, Q., Yu, Y.: Translated Learning: Transfer Learning across Different Feature Spaces. In: *Advances in Neural Information Processing Systems 21 (NIPS 2008)*, Vancouver, British Columbia, Canada, December 8-13 (2008)
5. Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning* 42(1-2), 177–196 (2001), doi:10.1023/A:1007617005950
6. Li, Z.X., Shi, Z.P., Li, Z.Q., Shi, Z.Z.: Automatic image annotation by fusing semantic topics. *Journal of Software* 22(4), 801–812 (2011), <http://www.jos.org.cn/1000-9825/3742.htm>
7. Tsai, D., Jing, Y., Liu, Y., Rowley, H.A., Ioffe, S., Rehg, J.M.: Large-scale image annotation using visual synset. In: *ICCV*, pp. 611–618 (2011)
8. Lowe, D.: Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
9. Yang, Y., Pedersen, J.: A comparative study on feature selection in text categorization

Automatic Dehumidifier Control System for Greenhouse Using Smart Phone

Young-Jae Lee¹, Kyung-Wook Park², and Eung-Kon Kim^{1,*}

¹ Department of Computer Science, Sunchon National University,
255 Jungang-ro, Suncheon, Jellanam-do, Republic of Korea

² Division of Culture Contents, Chonnam National University,
50 Daehak-ro, Yeosu, Jellanam-do, Republic of Korea
skyit89@nate.com, zergadiss73@chonnam.ac.kr,
kek@sunchon.ac.kr

Abstract. Agriculture and stockbreeding industry has been putting a lot of effort to optimize the growing environment of animals and plants. However, a considerable damage is occurring as it is exposed to the growth obstacle and harmful insects because the temperature and humidity control is not equipped properly. Measures using air circulation fan and industrial dehumidifier are being researched recently to resolve such problem. This paper suggests industrial dehumidifier automatic controlling system that automatically maintains optimized humidity using smart phone. The suggested system composes a small-scale sensor network with temperature and humidity sensor installed in dehumidifier. Also, it provides optimized temperature and humidity environment in growing environment of plants and animals.

Keywords: Dehumidification, Wireless Sensor Network, Remote control, Greenhouse, Smartphone.

Introduction

Currently, agriculture and stockbreeding is putting great effort to create the optimum growing environment of plants and animals, and they have reached a stage of quality and productivity improvement through optimum facility environment [1,2].

However, a considerable damage is still taking place in farming families from growth disturbance and disease and insects because the temperature/ humidity control is not equipped properly, and they are preparing countermeasures with increasing supply of air circulation fan or industrial dehumidifier, however, the necessity of dehumidifier that fits agronomic characteristics is being highlighted as the results fell short of its expectations[3,4].

For controlled horticulture, investment for increasing productivity and environment improvement is being avoided due to the increased working expense of farming families. Most of the cases were of not meeting requirements of controlled horticulture as high price of existing products, volume, and usage.

* Corresponding author.

For countermeasure of this matter, growth of productivity of harm houses by installing moderately priced IT convergence products on existing products, increase in rural household income through quality improvement, and technical development and distribution for environmental improvement and reduced management expense of farm houses are urgently needed.

This paper suggests remote monitoring and control system of air circulation-fan and dehumidification device using smart phone and the agricultural dehumidifier automatic control system using smart phone that automatically maintains optimized humidity in crops.

A suggested system composes small-scaled sensor network autonomously by installing the temperature and humidity sensor in existing air circulation fan or dehumidification device.

Also, it is possible to improve the productivity of crops and save management expense of harm housings by conveniently monitoring the greenhouse condition and controlling humidity automatically through smart phone.

Automatic Dehumidifier Control System for Agricultural Use Using Smart Phone

A proposed system suggested the smart phone system, and the diagram of entire system is as shown in Figure 1.

The system is composed of an integrated gateway module and smart phone monitoring software that manage the controlling module and sensor data of dehumidifier.

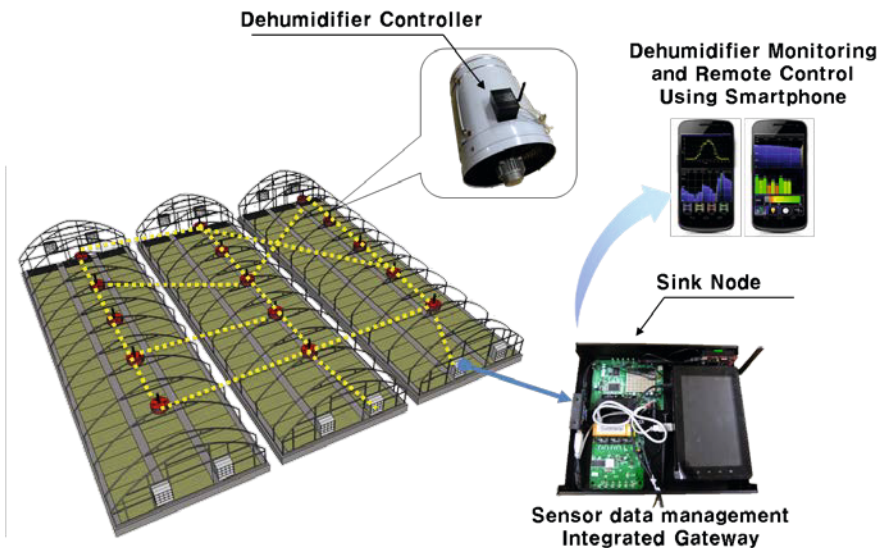


Fig. 1. Diagram of automatic dehumidifier control system for agricultural use using smart phone

Sensor Data Managerial Integrated Gateway. First, this paper suggests the android embedded based gateway that integrated the management functions of sensor data transmitter and internet connection management function.

Sensor data integrated gateway is composed with the communication module for communicating with the dehumidifier control system installed in the installation house from external internet network using Wibro modem and Wifi communication module, zigbee-based sensor data exchange module, and sensor data management and smart phone remote control service[5].

First, a diagram of communication between integrated gateway and smart phone is as shown in Figure 2.

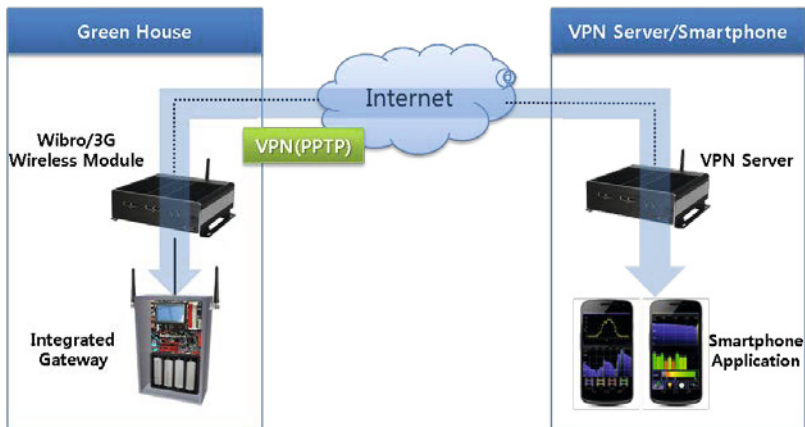


Fig. 2. 3G, Wibro, Wifi-based communication diagram

As it is composed to be connected to the external internet network using VPN, Wibro, 3G, and LTE and VPN, stable communication connection with smart phone of high-mobility becomes possible.

The integrated gateway that manages sensor data performs a sink node role of sensor network while storing received sensor data(temperature, humidity, intensity of illumination, and dehumidifier operating condition) from the network and providing them to smart phones.

Mobile web application server(WAS: Web Application Server) for convenient interlocking with smart phone was implemented and constructed to support android and IOS using standardized interface of HTTP and XML.

The integrated gateway provides the monitoring service, dehumidifier control service, environment setting service, and log function that direct the log in service, current condition of dehumidifier, and sensor data.

Figure 3 shows a developed sensor data managerial integrated gateway product.

A Sensor Node Attached Inside the Dehumidifier. The dehumidifier control module suggested in this paper is the zigbee based USN sensor node for collecting and transmitting temperature and humidity data and processes the dehumidifier control signal from smart phone.



Fig. 3. Sensor data managerial integrated gateway

The dehumidifier ON/OFF control module and a circuit diagram is as shown in Figure 4.

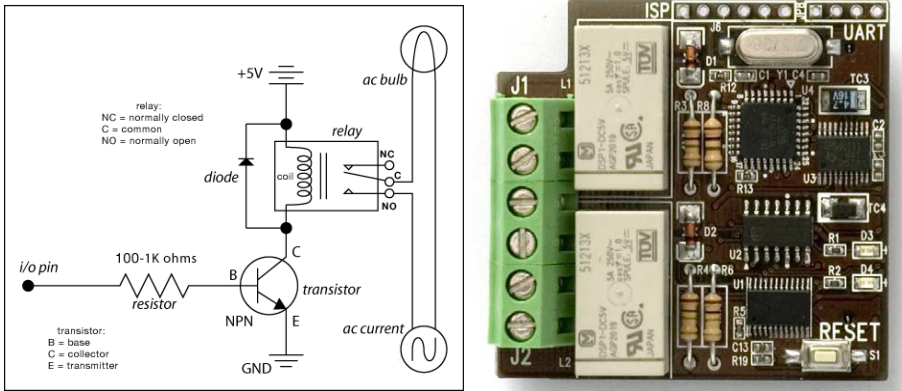


Fig. 4. Dehumidifier ON/OFF control module

The ultimately developed dehumidifier control module is as shown in Figure 5.

The system is constructed to compose a small-scaled sensor network autonomously by attaching the temperature and humidity sensor inside an existing air circulation fan or dehumidification device.

Smart Phone Monitoring and Remote Control Service. The dehumidification remote control application supports android smart phones and i-phones that are mostly used nowadays, and implemented the hybrid application using PhoneGap to apply various platforms in smart phone.

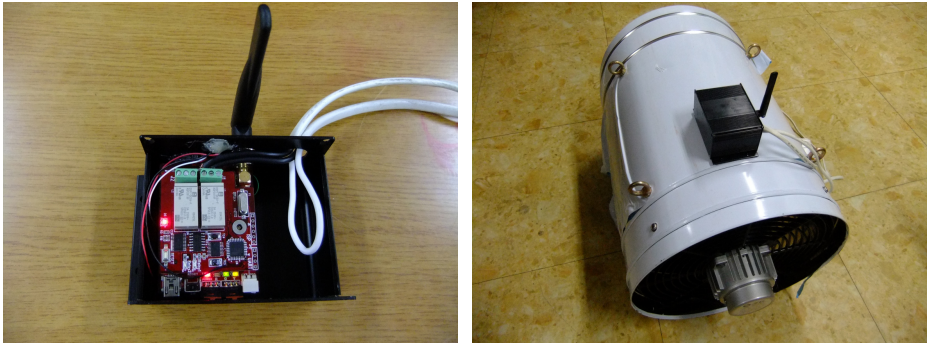


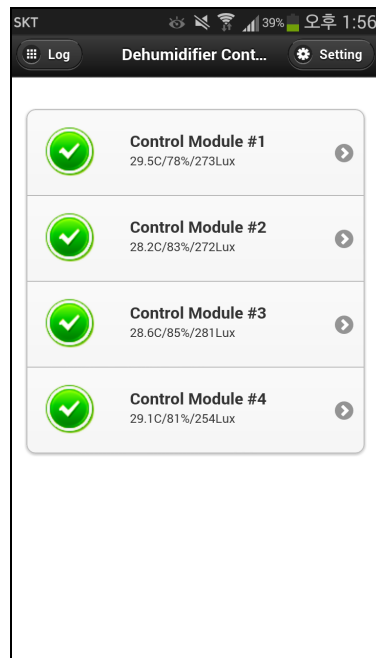
Fig. 5. Dehumidifier control module

Also, user interface of application was composed using HTML5 and Javascript. The embodied screen of dehumidifier remote control smart phone application suggested in this paper is as in the following..

Figure 6 is a screen of user authentication and running the main menu by entering ID and password for field managers to control the dehumidifier remotely.



(a) Log-in



(b) Main dehumidifier control

Fig. 6. User log-in and main menu

The dehumidifier control module setting and system setting are as shown in Figure 7.

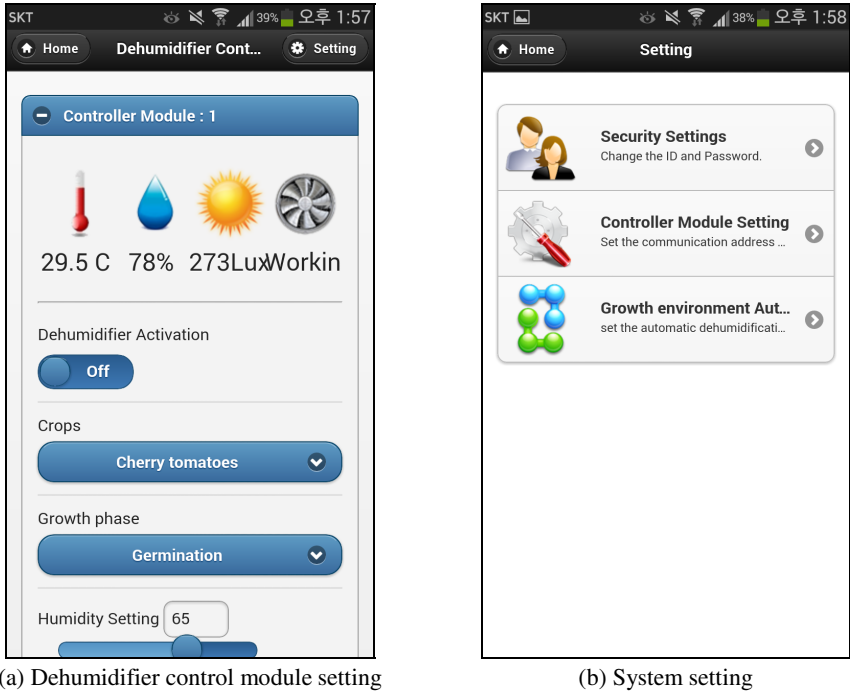


Fig. 7. The dehumidifier control module setting and system setting

The dehumidifier control module setting checks the current condition or sensor data, and is composed to set the dehumidifier ON/OFF and humidity controller.

The system setting is composed with the access security setting of changeable ID and password, controller module setting for setting the communication address of controller module, and growth environment automatic setting of setting-up the automatic dehumidification function according to the growth step.

The administrator setting of smart phone application and log menu are as shown in Figure 8.

Appointing or changing the site manager is possible in administrator setting, and relevant log information can be provided in the log menu composed with the log inquiry of inquiring the access function of system, dehumidifier operation log that inquires the ON/OFF record of dehumidifier, and sensor log that inquires the temperature, humidity, intensity of illumination data information.

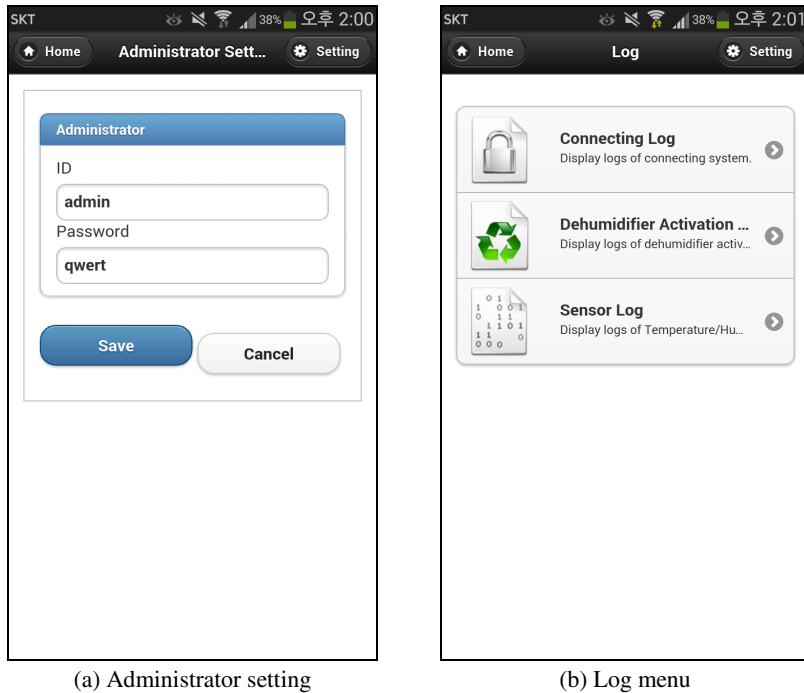


Fig. 8. Administrator setting and log menu

Conclusion

This paper suggested the automatic control system of agricultural dehumidifier using smart phone for the field manager to maintain optimized environment for crops through existing air circulation-fan or temperature and humidity sensor of dehumidification device.

The suggested system developed smart phone application on the basis of hybrid architecture for simultaneous support to be possible with one development in android, smart phone, or i-phone platform.

Also, intuitive and convenient usage is possible by remotely controlling the industrial dehumidifier with smart phone and providing data as current condition and sensor data of dehumidifier to administrators, and remote control of various agricultural equipment are possible with no additional installation cost by adding and operating small scale USN on agricultural dehumidifier.

The low cost USN construction, managerial technique, agricultural dehumidification monitoring and automatic controlling technique using smart phone are expected to be applied in automatic controlling system of other agricultural products in the future.

Acknowledgment. This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2013H1B8A2032217).

References

- [1] Barenbrug, A.W.T.: Psychrometry and Psychrometric Charts, 3rd edn. Cape and Transvaal Printers Ltd., Cape Town (1974)
- [2] Lee, C.K., Jung, I.G., Sung, J.H., Lee, B.Y., Chung, S.O., Park, W.K.: The current status analysis of precision agriculture research in USA and Japan. *J. Korean Soc. Int. Agr.* 17, 133–140 (2005)
- [3] Liu, Y., et al.: Design of Automatic Defrosting and Dehumidification System in Greenhouse. *Advanced Materials Research* 422, 196–199 (2012)
- [4] Niroomand, N., Zamen, M., Amidpour, M.: Theoretical investigation of using a direct contact dehumidifier in humidification–dehumidification desalination unit based on an open air cycle. *Desalination and Water Treatment* ahead-of-print, 1–11 (2014)
- [5] Baronti, P., et al.: Wireless sensor networks: A survey on the state of the art and the 802.15. 4 and ZigBee standards. *Computer communications* 30(7), 1655–1695 (2007)

Thermal Image-Based Disc Pads Diagnosis System in Grab Crane

Yeon-Jae Oh¹, Kyoung-Wook Park², and Eung-Kon Kim^{1,*}

¹ Department of Computer Science, Sunchon National University,
255 Jungang-ro, Suncheon, Jellanam-do, Republic of Korea

² Division of Culture Contents, Chonnam National University,
50 Daehak-ro, Yeosu, Jellanam-do, Republic of Korea
{oksug10, kek}@sunchon.ac.kr,
zergadiss73@chonnam.ac.kr

Abstract. Grab cranes are used for multi-purpose when the sand and soil are deposited into harbor wharf or the undersea construction is performed. Among the components of crane grab, the wire drum and disc brake pad are key expendables and have disadvantages that lot of heat is generated and very expensive when replacing them. In this study, the thermal image analysis for the disc brake, which works with wire drum of the crane is suggested. The suggested system performs the pad thermal diagnosis through the thermal image using the characteristics that the disc and pad surface temperatures are distributed abnormally before the brake failure and the disc pad damage. Therefore, the damage by the failure can be prevented by discovering the abnormality of the machine parts before failure and the life cycle of the pad and the cost can be extended and saved by operating the crane performing constant checkup for the overload.

Keywords: Thermal image camera, Thermal Image Diagnosis, Disk Pad Diagnosis.

Introduction

Recently, the importance of the thermal image camera is being increased in the diverse areas such as the industry, chemical engineering, semiconductor, shipbuilding, etc. In addition, by the development of smart devices, diverse mobile monitoring solutions are needed[1].

The continuous growth of our economy has been made by the navigation and port technologies and the shipbuilding industry together with the development of cutting edge technologies. Particularly, the marine transportation handles more than 98% of the world cargo volume and to deal with the cargo volume, the ships are growing bigger and faster[2], for which the navigational equipment and additional equipment[3] are installed in the ships. Among them, the grab cranes is used fir multipurpose when the

* Corresponding author.

sands and soils are deposited in the port wharf or performing the undersea construction. Among the components of grab crane, the wire drum and disc brake pad are the key expendables and have disadvantages that lot of heat is generated and very expensive when replacing.

In this paper, the thermal image analysis for the disc brake, which works with wire drum of the crane is suggested. The suggested system performs the pad thermal diagnosis through the thermal image using the characteristics that the disc and pad surface temperatures are distributed abnormally before the brake failure and the disc pad damage.

Although, in the existing diagnosis system, the checkup and maintenance are made after occurring the failure symptoms, since the disc pad diagnosis system grafted with the thermal image camera can identify the failure symptoms before occurring the failure, it can prevent the accident in advance and save the time to find the cause of problem by stopping the system. In addition, since the temperature of the monitored component is collected, it can predict when the failure may occur based on the database made for the past.

Related Research

Temperature Measurement through the Thermal Image Camera. All objects having temperature emit those heat with the infrared. All the objects having higher than absolute temperature of 0 (-273°C) radiate the rays of infrared domain. The infrared is radiated even to the ice, the object having very low temperature. The higher the temperature of the object, the more the infrared is radiated. Using such principle, diverse diagnoses are made using the thermal image equipment[4].

Infrared is the electromagnetic wave having the wavelength between 0.72~1000 μm , from the long wavelength of visible rays to the short wavelength of microwave. Since the infrared has close relation with the temperature of the objects, all the objects having the temperature higher than 0°C of absolute temperature are radiating the infrared constantly. These imaging devices are filming the thermal image on the surface of the object and use it as examining the state of the object by analyzing the temperature distribution and changes. According to the Planck's law, the radiant intensity (homogeneous radiant existence) of the black body can be obtained from Equation (1) [5].

$$\omega_{\lambda} = \frac{c_1}{\lambda^5} \left(e^{\frac{c_2}{\lambda T}} - 1 \right)^{-1} (W/cm \cdot \mu m) \quad (1)$$

where, λ is wavelength(μm), T is the absolute temperature of the black body (K), C_1 is $3.7402 \times (10^{-12} \text{W} \cdot \text{cm}^2)$, and C_2 is $1.4388 (\text{cm} \cdot ^\circ\text{C})$.

As an object, which absorbs the infrared projected into it completely, the black body is the object having greater radiant energy at certain temperature than any other objects.

Electromagnetic Wave Domain of Infrared Thermal Image. The infrared is a form of electromagnetic radiation having long wavelength than visible rays. For the other radiant waves, there are X-rays, ultraviolet rays, radio waves, etc. The range of the

electromagnetic radiation is determined by the frequency or wavelength. The range of the infrared detector or system is determined by the wavelength. The system, which detects the radiation in the band of 8 -12 μm , is generally referred to as 'long wavelength' and that can detect the radiation in the band of 3-5 μm is referred to as 'short wavelength'. The visible range of the electromagnetic spectrum is 0.4 - 0.75 μm [6].

Image Subtraction Method. Image subtraction is the method of processing the image and has 3 characteristics. First, it helps to identify small changes in the temperature. Second, it helps to suppress the influences of the heat reflection. Third, in case of applying consecutively, it helps, for example, visualizing the time derivative by the transient phenomena.

The image subtraction can be performed mainly by two methods. One is to subtract reference image from the each image in the recorded sequence and the other is the continuous process, that is, to subtract the each image from the previous image. The first method forms a new order as a result and shows the difference in the temperature between each image and the reference image. The comparison between the reference image and resulted image, in fact, is not helpful to find the location of hot spot but the subtracted image can indicate the location of abnormal heat source exactly. In the second method of image subtraction, the consecutive image subtraction is made, in which case, the difference signal of the nth image pixel (c, y) at the sequence $\Delta S(x, y, n)$ is calculated by following equation[4]. Since the image n and n-1 are separated by the time interval Δt , all of which are defined by the recording frame speed, the consecutive image subtraction refers to calculate the first derivative in the time domain. If the consecutive subtraction algorithm is applied to the image sequence consecutively subtracted second time, the second derivative of the temperature within the time domain will be calculated, which can provide the additional information of the time-based temperature action of the object to be examined and ,for example, is used in the pulse thermography.

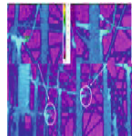
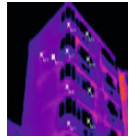
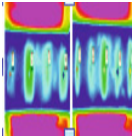
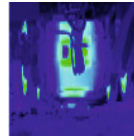
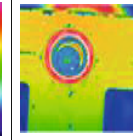
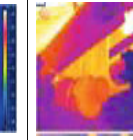
The consecutive image subtraction can be used to extract the characteristics of the moving objects in real-time when the object and the background are static or does not have characteristics.

Cases of Infrared Diagnosis System. If the infrared camera is used, the performance or the state of the object can be evaluated by examining the condition of the object using non-contact method. Therefore, it is actively used in diverse areas. Table 1 is shown the cases used the thermal image diagnosis system using thermal image camera[7-12].

Disc Pad Diagnosis System Using Thermal Image Camera

Grab cranes are used for multi-purpose when the sand and soil are deposited into harbor wharf or the undersea construction is performed. Among the components of grab crane, the wire drum and disc brake pad are key expendables and have disadvantages that lot of heat is generated and very expensive when replacing them. In this paper, the disc pad diagnosis system using thermal image is suggested. The suggested system performs the thermal image monitoring for the pad thermal analysis using the characteristics that the disc and pad surface temperatures are distributed abnormally before the brake failure and the disc pad damage. Figure 1 is shown the system layout for the brake and disc diagnosis and monitoring using thermal image.

Table 1. Cases of Thermal Image Diagnosis System

Power board/substation diagnosis using time series data of thermal image camera [7]	Inspect the aging and tearing of the apartment with the thermal images [8]	Detect the defects in the thin wall piping with small calibre in the nuclear power plant [9]	Detect and analyze the abnormal heat generation at the substructure of railway car [10]	Diagnose and measure the dynamic load of ball bearing [11]	Inspect the industrial plant condition using the thermal image camera [12]
					

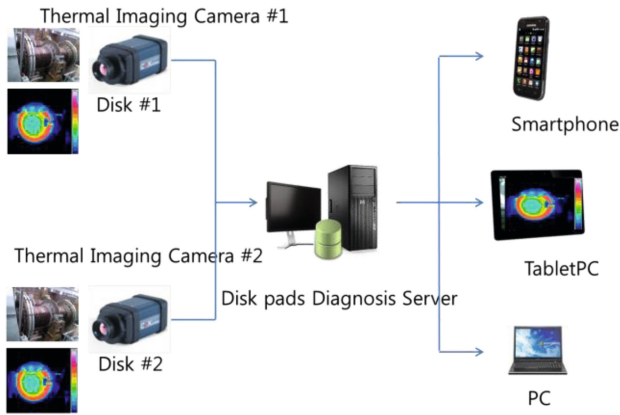


Fig. 1. System structure

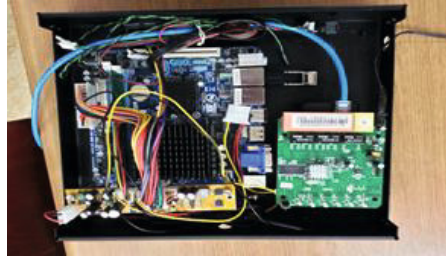
The image of the disc pad of the grab cranes are fed into the thermal image control and diagnosis server using thermal image camera. The thermal distribution chart of the thermal disc pad fed is compared with the existing data and can be verified through the smart phone, tablet PC or control PC of the user. In this moment, if the overheating may occur above than designated temperature, the notification message will be sent to the controller. Figure 2 is shown the screen fed the images using thermal image camera, wireless image transmission module.

The disc pad diagnosis system using thermal image us basically composed with the thermal image camera to monitor the heating condition of the disc pad and surrounding components and the communication equipment to transmit the thermal image data to control/diagnosis server. The thermal image control server receives the images of the thermal image camera installed at each disc through TCP/IP, stores them with the disc No. operator, time information, and at the same time, performs the streaming service that transmits the images to the smart phone or tablet PC, which request the monitoring.

To analyze the thermal image, the software module to extract the feature point of the temperature distribution was developed. Figure 3 shows the screen performed the login, select disk, transmission test of the thermal image camera's image for the disc and pad of current grab crane.

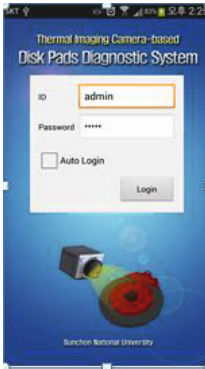


(a) Screen Fed Images from Thermal Image Camera

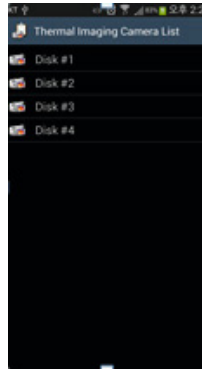


(b) Wireless Image Transmission Module

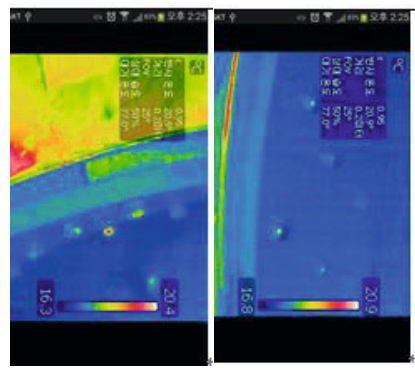
Fig. 2. Screen Fed Images from Thermal Image Camera, Wireless Image Transmission Module



(a) Login



(b) Select Disk



(c) Thermal Images

Fig. 3. Screen shots of login, select disk and thermal image transmission test

The detailed specification of the thermal image camera used in the thermal image monitoring system is shown in Table 2.

Table 2. Detailed Specification of Thermal Image Camera

Description	Specification
Resolution	320×240
Detector type	LWR <80mK@f/1, 60Hz, 300K
Spectral Range	8 to 14μm
Measurement Range	-20°C to 120°C, up to 650°C
Frame Rate	NTSC(60Hz), PAL(50Hz)
Ethernet	10/100base, TCP/IP
Lens	20mm Motorized Focus lens

It determines the abnormality of the thermal distribution of the disc and pad by performing image analysis according to the designated rules for the images transmitted from the thermal image control server. Table 3 is shown the rule for the image analysis.

Table 3. Rules for Image Analysis

Rule	Explanation
Disc Overheat Temperature (300°C)	Designate the maximum disc temperature
Disc pad thermal distribution (100°C)	Designate range between maximum and minimum values of the measured temperature on the disc surface
Pad overheat temperature (300°C)	Designate the maximum temperature of the pad surface
Abnormal pad thermal distribution (100°C)	Designate the difference between the maximum and minimum temperature of the pad surface

The thermal image analysis of the disc and brake pad compares and analyzes the difference between the maximum temperature and the minimum temperature from the thermal image of the disc and the brake pad area and the distribution of the temperature by time and determines as abnormality if they may out of the rules. It analyzes the thermal image according to the rules established and if determines abnormal situation, sends the notification message to the relevant controller. The notification message is composed with the contents of the message, time of occurrence, disc No. and thermal image and the types of notification messages are shown in Table 4.

The application for the smart phone and tablet PC was developed in order for the operator to monitor the thermal images of the disc and disc pad easily during the operation. Table 5 is shown the functions of the smart phone application.

Table 4. Type of Notification Message

Notification Message	Explanation
Disc Overheat	Detect the higher value than designated temperature from the disc surface
Abnormal disc thermal distribution	Detect the unusual pattern of thermal distribution from the disc surface
Pad Overheat	Detect the higher value than the designated temperature from the pad surface
Abnormal pad thermal distribution	Detect the unusual pattern of thermal distribution from the pad surface

Table 5. Functions of Smart Phone Application

No	Description	Function	Explanation
1	Log-in	User Log-in	Function to authenticate the user
2		User Registration and Modification	Register user ID and password
3	Monitoring	List of thermal image camera	Display the list of thermal image cameras
4		Thermal image monitoring	Visualize the images of the thermal image camera selected by user
5	Notification Service	Notify abnormal situation	Display the notification message if the abnormal situation is transmitted from the server

Conclusion

All the machine parts or electric components have characteristics that the temperature is rising before they are failed or damaged. Since the thermal image camera detects the area occurred the heat and display it with images, it has advantage that finds such machine parts and electric components before they may be failed. Therefore, the thermal image camera displays the out-of-order signal of the key components such as bearing shaft, gear, brake, pad, etc., immediately with high reliability and allows performing preventive maintenance or replacing the relevant component in advance.

In this paper, the thermal image diagnosis system for the disc brake pad, which works in the wire drum of grab crane was suggested. The suggested system performs the pad heat diagnosis analysis through the thermal images using the characteristics that the disc and pad surface temperature is distributed abnormally before they may be failed or damaged. Particularly, by sending the notification message to the controller when the overheat or overload occurs in the disc and brake pad, it always can extend the life cycle and save the cost by performing the checkups for the overload and operating the crane

Acknowledgment. “This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2013 H1B8A2032217).”

References

- [1] Gleiter, A., Riegert, G., Zweschper, T., Busse, G.: Ultrasound lockin thermography for advanced depth resolved defect selective imaging. In: European Conference on Nondestructive Testing, ECNDTconference Proceedings, Berlin, Germany, September 25-29 (2006)

- [2] Lee, Y.S., Ahn, Y.J.: A Study on the Standard Ship's Length of Domestic Trade Port. *Journal of the Korean Society of Marine Environment & Safety* 19(2), 164–170 (2013)
- [3] Jeong, K.Y., Kim, E.K.: Image Guidance System for Working with Abalone Park. *Journal of The Korea Institute of Electronic communication Sciences* 9(3), 369–376 (2014)
- [4] Kim, J.Y.: Thermo-anaiysis of machning center main-axis thermo-displacement for infrared rays thermo-image camera. In: *Proceedings of the Korean Society of Machine Tool Engineers Conference*, vol. 2001, pp. 125–130 (2001)
- [5] Kim, D.Y., Hong, D.P., Yu, C.H.: Concdtion Monitoring under In-situ Lubrication Status of Bearing Using InfrnredThennography. *Journal of the Korean Society for Nondestructive Testing* 30(2), 121–125 (2010)
- [6] Moore, P.O.: *Nondestructive Testing Handbook. Infrared and Thermal Testing*, 3rd edn., vol. 3, pp. 223–246. ASTM (2001)
- [7] Hwqng, S.S., Bae, Y.C.: Diagnosis of power supply using time-series of infrared camera. *Journal of The Korea Institute of Electronic communication Sciences* 7(6), 144–147 (2012)
- [8] Cheong, C.H.: Study on External Wall Surface Temperature Pattern of Old Residential Building in Jinju City. *Architectural Institute of Conference Proceedings published* 34(1), 265–266 (2014)
- [9] Yun, K.W., Kim, D.L., Jung, H.C.: Application Defects Detection in the Small-Bore Pipe Using Infrared Thermography Technique. *Journal of the Korean Society for Nondestructive Testing* 33(1), 34–39 (2013)
- [10] Kang, B.B., Kim, J.G., Jimg, C.M.: Thermal image analysis method for the detection of the abnormal heat under the railway vehicle 13(1), 1347–1360 (2013)
- [11] Hong, D.P., Kim, H.J., Kim, W.T.: On Diagnosis Measurement under Dynamic Loading of Ball Bearing using Numerical Thermal Analysis and Infrared Thermography. *Journal of the Korean Society for Nondestructive Testing* 33(4), 354–359 (2013)
- [12] Jeon, B.J., Kim, T.W., Kim, S.G.: Panorama Image Processing for Condition Monitoring with Thermography in Power Plant. *Journal of the Korean Society for Nondestructive Testing* 30(2), 98–103 (2010)

Image Acquisition System for Construction Inspection Based on Small Unmanned Aerial Vehicle

Sungsuk Choi and Eungkon Kim*

Department of Computer Science, Suncheon National University, Korea
{mstar2000, kek}@suncheon.ac.kr

Abstract. As for large scaled skyscrapers, it is time-consuming, requires a significant amount of expenditures, and too risky for an inspector to examine in person with tools on hands. Therefore, unmanned aerial vehicle is specifically required to inspect the facilities in order to deal with limited access of an inspector, to examine multiple structures, and to confirm whether the inspected areas are damaged in each season or daily time frame. This study is intended to specifically develop the small aerial vehicle-based Image acquisition system for the inspection of structures aiming to identify the cracks of the building. Hereupon, safety inspection was conducted on the building. It is also intended to utilize the suggested system in this study to safety inspection in diverse fields of industry in the future.

Keywords: Small unmanned aerial vehicle, Image acquisition system, Monitoring facilities, Safety inspection, Image processing, Edge detection, Building crack.

Introduction

It is recently in an increasing trend that structures are damaged and collapsed. Especially, it is time-consuming, requires a significant amount of expenses, and too risky for an inspector to examine large-scaled skyscrapers in person with tools on hands. Therefore, it is difficult to provide an accurate diagnosis.

Therefore, it is urgently required to inspect structures by using unmanned aerial vehicle in order to deal with limited access of an inspector, to inspect multi-structures at a time, and to confirm whether the inspected areas are damaged in each season or daily time frame.

Unmanned aerial vehicle does not require a controller to board and hence is remotely controlled on the ground by a particular person. Therefore, it is called as 'drone' [1].

In the past, it was used for the surveillance on the target area or the military purpose. However, it recently started being utilized by the private sectors and also for research inspection including many other diverse purposes [2].

In this study, it is specifically suggested to use a small aerial vehicle-based Image acquisition system for inspection on structures. Suggested system uses Image acquisition algorithms designed to accurately detect whether a particular structure is

* Corresponding author.

damaged through videos or images recorded by using an aerial vehicle developing the program for users to conveniently operate. In addition, it is received with various sensor data making it feasible to conveniently identify the current status and designing the program that visualizes data.

This system is intended to install video camera on the small aerial vehicle approaching to the structure of skyscrapers through remote control, identifying the problem, and making it possible to precisely inspect and diagnose more accurately via image processing on the ground.

Related Research

Definition and Applied Technology of Small Aerial Vehicle. Unmanned aerial vehicle (UAV) is a flight vehicle that man does not board on. Focusing on the fact that there is a person remotely controlling it on the ground, it is often called as uninhabited aerial (air) vehicle. It is normally categorized by the duty or flight altitude and scale [3].

As for the advantage of unmanned aerial vehicle, it is capable of performing the risky and difficult tasks that a human was in charge of in the past. Floating the vast area for a long time, it is able to supervise the risk of forest fire or to navigate the area that a human cannot enter due to contamination from radioactivity. Furthermore, it is also appropriate role of unmanned aerial vehicle to reconnoiter the area with high risk of being brought down by an antiaircraft missile [4].

Diverse scopes of study are being proceeded on the controlling technology of unmanned aerial vehicle to apply it on various fields including document recording or crop-dusting [5]. In addition, the image analyzing technology is currently being developed in the advance level that is capable of replacing physical sensors including the fire or movement supervision. [6].

It is in an increasing trend that maintenance and safety inspection of the structures are in a limelight mostly in Japan, Germany, and England. Hereupon, Fraunhofer ISE in Germany and Cyberhawk Company in England have developed and commercialized a small aerial vehicle for safety management of the wind power structures.

This study aims to install the camera on a small aerial vehicle recording the status of the structure in a remote control and inspecting the safety of structures through image process.

Image Process Method and Algorithm. Image process indicates all ranges of the work producing and processing the image by using computer and also is related to all tasks related to the image including interpreting and recognizing the image.

Detection algorithm using the image process is being used in a wide range of industrial fields.

Image process was used to detect the cracks on the surface of concrete slab that specifically utilized the fuzzy inference mechanism and SOM algorithm [7]. In addition, image process was also utilized when developing the system that inspected on the bridge by using vision-based crack detecting algorithm [8].

This study is intended to use the changes in brightness converting the pre-recorded video to the gray scale and to develop the image acquisition system that detects damaged parts of skyscrapers by using the edge detection algorithm.

Image Acquisition System

System Overview. In this study, it is specifically suggested to program by utilizing a small aerial vehicle so that an operator is able to conveniently and accurately inspect damaged areas of structures through recorded videos or images in the use of image acquisition system. In addition, it is also suggested to develop a program that visualizes the data for easy identification of the status after receiving various types of sensor data.

The overall system flow is shown in the Fig. 1.

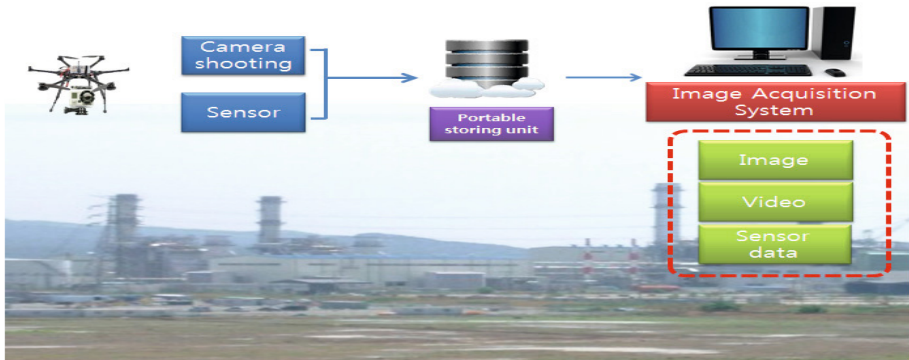


Fig. 1. Outline of Image Acquisition System

Procedures of Acquiring Image. Image processing technology is required to acquire the image.

When the images delivered to the video camera or thermal infrared camera are damaged or ill-defined or when intending to look at the preferred areas in details, it is able to efficiently manage the data by using image processing technology.

In this study, it is intended to develop the API-based applied program for Window by using algorithm that detects the damaged areas of buildings through the procedures of acquiring image as shown in the Fig. 2.

Suggested system separates the image processing procedures and video processing procedures.

Development Environment. Development environment of the program that was used to acquire the pre-entered video images is as follows in the Table 1.

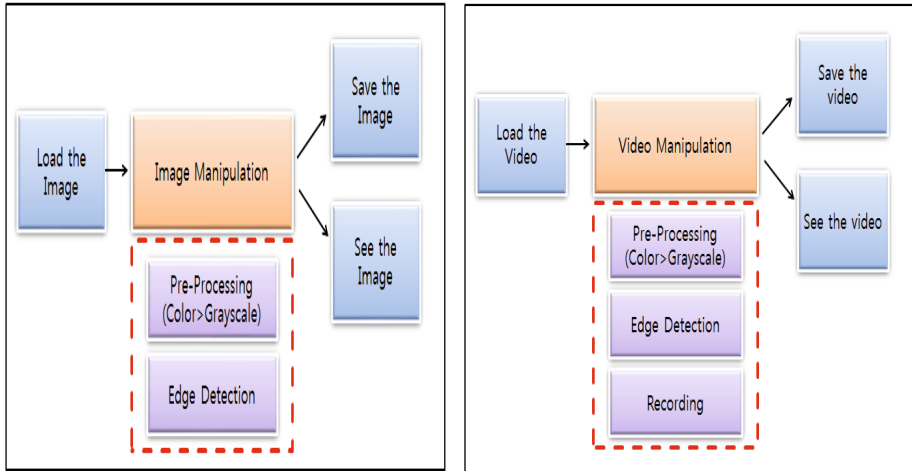


Image Processing Procedures

Video Processing Procedures

Fig. 2. Procedures of Acquiring Videos

Table 1. Development Environment

Classification	Detailed Environment
CPU	Intel(R) Core(TM) i3 CPU M330 @213Ghz
RAM	2.0GB
OS	Windows7 Service pack 1
Language	C, Visual C++
Develop Tool	Microsoft visual studio2010
Library	Window API, Opencv2.3.1, Chartlibrary
Device	LG Notebook

Image Processing Techniques. This system uses the aerial vehicle converting the recorded color image into gray scale and inspecting the safety or whether a certain part of the structure is damaged. Therefore, it is important to detect the accurate edge.

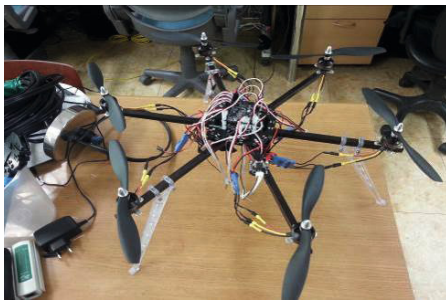
Especially, it is required to detect edges when detecting the cracked areas inside of the building. Most of the edges detection masks are sensitive with noises. Therefore, most of the cases tend to regard a low level of noise as edge and detect it. There is an edge detecting technique that uses canny mask supplementing such a weakness. However, it aims to detect strong edges that are not sensitive with noises.

Therefore, this system aims to identify damaged areas inside of the building by using Canny edge detection algorithm

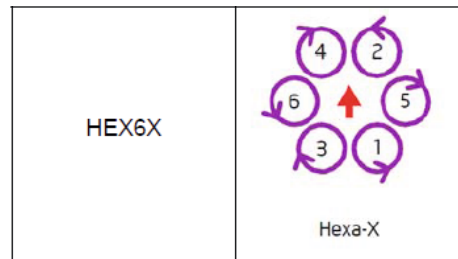
Experiment and Realization

Experiment Environment. Hexa-copter, a type of small aerial vehicle, was used to identify whether a part of skyscraper was damaged to experiment and realize the plans in this study.

Hexa-copter board being used in the experiment was Black 32 consisting of gyro, acceleration pressure, and terrestrial magnetism sensor to provide a stable status of flight based on MultiWii open source.



Hexa-copter used in the experiment



Setup of flight mod

Fig. 3. Hexa-copter used in the experiment and flight mode

Result of Realization. In this study, Window API was realized that video images recorded by a small aerial vehicle was loaded to inspect the safety of structures intending to experiment as to whether a part of inside the structure was cracked.

Realized image acquisition system is comprised of two major programs. First one is to load the color image converting it to gray scale and also to edge detection to identify whether the structures are cracked either inside or outside.

According to the detection of crack image inside and outside of the building, it was relatively convenient to detect it. However, there were several cases that a tiny part of crack was not well detected.



Fig. 4. Recorded images and also hexa-copter used in the experiment



Fig. 5. Experiment of Detecting the Crack of Buildings

The second image represents the program that loads the video image making it possible to see the image by processing immediately. Furthermore, it is available to capture and save the screen and also to receive the input from a camera if needed.



Fig. 6. Realize the Video Images

Conclusion and Follow-Up Study in the Future

In this study, a design of the small aerial vehicle-based image acquisition system for inspection of structures was suggested. Furthermore, the image acquisition system was realized as well.

Image processing algorithms were used to process the video image that was acquired by aerial vehicle removing the noises and leading to determine whether the structures were damaged through detection of edge.

It is currently planned to proceed follow-up study intending to acquire very safe and objective test procedures and also results compared to the previous method by providing more of convenient work environment by installing various types of sensors, thermal infrared camera, and various types of sensors on the small aerial vehicle by visualizing the data in the future.

Therefore, it is ultimately intended to apply it on the safety inspection of transmission towers, plant smokestacks, bridges, and piers that are dangerous and difficult for a person to approach based on aforementioned design.

It is much anticipated for the suggested system in this study to be highly useful on the safety inspection in diverse fields of industry in the future.

Acknowledgment. This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2013H1B8A2032217).

References

- [1] Sullivan, J.M.: Revolution or Evolution? The Rise of the UAVs, June 8-10, pp. 94–101. Cornell Univ.Ithaca, NY (2005)
- [2] Hoffmann, G.M., Huang, H., Waslander, S.L., Tomlin, C.J.: Quadrotor Helicopter Flight Dynamics and Control: Theory and Experiment. In: AIAA Guidance, Navigationand Control Conference and Exhibit (2007)
- [3] Bone, E., Bolkcom, C.: Unmanned aerial vehicles: Background and issues for congress. Report for Congress (2003)

- [4] Saripalli, S., Montgomery, J.F., Sukhatme, G.S.: Visually-guided landing of an unmanned aerial vehicle. *IEEE Transactions on Robotics and Autonomous Systems* (2003)
- [5] Odido, D., Madara, D.: Emerging Technologies: Use of Unmanned Aerial Systems in the Realisation of Vision 2030 Goals in the Counties. *International Journal of Applied* 3(8) (2013)
- [6] Castillo, P., Dzul, A., Lozano, R.: Real-Time Stabilization and Tracking of a Four-Rotor Mini Rotorcraft. *IEEE Trans. Automat.* 12 (July 2004)
- [7] Kim, K.-B.: Extraction of Concrete Slab Surface Cracks using Fuzzy Inference and SOM Algorithm. *CI* 49(2), 38–43 (2012)
- [8] Abdel-Qader, I., Abudayyeh, O., Kelly, M.E.: Analysis of Edge-Detection Techniques for Crack Identification in Bridges. In: *ASCE* (2003)

Portable Video System for Farm Growth Monitoring

Mi-jeong Park and Eung-gon Kim*

Department of Computer Science, Suncheon National University,
255 Jungang-ro, Suncheon-si, Jellanam-do, Republic Of Korea
{mj21, kek}@suncheon.ac.kr

Abstract. Various marine lives are mainly cultured from fish farms. This study focused on the abalone among many marine lives to research on the underwater monitoring system which monitors its growth and development and the amount of sludge. The insufficient feeding in the farm causes growth inhibition of abalone and over feeding causes contamination. The portable video system to monitor farm culturing proposed in this study would make scientific contribution to reduce cost and contamination by controlling the feeding according to the growth status of abalone.

Keywords: Image acquisition system, Monitoring facilities, Image processing, Smartphone, Nursery, PTZ, Augmented Reality.

1 Introduction

Overcrowded farming and over feeding exceeding appropriate facility volume increase organic loading in the surroundings and bottom of farm. Even its impact is limited only to fish farm, but equivalent to the impact of waste water disposed in the land.[1] Fishing nursery in our country is being seriously contaminated by contamination from the land and fish farm. Due to serious contamination because of aged farm and bottom materials caused by repeated cultivation, the productivity becomes lower. Therefore, it is urgent to improve bottom materials in the surroundings of farm and management of environment. Due to bottom materials and contamination of ecological environment, the mortality rate of abalone reached about 50%. Therefore, it is urgently needed to make efforts to improve ecological environment of its surroundings. [2]

Technology and product to monitor the degree of contamination is highly required but the current technology to monitor marine ecological environment should be installed directly in the farm which requires initial installation fee and high maintenance cost. As it is difficult to install in the current operating farm and has blind spot where it cannot monitor, it is not practical. [3]

This study focused on the abalone farm among the farms. As most of current underwater monitoring system is fixed type, it is easily eroded by salt and difficult to maintain because of high maintenance cost. However, using abalone growth monitoring system, appropriate feeding can be made according to the growth status while increasing the production. In addition, appropriate feeding control would reduce environmental contamination.

* Corresponding author.

2 Portable Video System for Monitoring Farm Culturing

This study focuses on the underwater video system to monitor growth environment of abalone farm. Figure 1 shows system concept diagram.

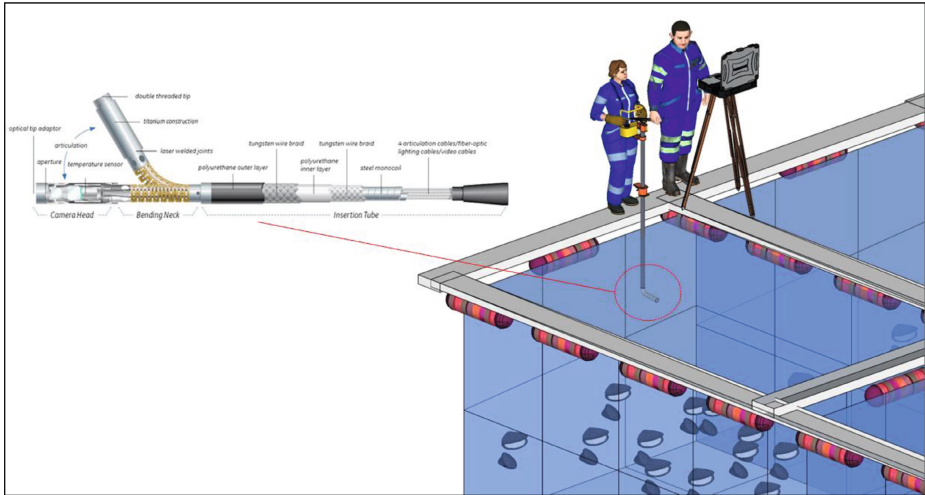


Fig. 1. System diagram of portable video system

This system is portable and can monitor even the blind spot. And it is practical as it can perform monitoring periodically if necessary.

Figure 2 shows system configuration of underwater video system of abalone farm growth monitoring.

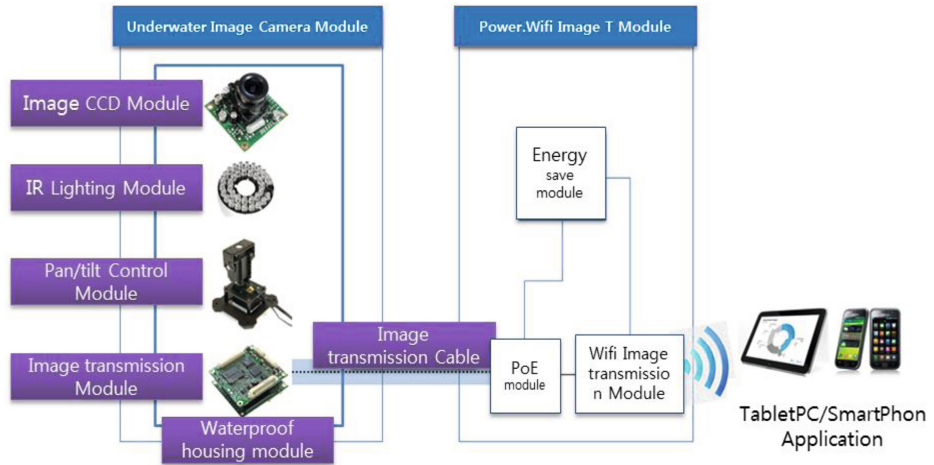


Fig. 2. System configuration of underwater video system of abalone farm growth monitoring

This system comprises video CCD module, IR Led light, Pan/Tilt control module, Video transmission module and water proof housing module. It consists of energy saving module which transmits the video captured by the underwater video camera module and wireless video transmission module. This video system can be displayed and controlled by smartphone or tablet PC.

3 Underwater Video Camera System

Video camera with 2M pixels can monitor the growth status of abalone in 5~10 meter in depth. It is designed with LED with lower illumination not to interfere ecological environment. And the underwater cable module was made to transmit the video from the camera to workers in the sea. Cable is designed to perform video transmission and support camera location.

Table 1. Key module specification of underwater video camera

Module	Specification	
Video CCD Module	Sensor	CMOS
	Resolution	2Mega Pixel
	Codec	MJPEG, H.264, MPEG4
	EEL	4.3mm
	DC volt	DC 12V
	Power	7W
Video Transmission Module	Network protocol	HTTP, TCP/IP,
	Speed	100Mbps
Energy Saving Module	Input power	12V
	Output power	12V, 5V Dual
	Capacity	1000mA
Wireless Transmission Module	CPU	Mediatek RT5350
	Consuming power	5W
	Standard	IEEE802.11n
	Wireless RF	2.4GHz
	Transmission speed	150Mbps
	Transmission power	16dBm \pm 2dB
	Access Method	CSMA / CA
	Transmission method	DSSS, OFDM, Half Duplex

3.1 Pan/Tilt Control Module and Waterproofing Housing Module

PTZ camera was used which can acquire 4 directions video. To avoid blind spot, Pan, Tile function were mounted. Table 2 shows its specification.

Table 2. Pan, Tilt specification

Description	Specification	
Pan	Rotating angle	335°
Tilt	Rotating angle	-30°~ 90°
Protocol	HTTP	

As the abalone growth monitoring system shoots underwater, it should be well waterproofed. It endures max 10 meter depth with 10kg weight of stainless housing module.

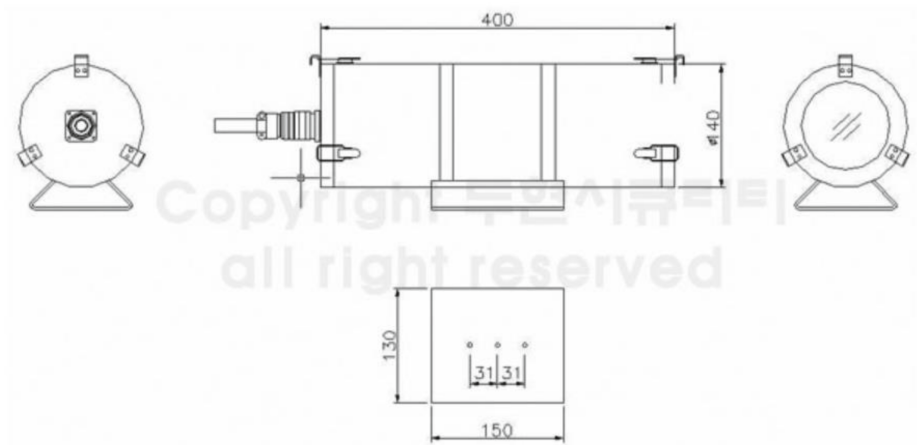


Fig. 3. Stainless housing module

3.2 Power and Wireless Video Transmission Module

Power and wireless video transmission module consists of underwater camera, infrared LED, Pan/Tilt module, energy saving module which is supplied from wireless video transmission module, wireless video transmission module which transmits video to smartphone and table PC, PoE and other switches. The housing of power and wireless video transmission module was made 20 meter length communication cable.



(a) Shape of power and wireless video transmission module

(b) Inside of power and wireless video transmission module

Fig. 4. Power and video transmission module

3.3 Development of Application

Android based smartphone application was made to visualize the video transmitted from the underwater video system. Most of android smartphone released in the country are compatible.

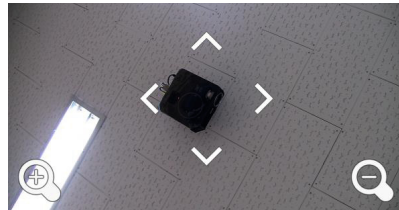
Once the application is executed after connecting to WIFI which belongs to SSID of video transmission module, the video is transmitted simply.

It supports both the table PC and Smartphone without purchasing any other special terminal, which reduces the product price.

The execution result of developed application is as shown in Figure 5.



(a)Initial screen



(b)Execution screen 1



(c) Execution screen 2

Fig. 5. Application execution screen

3.4 Produced Monitoring System

The farm growth monitoring system made with technology developed through this study is as following.



Fig. 6. Developed farm growth monitoring system

4 Results

The abalone farm in Haenam was selected as a subject to develop the portable underwater video monitoring system for growth monitoring. It is a system to control the growth environment of abalone and feeding amount and it is a system which detects the degree of contamination and amount of sludge at the bottom due to over-feeding. This system has advantage to monitor the blind spot which is disadvantage of fixed type. As it is portable, it covers wide range of farm for monitoring.

By performing appropriate feeding amount according to the growth status of abalone using this system, the production of abalone can be increased. By controlling the feeding amount according to the growth status not by experience based feeding, it is expected that the cost and contamination would be reduced. Not limited to the abalone farm, the future research should be made on the system which can monitor in most of farms which can protect serious contamination, red tide or group perish by mounting various sensors.

Acknowledgment. This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2013H1B8A2032217).

References

- [1] Apitz, S.E., Elliott, M., Fountain, M., Calloway, T.S.: European Environmental Management: Moving to an ecosystem approach. *Integrated Environmental Assessment and Management* 2, 80–85 (2006)
- [2] A study on the implementation of sabbatical year of farm 2010. 4, Korea Maritime Institute, the Ministry of Agriculture, Forestry and Fisheries (2010)
- [3] Sik, J.Y., Chan, L.W., Bae, K.J., Jin, H.S., Cheol, K.H., Suk, K.C.: Establishment of environmental assessment criteria of shellfish farms using total organic carbon in sediments and benthic polychaete populations. *Journal of the Korean Society of Marine Environment & Safety* 19, 430–438 (2013)
- [4] Jeong, P.M., Nam, S.J., Wook, P.K., Kon, K.E.: Portable monitoring system for monitoring the growth of abalone. *The Korea Institute of Electronic Communication Sciences* 8(1), 403–405 (2014)

Design of Flight Stabilization System for Acquisition of UAV Based Monitoring Image

Oh-hoon Cho and Eung-Kon Kim*

Department of Computer Science, Suncheon National University,
255 Jungang-ro, Suncheon, Jellanam-do, Republic of Korea
{worldnet, kek}@suncheon.ac.kr

Abstract. Monitoring technique using UAV(Unmanned Air Vehicle) means a technique of monitoring by mounting camera on UAV in case of having to monitor at an altitude where it is unable to shoot objects with general ground photographing equipment. Monitoring method by using UAV has an advantage of accessibility and speed compared with vehicle and monitoring of diversified compositions is possible. In addition, in order to perform UAV-based monitoring, it is very important for monitoring manager to collect image without shaking by manipulating UAV even though such manager does not have any professional knowledge. This thesis suggests a method of research on a design of flight stabilization system for acquisition of UAV-based monitoring image.

Keywords: Small unmanned air vehicle, Flight stabilization, Image acquisition system, Monitoring facilities, Image processing.

Introduction

Recently, at the time of occurrence of forest fire or disaster, site situation is identified by using helicopter in most cases.

However, a research on intelligent type robot by which site situation is monitored in real time by mounting LTE wireless T/R system on UAV is being progressed actively. Multi-purpose watch and monitoring technology for traffic control, collection of weather information, monitoring forest fire prevention and crackdown on illegal architectures in addition to lifesaving, in particular, have been popularized for its advantage of being able to be utilized in diversified fields. Therefore, as importance of collecting stabilized image without shaking and data by manipulating UAV in a wireless mode even though such personnel does not have any professional knowledge or experience in order to perform monitoring by using UAV is being strongly emphasized, a research on flight stabilization system design for acquisition of UAV-based monitoring image was performed in this thesis[1-5].

Goal of this research is to design and suggest flight stabilization system that performs research on moving control module for identifying distance between UAV and monitoring target in real time by using ultrasonic sensor and maintaining pre-set altitude and distance automatically and monitors targets.

* Corresponding author.

Related Research

Definition of UAV-Based Monitoring. UAV-based monitoring is performed by UAV that is mounted with small camera, communication equipment and various sensors or other equipments for operating such camera and sensors being remotely controlled by monitoring personnel on the ground or operated automatically according to pre-set control program. People does not directly get aboard this vehicle but if monitoring personnel controls UAV directly, indirectly by programming control contents for a specific work, UAV identifies arrival location and performs various works of its own accord based on information collected by several sensors mounted on UAV.

As its advantage, monitoring is possible at any place without being bound by space, data securing for monitoring in a short time at reasonable cost is convenient and accessibility to a place where people is hard to approach is allowed as well.

Currently, as a UAV type being mostly utilized for monitoring in the technical field of UAV, there are unmanned helicopter type being frequently used in producing video of film or CF as it is suitable for low, medium altitude and convenient for operating with maneuverability within limited space, electric-powered aerial vehicle that is suitable for photographing wide area such as long shot of wide feeling as its long-distance flight is possible and unmanned airship being utilized for civil T/K construction site and advertising as it is suitable for high altitude shooting and has high stability and in addition to these, paraglider type having relatively cheaper price and high maneuverability.

Among these, multicopter-based UAV utilizing advantage of cost-price, miniaturized helicopter type is being utilized frequently in general.

Type of Multicopter. Multicopter was developed by supplementing disadvantages of helicopter and stable attitude control is possible without rotation of fuselage by backlashing each propeller.

As the more number of propeller is plenty, the more is its power getting strong, it has an advantage that take-off/landing is easy and stable flight control is possible based on its multiple propellers[6].

Multicopter is divided into multicopter and microcopter depending on its size and depending on number of rotor that gains lift and propulsion, it is divided into dualcopter, tricopter, quadcopter, hexacopter and octocopter. And a system for controlling UAV comprises ground base station, communication network and UAV.

Rotating Mode of Blade. In case of general UAV, a disadvantage of having to hover wide space is exposed in order to stay at a certain position in the space but blade plays a role of propeller that enables take-off/landing flight of helicopter and it is able to perform diversified works including hovering and stable flight in a limited space.

Fig. 1 shows overall blade arrangement in a form of quadrangle and in case of blade comprising 6, diagonally positioned blades facing with each other are designed to be rotated in the same direction and neighboring blades inversely.

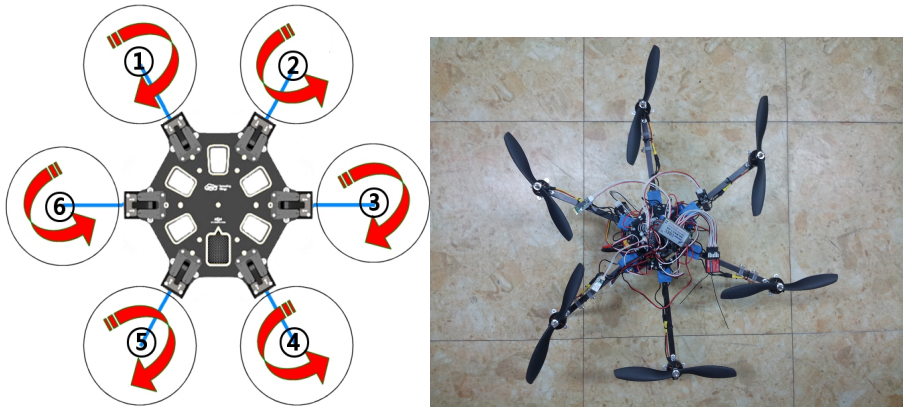


Fig. 1. Rotating mode of blade

At this time, ascending and descending of UAV could be guided by regulating rotating speed power of overall blade depending on rotation mode of blade and by changing speed of front, rear blade (on the assumption that frontal side is X-axis), pitch moment is created and if speed change is regulated by binding left/right and front/rear of roll moment as a pair based on speed change of left/right blade, Yaw moment is generated. In other words, in case of quadcopter like UAV, its operation principle is simple compared with existing helicopter and additionally, as Gyro sensor for attitude control, Terrestrial Magnetism sensor for operator-based control and Barometric pressure sensor for altitude hold are mounted, this vehicle could be controlled skillfully without specialized knowledge for attitude control[7].

In particular, recently, as expensive UAVs are furnished with various diversified sensors, perfect GPS Position Hold, stable flight and excellent gimbal control function, its convenience is being increased further[8-9].

Fail safe function by using GPS Holding technology is a technology of preventing a risk of missing aerial vehicle under the situation that basic T/R signal between aerial vehicle and base station is interrupted and its typical function is to provide aerial vehicle with a function of auto hovering (auto pilot) or of safely going home to and landing at its original take-off place until aerial vehicle is bound by T/R signal again.

Research Contents and Method

Research contents. Biggest advantage of multicopter is that its control is simple than existing helicopter and it is furnished with most of the advantages of aerial vehicle. Hovering function is provided and even in narrow working area like indoor, its positive working performance is possible and without changing direction, omni-directional movement (front/rear, left/right, up/down) is allowed.

This thesis designed flight stabilization system that performs accurate, convenient control of UAV movement in order to obtain precise image of monitoring using UAV.

As concrete R & D contents, in order to perform accurate monitoring by using UAV, a movement control method of identifying distance between UAV and structures in real time through ultrasonic sensor, moving setting distance up/down, left/right under auto

mode and moving and stopping UAV at designated altitude was explored. As its control method, stable flight attitude for maintaining real time distance and levelling was controlled by mounting ultrasonic sensor and by mounting camera, acquisition possibility of high definition image was identified and performance of flight stabilization system was analyzed through site application.

Study Method. Flight stabilization system was designed by collecting technologies relevant to target monitoring and latest trend of flight attitude control and autonomous flight and analyzing relevant theses. and regarding design of aerial vehicle's stabilization system design, a research was performed under the advice of professors of relevant major of Sunchon National University and outside experts and site application and performance analysis were performed under the advice of researchers of participating enterprises.

In this thesis, a method of enhancing attitude control performance of UAV was suggested based on hexarotor type UAV that has advantageous merit in terms of miniaturization and performance compared with its price, enables convenient vertical take-off and landing and of which kinematic structure is simple.

A Study on Flight Stabilization

Ultrasonic Sensor. Ultrasonic sensor is mainly used for measuring distance by using its features of emitting short high frequency pulse for a certain period of time and being transmitted with speed of light in an atmosphere.

Principle of estimating distance from target point is to estimate distance by using time difference between time of emitting pulse from ultrasonic sensor and time of echo signal that returns after colliding with a target object and at present, it is also acknowledged as a new standard technology of automation.

Ultrasonic sensor of microsonic is fit for measuring target distance from 20mm to 10m and it could detect targets under dusty air or ink sprayed condition. In addition, even though film deposition is made on sensor surface, its function is not affected thereby and sensor having blind zone of app. 30mm and that with very narrow beam are able to detect even extremely minute process reliably.

Autonomous Flight Attitude by Utilizing Ultrasonic Sensor. UAV was controlled by identifying its autonomous attitude in a way of utilizing features of ultrasonic sensor like this so that its stable flight is enabled and at the time of its take-off/landing, UAV was controlled by sensing 3cm-3m in a way of utilizing ultrasonic wave of 40KHz in order to ensure its stable take-off and landing.

In order to use ultrasonic sensor, NaviBoard is required. NaviBoard is an Expansion board that plays a role of estimating current inclination angle by using proper attitude algorithms (DCM, Kalman Filter, etc.) by mounting 3-axial Gyro sensor, Terrestrial Magnetism sensor, Acceleration sensor and Barometric pressure sensor in addition to GPS.

By making divergence angle of ultrasonic wave as 0° . it could be realized that inclination measurement control of UAV is possible depending on inclination angle of the ground surface.

Based on an experiment like this, inclination of UAV was made to be controlled by ultrasonic sensor only under the condition of maintaining distance of 100cm from an obstacle like rice paddy and fixing distance of 100cm from the ground surface also.

The system was programmed in a way that 4 sets of sensor were made to control blade motor for 10 times per second by alternately sensing obstacle for each 1 time in 0.1 second so that it may specify distance of front/rear, left/right obstacles after mounting 8 sets of ultrasonic sensor on the lower part of hexacopter center and in parallel, remaining 4 sets of sensor for measuring inclination were made to control blade motor for 10 times per second by alternately sensing obstacles for each 1 time in 0.1 second.

At this time, reliability was enhanced by filtering value below 95cm above ground and measured value over 105cm above ground among the values being measured continuously considering size of UAV of which diameter is 60cm.

Fig. 2 is a system schematic drawing for flight stabilization. Control system of UAV that receives control signal from base station was made to receive data signal of distance, altitude and inclination angle being collected from ultrasonic sensor for measuring distance and inclination with obstacles mounted on expansion board. Therefore, blade motor RPM was made to be controlled by comparatively analyzing constant distance, altitude and inclination of UAV based on programmed control program.

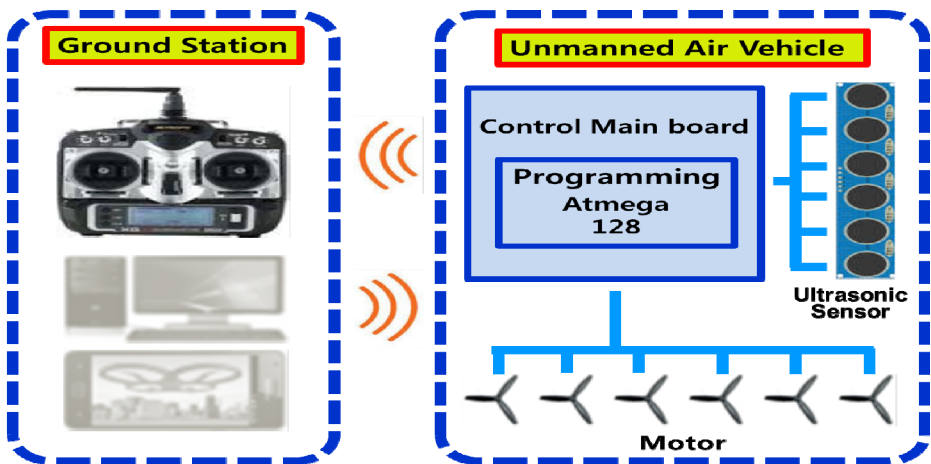


Fig. 2. Schematic drawing of flight stabilization system

Measurement Experiment of Ultrasonic Sensor Inclination. Fig. 3 shows inclination experiment of UAV based on ground surface inclination angle of 0° , 5° , 10° , 15° under the condition that distance from wall obstacles is maintained at 100cm.

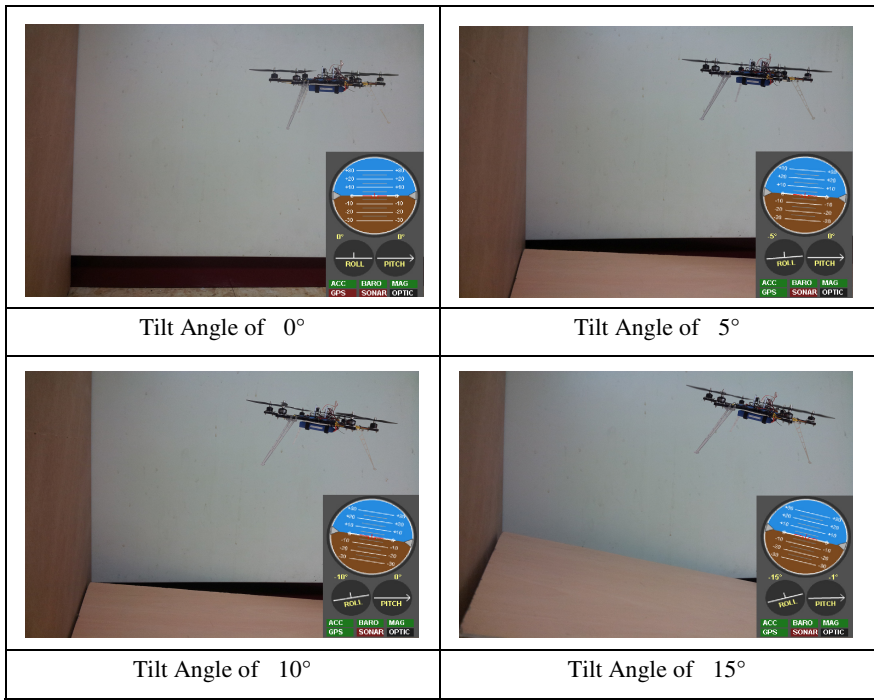


Fig. 3. UAV inclination experiment

Table 1. UAV inclination experiment

Altitude and Obstacle distance of Air Vehicle	Tilt Angle to the Ground	Tilt Angle of Air Vehicle
95~105 Cm	0°	0~1°
95~105 Cm	5°	4~6°
95~105 Cm	10°	9~11°
95~105 Cm	15°	14~16°

In above Table 1, it could be realized that stable attitude control of UAV inclination is possible by UAV inclination experiment based on maintaining distance from obstacles and ground inclination angle by using ultrasonic sensor and therefore, motor rotation speed control of blade could be utilized.

In order to measure accurate angle, there was a working problem that angle between ultrasonic sensors is required to be coincided accurately and reliability for accuracy value may be somewhat lowered due to currently furnished cheap observing equipment but as an experiment was performed through same method of 5 times in order to maintain precision as closely as possible, it is considered that reliability for an experiment objective would be sufficient even with its approximate value.

Blade Motor RPM Control. Blade motor RPM was made to be controlled in a way of controlling flight condition of UAV after sensing distance from obstacles and inclination, altitude of UAV by using ultrasonic sensor. Blade motor RPM was made to be controlled based on value difference being obtained by converting value of ultrasonic sensor into distance value in order to control RPM after measuring it with PWM wave form. In other words, stable autonomous flight of UAV was made to be maintained by controlling RPM of each blade based on distance from obstacles and inclination of UAV.

Conclusion

At present, a lot of UAVs including UAV for military purpose, multicopter type, airship type and paraglider type being frequently used for industrial or hobby purpose and micro type UAV resembling birds or insects have been developed but a lot of limitation are entailed for utilizing them in an actual world.

Typically in case of helicopter, as its mechanical analysis for its automation is difficult, control based on experiment rather than control by formula is attempted. Under this situation, in order to develop aerial vehicle that could be conveniently analyzed based on flight dynamics and enables autonomous flight without difficulty, control experiment was performed by manufacturing aerial vehicle as follows. And as a result of determining control function status through this experiment, sufficient control function was represented and so, it was judged that this test vehicle could be developed further in the future.

As a future research development direction, a method of autonomous flight to destination by bypassing obstacles under autonomous diagnosis and judgment when unexpected obstacle was encountered on the way to destination for monitoring by utilizing ultrasonic sensor was suggested and by performing a lot of sustained researches, flight stabilization and autonomous flight system for acquisition of UAV-based monitoring image are intended to be implemented.

Acknowledgment. This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2013H1B8A2032217).

References

- [1] Verma, A., Sawant, H., Tan, J.: Selection and Navigation of Mobile Sensor Nodes Using a Sensor Network. *Pervasive and Mobile Computing* 2, 65–84 (2006)
- [2] Hwang, J.H., Hwang, S., Hong, S.K., Yoo, M.G.: Attitude Stabilization Performance Improvement of the Quadrotor Flying Robot. *J. ICROS* 18(6), 608 (2012)
- [3] Yoon, S., Soysal, O., Demirbas, M., Qiao, C.: Coordinated locomotion and monitoring using autonomous mobile sensor nodes. *IEEE Transactions on Parallel and Distributed Systems* 22(10), 1742–1756 (2011)
- [4] Mei, Y., Xian, C., Das, S., Hu, C., Lu, Y.-H.: Sensor Replacement Using Mobile Robots. *Computer Communications* 30, 2615–2626

- [5] Abdel-Mageid, S., Ramadan, R.A.: Efficient deployment algorithms for mobile sensor networks. In: 2010 International Conference on Autonomous and Intelligent Systems (AIS), pp. 1–6 (2010)
- [6] Hwang, J.H., Hwang, S., Hong, S.K., Yoo, M.G.: Stabilization performance improvement of the Quadrotor flying robot. In: ICROS Annual Conference 2012, vol. 41, pp. 7–9 (April 2012)
- [7] Lange, S., Sunderhuaf, N., Protzel, P.: Autonomous Landing for a Multirotor UAV Using Vision. In: International Conference on Simulation Modeling and Programming for Autonomous Robots, pp. 482–491 (2008)
- [8] Ruesch: Dynamics Identification & Validation, and Position Control for a Quadrotor. Semester-Thesis, Eidgenössische Technische Hochschule Zurich (2010)
- [9] Chang, S.-H., Choi, S.-W., Koo, S.-O.: Development of the Scaled Vehicle of Smart DAV. KARI 6(2), 236–244 (2007)

The Smart and Secure Protocol for Mobile Office Environments

Hyun-A. Park¹, JongSung Park², Jee-In Kim³, and JaeHyun Park^{4,*}

^{1,3} HCI Lab., Konkuk University/120 Neungdong-ro, Gwangjin-gu, Seoul, Korea

⁴ 2-12-1, Tokyo Institute of Technology, Ookayama, Meguro-ku, Tokyo, Japan
Kokokzi@naver.com, jjong97@gmail.com, jnkm@konkuk.ac.kr,
park.j.ai@m.titech.ac.jp

Abstract. In this paper, we deal with largely two problems in the networked collaborative computing environments between the cloud computing service and TPM chip as a mobile convergent technology. Firstly, we solve the security problem from inside attackers, which has been social issues, for example, 45 % of the attacks are conducted by insiders according to [18]. The substantial reason of this problem is that server managers are regarded trustworthy. Hence, we propose Encrypted DB Retrieval System whose server manager cannot access on real data (plaintexts) in Mobile Office Environments of the cloud datacenter. Secondly, cloud computing has limitless computing resources; however, it faces with the vulnerability of security. On the other hand, the TPM technology has been regarded as a high level of physical security; however, it has the severe limitation of use such as hardware constraints or limited amount of non-volatile memory (NVRAM). To produce synergic effects between the two technologies, we combine two applications (cloud datacenter service, TPM chip) as a mobile convergent technology. Consequently, this system is secure against both of the insiders and outsiders, the cloud computing service can improve security weaknesses.

Keywords: TPM, cloud computing, insiders, collaborative computing.

1 Introduction

In the radically changing computing societies, a variety of information technologies have brought up new types of IT-enabled product and service innovations in our daily lives. The cores of these IT and digital innovations have used highly advanced mobile, cloud, or big data technologies for identifying the innovative solutions in the networked collaborative computing environments. For example, in South Korea, government-led IT policies have established well-organized socio-technical infrastructures using Internet, mobile technologies, and the usages of mobile devices since the last 10 years. Based on this IT infrastructure and increasing usages of smartphone and smart devices, they have currently transformed the information environment paradigm

* Corresponding author.

from wired to wireless or to integrated information environments. Yet, these radically developed IT-driven changes in the shifting paradigm have encountered a dilemma - That is the security problem.

For example, South Korea's three leading credit card companies' customer data were revealed in 2013, which was caused by insiders. Moreover, although current DB systems are encrypted, the 'searching' process is done under the state of 'decrypted'. That is the time attackers try to access or reveal some information from the DB.

As a solution, we propose Encrypted DB Retrieval System that a server manager cannot access on real data (plaintexts) in the Mobile Office Environments of the cloud datacenter. All of the data stored in DB and even querying keywords should be encrypted. There should be no decryption in a server through all processes. In addition, even users do not know their own secret key, but they can decrypt real data by using the masked keys every time. Hence, our proposed system is secure against both of the insiders and outsiders.

Main Methods

- **Security Client:** The role of TPM (Trusted Platform Module) chip - Supporting cloud security vulnerabilities by TPM chip applications. We focus on the importance of converging technologies in order to produce synergic effects between two technologies. In other words, core aspects related to security are run by TPM and additionally PC. This is because the cloud computing service entails limitless computing power as in a pay-as-you-go arrangement [1]; however, it remains in security risks. On the other hand, TPM application technology has a high level of physical security but it has severe limitation of use such as limited command processing function [2,9]. Hence, we use the TPM embedded PC as a Security Client in the Mobile Office Environments.
- **The Masked Keys:** Secret keys should not be released from the TPM chip in PC. This is our Security Goal. To achieve this, we generate a random number every session and mask the real secret keys with that random number. Then, the masked keys can be transferred from the TPM chip and to a terminal device or cloud services. A user or a cloud service manager implements the computations with the masked keys according to the given protocols. Therefore, nobody can know the real secret keys except for the security client (TPM chip in PC). Compared to the values, the results of the masked keys are the same as the results of the real secret keys.

Results and Contribution

- **Similar level of security to 'One time encryption'** – TPM generates a unique random number every time and masks the real secret keys, then computes the results under the encrypted state with the masked keys. Hence, a user/device cannot know the real secret keys. The decrypted results are the same as the ones decrypted with real keys.
- **Secure key management system** - Real secret keys are stored in TPM which provides high level of physical security, and the real secret keys are not be released from the TPM.

- Secure Scheme against both of the insiders and outsiders - There is no decryption in a server, because the server manager does not know anything about secret key in our scheme.
- Our scheme is converging technologies in order to produce synergic effects between the cloud computing service technology and the TPM technology by assembling their technological advantages. Therefore, if our scheme is to be commercialized, the social security problems such as customer data release caused by insider attacks would be solved.

2 Application Scenario and Model

2.1 Application and Participants

Our application scenario is Mobile Office Environments with storage system in the cloud datacenter. In the cloud computing environments, the problem is organizations or individuals cannot trust the datacenter server managers, and these organizations or individuals cannot be sure of data integrity from the unauthorized accesses or attacks. Therefore, they require a security solution, which provides authentication, access control, encryption, and so on. In order to do this, we select TPM as a possible solution approach. In our scheme, TPM acts the important role of a security client as a kind of TTP as well as Integrity Measurements. In this paper, two kinds of clients are defined: a general and a security clients. We grant TPM a role of a security client. Participants of our schemes are made up of 1) a user / a terminal device, 2) TPM chip in PC, and 3) the datacenter. We define a terminal device (TD) as a mobile device. Thus, users can authenticate themselves and can access the cloud computing services.

- 1) User(U_i) / Terminal Device (TD_i) is a general client. A user can access the cloud services by the terminal device and this general client manages all cloud service system processes only except for security parts. Hence, even a user cannot know his own secret key but he only can use his masked keys whose result is the same as the real secret key.
- 2) TPM chip in PC (AR_{si}). TPM chip plays a special role of a security client and PC (AR_{si}) also does the role of “authority of registration” as a TTP (Trusted Third Party). The special tasks involve as follows.
 - TPM chip stores secret keys and all of the information related to the secret keys are given at enrolment time. The authentication information such as PIN and $h(P_u)|P_u$ should be stored. Each attribute secret key and partition Tables for numeric data should be also stored in TPM chip.
 - After a user authentication, TPM makes a legitimate user encrypt data with the masked secret keys and the other information.
 - PC with TPM masks decryption keys with a random number, and transfers the masked keys and the results from the datacenter server to a terminal device.
- 3) Datacenter Server (DS) is the storage server of cloud services. It can be private or public. We assume the server manager is untrustworthy. Thus, all of the sensitive information should be encrypted and there is no decryption process in a datacenter.

2.2 Notations

<ul style="list-style-type: none"> • p_j : each partition for numeric data attribute A_i • PI_j : Partition identifier • p_u : user u's PIN(Personal Identification Number) • $1 m_1^t$: the expression for text data m which is located in the first column(attribute). • $p+q m_{p+q}^n$: the expression for numeric data m which is located in the $(p+q)$-th column(attribute). • t_i : tuple identifier • $1 m_1^t, \dots, p m_p^t, p+1 m_{p+1}^t, \dots, p+q m_{p+q}^t$: the expression for one tuple 	<ul style="list-style-type: none"> • $k_j \in K \in \{0,1\}^k$: keys set for each attribute • f_k : pseudorandom function with secret key k(TI secret key) • K_j^1, K_j^2, K_j^3 : the masked keys with random number to real key k_j • F_j : attribute for text data • P_j : attribute for numeric data • g : a generator of a group G • $g^{k_j \cdot f_k(k_j)} m_{i,j}$: column-level encryption of data m
---	---

3 TPM Controlled Encrypted SQL Protocol in Cloud Datacenter

3.1 System SetUp and Enrollment

Idf is an identification function which maps each partition p_j for a numeric data attribute A_i to an identifier $idf_{A_i}(p_j)$ like this; $idf: A_i \rightarrow idf_{A_i}(p_j) = PI_j$. A mapping function MAP_{B, A_i} maps a value v within an attribute A_i 's domain to the partition identifier PI_j to which the v belongs: $MAP_{B, A_i}(v) = idf_{A_i}(p_j)$. The secret keys for each attribute and partition Tables for numeric data attributes are offered in the user's TPM at enrollment time.

3.2 Authentication and Query_Preparation

$[U_i \rightarrow TD_i]$ 1. Input; $p_u, R(1|m_1^t, \dots, p|m_p^t, p+1|m_{p+1}^n, \dots, p+q|m_{p+q}^n)$: relation.

$[TD_i]$ 2. Compute and Transfer to a TPM chip;

$h(p_u) \parallel E_{p_u}(R(1|m_1^t, \dots, p|m_p^t, p+1|m_{p+1}^n, \dots, p+q|m_{p+q}^n))$

$[TPM\ chip\ in\ PC\ (AR_{si})]$ 3. Verify in PC ; $h(p_u)' = h(p_u)$.

Decrypt in TPM; $D(E_{h(p_u)}(k_j)) = k_j, (1 \leq j \leq p+q)$

generate γ and a tuple identifier t_i in TPM.

Compute in PC ; $K_j^1 = k_j \cdot \gamma, K_j^2 = k_j \cdot f_k(k_j) \cdot (1-\gamma), K_j^3 = g^{k_j(k_j)}$.

Masked keys for a tuple t_i ; $f_k(t_i) K_1^1, K_1^2, K_1^3, \dots, K_{p+q}^1, K_{p+q}^2, K_{p+q}^3$.

In TPM -ESQL, the data which consists of DB Tables are largely classified into two: text data and numeric data. Let R be a relation which has the following attributes set: $R = \{r_1, r_2, \dots, r_n\}$. $1|m_1^t$ is a text data m for the first attribute and $p+1|m_{p+1}^n$ is a numeric data m for the $(p+1)$ -th attribute. $k_j (1 \leq j \leq p+q)$ are secret keys for each attribute. $j | E_{h(p_u)}(k_j) 1 \leq j \leq p+q$; This is the style for stored in TPM.

- 1) A user inputs his/her PIN p_u and the number of n data belonging to n attributes, to TD_i .
- 2) TD_i computes the hash value $h(p_u)'$ and encrypts the n data with p_u .
- 3) $h(p_u)'$ is the received data. $h(p_u) \parallel p_u$ is stored in the TPM. If the verification is satisfied, the TPM decrypts k_j (secret keys for each attribute), generate random numbers γ, t_i . With random number γ , PC computes the masked keys K_j^1, K_j^2, K_j^3 to hide the real key $k_j, (1 \leq j \leq p+q)$.

3.3 DB Encryption

Table 1. Partition Table for the J -th attribute p_j in a relation R

Table 1. Partition Table for the j -th attribute P_j in a relation R	
R.P _j	
Partitions	PI(Partition ID)
[value ₁ , value ₂]	PI _{1,1}
(value ₂ , value ₃]	PI _{1,2}
.....
(value _l , value _{l+1}]	PI _{1,l}

$$idf_{R,P_j}([value_1, value_2]) = PI_{j,1}, \quad MAP_{R,P_j}(value_y) = PI_{j,1}, \text{ if } value_y \in [value_1, value_2]$$

The encrypted relation R in a datacenter server has the following attributes:

- TID: $f_k(t_i)$. This is an encryption of a tuple identifier t_i and k is a TPM’s secret key.
- Column-level encrypted attributes $F_j(1 \leq j \leq \rho + q) : F_j = g^{k_j \cdot f_k(k_j)} m_{i,j}$, where k_j are the secret keys for the j -th attribute and $m_{i,j}$ are the values for the j -th attributes in the i -th tuple. F_j is a text data if $1 \leq j \leq \rho$, and a numeric data if $\rho + 1 \leq j \leq q$. That is, $F_j(1 \leq j \leq \rho + q)$ consist of ρ text data and q numeric data, where equality selections, equi-joins, grouping, etc. are operated.
- Partitioning attributes $P_j(\rho + 1 \leq j \leq q)$: Range queries for numeric data are operated by partition indexes for each attribute value. Namely, $PI_{j,x} = MAP_{R,P_j}(m_{i,j})$, where a numeric data value $m_{i,j}$ is again mapped to a partition index $PI_{j,x}$ (the x -th partition index of the j -th attribute, $1 \leq x \leq l$). Table 1 shows the partition Table for the j -th attribute P_j . Given the above attributes, the relation R^S in a datacenter server is as follows: $R^S(TID, F_1, F_2, \dots, F_{\rho+q}, P_1, \dots, P_q)$.

[TPM chip in PC (AR_{si})]

3. Compute in PC(Column-level Encryption); $(K_j^3)^{K_j^1} g^{K_j^2} m_{i,j} = g^{k_j \cdot f_k(k_j)} m_{i,j}, (1 \leq j \leq \rho + q)$
4. Do Partition and Transfer to DS; $R^S(TID, F_1, F_2, \dots, F_{\rho+q}, P_{\rho+1}, \dots, P_{\rho+q}) = (f_k(t_i), g^{k_1 \cdot f_k(k_1)} m_{i,1} \dots g^{k_{\rho+q} \cdot f_k(k_{\rho+q})} m_{i,\rho+q}, PI_{\rho+1,1}, \dots, PI_{\rho+q,x})$.
- 4) For all data(attributes), the PC does the column-level encryption.
- 5) For the data $(\rho + 1 | m_{\rho+1}^n, \dots, \rho + q | m_{\rho+q}^n)$; numeric data, PC gets the partition indexes $PI_{j,x}, (\rho + 1 \leq j \leq q)$ from the partition Table in TPM and sends the additional “PIs”(partition ID) to DS. If the input data from TD_i are new, TPM generates “PIs” and stores them as a new partition Table and then transfers the query to DS.

3.4 Querying

- [TD_i] 6. Generate; Q(query)
 Compute and Transfer to PC with TPM chip; $h(\rho_u Y | E_{\rho_u}(Q))$

[TPM chip in PC (AR_{si})] 7. Verify; $h(p_u)' = h(p_u)$.
 Decrypt; $D(E_{p_u}(Q)) = Q$.

6, 7) This process shows that TD_i generates ‘Query’ and sends it to DS cryptographically.

Query Decomposition

- ◆ Certain Query Q_c^s : Certain Query selects the tuples which can be assured about the fact that they satisfy the originally given condition or not. Aggregation operations are possible in a server.
- ◆ Guessing Query Q_g^s : The selected tuples to this query cannot be assured about the fact that they satisfy the original given condition unless they are decrypted. Hence, a datacenter server outputs the encrypted values for the attribute without aggregation operations. That is, if the query is for P_{p+j} , a server outputs the matching attribute values themselves, $g^{k_{p+j} \cdot k_{(p+j)}} m_{i,p+j}$ for F_{p+j} in the tuple.

A. Query for Equality Test

[TPM chip in PC (AR_{si})] 8. Compute in PC; $Q^s = Q_c^s = g^{k_{j \cdot k_j} \cdot k_j} v$.

Q expresses various types of queries. If the query is for an equality test, it can be expressed as jlv , where j is an attribute to be searched and v is a text or numeric data. PC should transform this to a query Q^s for a datacenter server. In equality tests, Q^s consists of only certain queries Q_c^s , not guessing queries Q_g^s . A PC encrypts v with the secret key k_j for j -th attribute: $g^{k_j \cdot k_j} v$. This is a query Q^s to a datacenter server.

8) In Equality Test, $Q(= jlv) \rightarrow Q^s = Q_c^s$

B. Query for Comparison Test

[TPM chip in PC (AR_{si})] 8'. Produce(Query Plan); $Q^s = Q_c^s \vee Q_g^s$.

8') Comparison tests are for range or MAX query of numeric data. In our encrypted search system, we use a partition method like Table 1. PC transforms a query Q into a query $Q^s = Q_c^s \vee Q_g^s$ with the partition Table.

3.5 Searching

[DS] 9. Search; $Q^s = Q_c^s \vee Q_g^s \rightarrow$ Result; $R^s = R_c^s \vee R_g^s, R_c^s = \{f_k(t_i), g^{k_j \cdot k_j} m_{i,j}\},$
 $(p+1 \leq j \leq q) \in F_{i,j}$

Transfer to a Security Client; Result.

R_c^s is a final aggregation result for a certain query and R_g^s is a result for processing a guessing query.

[TPM chip in PC (AR_{si})]

10. Operate Aggregations in PC (guessing query); m (value m in Q) \rightarrow PI

Decrypt in TPM; $f^{-1}(f_k(t_i)) = t_i, g^{k_j \cdot k_j} m_{i,j} = (g^{k_j \cdot k_j} m_{i,j}) \cdot g^{-\{k_j \cdot k_j\}} = m_{i,j}$,

$(p+1 \leq j \leq q, \text{ for some } j)$ and process final result.
 Generate in TPM; α .
 Compute in PC; $C_1 = g^{k_j \cdot f_k(k_j)} V_{i,j}, C_2 = g^{f_k(k_j)}, C_3 = -(k_j + \alpha), C_4 = f_k(k_j) \alpha$
 $C_5 = g^{k_j \cdot f_k(k_j)} m_{i,j}$
 Transfer to a TD; $R_c^{TD} = R_c^{TD} \vee R_g^{TD}, R_c^S = R_c^S = \{C_1, C_2, C_3, C_4\}, R_g^{TD} = \{C_5, C_2, C_3, C_4\}$

- 9) DS searches for certain or guessing query.
- 10) For guessing query, at first, the value m in Q is mapped to a partition index PI. Because all the partition Tables are stored in TPM, aggregations are operated in the PC with TPM. If the result of the first aggregation is not assured, the PC can decrypt some expected values(results). Then, finally the PC encrypts the results(certain and guessing query) with the generated random number and masked keys $C_2; C_3; C_4$.

3.6 Decryption

[TD_i] 11. Decrypt; $C_1 \cdot C_2^{C_3} \cdot g^{C_4} = V_{i,j},$
 $C_1 \cdot C_2^{C_3} \cdot g^{C_4} = (g^{k_j \cdot f_k(k_j)} V_{i,j}) \cdot (g^{f_k(k_j)})^{-(k_j + \alpha)} \cdot g^{f_k(k_j) \alpha} = V_{i,j},$
 $C_5 \cdot C_2^{C_3} \cdot g^{C_4} = (g^{k_j \cdot f_k(k_j)} m_{i,j}) \cdot (g^{f_k(k_j)})^{-(k_j + \alpha)} \cdot g^{f_k(k_j) \alpha} = m_{i,j}.$

11) To the given above protocol, we can get real data with masked keys. It is the same value as the decryption of 10 ($g^{k_j \cdot f_k(k_j)} m_{i,j}$), which is used with real secret.

Fig. 1. The whole process of TPM-ESQL.

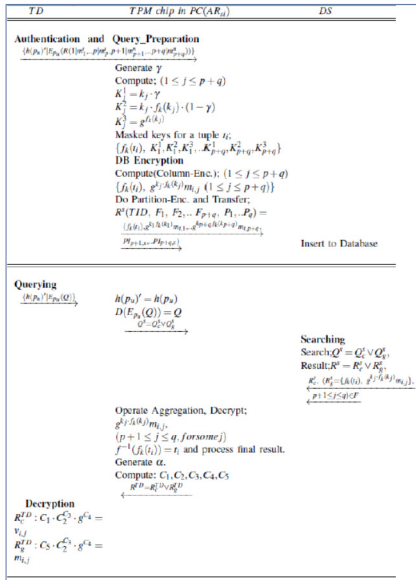


Table 2. PERFORMANCE OF STEPS 1-3

Total Time		376.18	
Mobile phase	83.18	PC phase	293

unit: ms

Table 3. PERFORMANCE OF STEPS 4-5

Total Time (PC Phase)	37.38
-----------------------	-------

unit: ms

TABLE 4. PERFORMANCE OF STEPS 10-11

Total Time		35.9	
PC phase	25.5	Mobile phase	10.4

unit: ms

4 Experiments

This paper deals with many technologies of the combined authentication, access control, and data search. The main purpose of our paper is not for the design of building block algorithms. It is true that the performance of our scheme definitely depends on the subordinate algorithms and how much money you pay for the cloud services, only except for some parts. By these reasons, the comparison and analysis on performance with other papers are inappropriate. We discuss the performance of our scheme with separate six phases;

1. Steps 1-3): Authentication and Query Preparation, 2. Steps 4-5): DB Encryption, 3. Steps 6-7): Querying, 4. Steps 8 & 8'): Query for Equality and Comparison Test, 5. Steps 9): Searching, 6. Steps 10-11): Decryption.

We implement only three parts, i.e., Steps 1-3), Steps 4-5), and Steps 10-11), which have relatively much influence on our performance. In fact, our whole performances are determined much more by environmental factors such as network stability and speed, mobile phone capability, server's computing power other than our proposed scheme. Furthermore, there was no prior study to be compared with our scheme, which considers most of the processes over authentication, access control, searching, private key management, etc. Consequently, respective analysis for each step seems to be quite proper for our scheme.

The actual performance of Steps 6-7), 8&8'), 9) is up to mobile capability, the amount of data which a user wants to search, or environmental factors other than by the schemes we designed newly. Therefore, these parts are excluded from our experiment.

We experiment with both on a PC and a mobile phone. The processing power of our mobile phone is Qualcomm APQ8064 1.5 GHz Quad CPU Core on Snapdragon, 32GB eMMC,

LPDDR 2GB Our personal computer is Intel(R) Core(TM) i5-4570 CPU 3.20 Ghz processor and 8GB RAM and OpenSSL cryptography modules for cryptographic operations. As for the TPM, Infineon OPTIGA TPM SLB 9660 is run in our experiments and the chipset is SLB 9660-1.2, compliant to TPM, 1.2 Rev. 116, LPC interface 24/33MHz.

We set $p=10$, $q=5$, i.e. totally all data $n(p+q)=15$.

- Steps 1-3): The computing time of this step consists of mobile phase and PC phase. As the step for authentication and preparation for queries, masked keys are generated. Because we used common algorithms and commands from OpenSSL cryptography modules and TPM Command Table, the performance depends on the subordinate algorithms.

- Steps 4-5): This step is composed of only PC phase. The step 4 consists of our newly designed protocol for all data's column-level encryption. The step 5 is composed of "getting PI(partition index)" for numeric data in TPM and transferring to DS. The result is shown in Table 3.

- Steps 10-11): This phase is also composed of mostly our newly designed protocols. The step 10 is for the decryption process for guessing query of numeric data and generation of a in TPM, and the masking process of final results with random number a in PC. The step 11 is the final decryption process by masked values C_1, C_2, C_3, C_4, C_5 in terminal device. The result is shown in Table 4.

4.1 Performance Analysis

As shown in the above Tables, the configuration of our protocol has no problem to run in a mobile device. The total time of Table 2 seems to take a little long time (376.18 ms), but this phase is the time for authentication and uploading of takes only 10.4 ms mobile phase as shown in Table 4.

But, in the middle of experiments, we found some problems in processing TPM commands because there are too small kinds of available commands. For example, AES Decrypt does not have so many commands, only except for TPM LoadContext (1024 bytes) in (TPM 1.2 Rev. 116). It shows that we need to make a plan everytime for complex processing configuration to run the TPM embedded PC. Actually, it is expected that the performance would be improved after TPM v2.0 is commercialized.

5 Discussion

5.1 Security

The cloud computing technologies offer a variety of IT infrastructures through extending IT environments outward by their migration policies. As cloud computing technologies develop, security and reliability has become the core issues of the information protection. In the cloud computing environments, the risk of the information leakage is quite high because software, data, and most of the IT resources can be provided as services. Like this, virtualized IT resources make the traditional concept of boundary to protect from attackers “blurred” in the networked information societies. In this section, we seek to solve the security problems of the cloud services with the TPM and encryption as follows; 1. Security Client: We assigned a TPM to the role of high level of security client; 2. Intractability of Decryption by a Server; 3. Entity Authentication at Access Time; 4. Access Control by Encryption and Authentication; 5. Managing Real Secret Keys in the TPM: Only except for the TPM security client, real secret keys for encryption and decryption should not be known to anyone. Even a user does not know his/her own key; 6. Masking Real Secret Keys with Random Number; 7. Data Protection by Encryption.

Consequently, our scheme can prevent abuse or misuse of personal information by a server manager, and it can protect our system from privacy infringement.

5.2 Efficiency

Our goal for efficiency is largely two; 1) less complex processing configuration in the TPM embedded Service System, 2) the minimum computation in a TPM and the maximum computation in a cloud. This is because that the TPM does not have so many kinds of available commands and limited computing power, while the cloud has powerful computing power and resources. The followings show our efficiency.

1. Encryption Process and Search Process

- Column-level Encryption : For all data including text and numeric data, we implement Column-level Encryption in PC. The computation in PC is appropriate for efficiency, because this encryption requires exponential calculation. Table 2 shows that it has no problem. Searching is possible for only Certain Queries in a Server.

- Partition Encryption : This encryption method is used for only numeric data in TPM. Searching is possible for both of Certain and Guess Queries, while Guessing Queries cannot be implemented in the server but the TPM. This method needs additionally storing partition Tables, but its computation is extremely light compared to other schemes such as homomorphism.
2. Efficient Calculation - To achieve less complex processing configuration in the TPM embedded Service System, we implement heavy computation like exponential calculation in a PC or a Server. As to Searching process, only Guessing Queries for numeric data should be implemented in the TPM chip embedded PC.
 3. Total Performance : It depends on the amount of data stored in the DS and the amount of the price we pay for Cloud datacenter services. This is because most of the heavy computations are implemented in a datacenter server.

6 Conclusion

As a result, this paper includes three contributions as follows: (1) it highlights the importance of security issues in radically developing technological cloud computing era by expanding the coverage of security more broadly; (2) it identifies a future converging technological solution by assembling the advantages between the cloud computing service technology and the TPM technology; (3) it identifies the potential issue of implementing Integrated Security Management (ISM) systems rather than those limited to database security. For the future research, the collaborative computing environments among establishing, emerging, and future computing might offer a variety of research potential for the communities of security and mobile computing. Therefore, we need to consider the security issues as the core of the collaborative computing environments over time.

References

1. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M.: Above the Clouds: A Berkeley View of Cloud Computing White paper (2009)
2. Chen, C., Raj, H., Saroiu, S., Wolman, A.: cTPM: A Cloud TPM for Cross-Device Trusted Application. In: 11th USENIX Symposium on NSDI 2014, Seattle, pp. 187–201 (2014)
3. Cheon, J., Kim, W., Nam, H.: Known-plaintext cryptanalysis of the Domingo-Ferrer algebraic privacy homomorphism scheme. *Information Processing Letters* 97, 118–123 (2006)
4. Ferrer, J.D.: New privacy homomorphism and applications. *Information Processing Letters* 60(5), 277–282 (1996)
5. Ferrer, J.D.: Provably Secure Additive and Multiplicative Privacy Homomorphism. In: Chan, A.H., Gligor, V.D. (eds.) *ISC 2002*. LNCS, vol. 2433, pp. 471–483. Springer, Heidelberg (2002)
6. Gregg, M.: 10 Security Concerns for Cloud Computing. *Expert Reference Series of White Papers* (2010)
7. Hacigumus, H., Iyer, B., Li, C., Mehrotra, S.: Executing SQL over Encrypted Data in the Database - Service-Provider Model. In: *ACM SIGMOD*, pp. 216–227 (2002)

8. Hacıgümüş, H., Iyer, B., Mehrotra, S.: Efficient Execution of Aggregation Queries over Encrypted Relational Databases. In: Lee, Y., Li, J., Whang, K.-Y., Lee, D. (eds.) DASFAA 2004. LNCS, vol. 2973, pp. 125–136. Springer, Heidelberg (2004)
9. Kotla, R., Rodeheffer, T., Roy, I., Stuedi, P., Wester, B.: Pasture: Secure Offline Data Access Using Commodity Trusted Hardwar. In: 10th USENIX Symposium on Operating Systems Design and Implementation (OSDI 2012), Hollywood, pp. 321–334 (2012)
10. Mykletun, E., Tsudik, G.: Aggregation Queries in the Database-As-a-Service Model. In: Damiani, E., Liu, P. (eds.) Data and Applications Security 2006. LNCS, vol. 4127, pp. 89–103. Springer, Heidelberg (2006)
11. Park, H., Hong, J., Park, J., Zhan, J., Lee, D.: Combined Authentication based Multi-Level Access Control in Mobile Application for DailyLifeService. *IEEE Transactions on Mobile Computing* 9(6), 824–837 (2010)
12. Park, H., Zhan, J., Blosser, G., Lee, D.: Efficient Keyword Index Search over Encrypted Data of Groups. In: *ISI 2008*, pp. 225–229 (2008)
13. Qayedi, A.A., Adi, W., Zahro, A., Mabrouk, A.: Combined Web/Mobile Authentication for Secure Web Access Control. *Wireless Comm. and Networking Conf. 2*, 677–681 (2004)
14. Riccia, R., Chollet, G., Crispino, M.V., Jassim, S., Koreman, J., Dimas, M.O., Salicetti, S.G., Rodriguez, P. S.: SecurePhone.: A Mobile Phone with Biometric Authentication and e-signature Support for Dealing Secure Transactions on the Fly. In: *SECRYPT*, pp. 9–16 (2006)
15. Song, D., Wagner, D., Perrig, A.: Practical techniques for searches on encrypted data. In: *IEEE Symposium on Security and Privacy*, pp. 44–55 (2000)
16. Wagner, D.: Cryptanalysis of an Algebraic Privacy Homomorphism. In: Boyd, C., Mao, W. (eds.) *ISC 2003*. LNCS, vol. 2851, pp. 234–239. Springer, Heidelberg (2003)
17. <http://searchcloudapplications.techtarget.com/feature/Cloud-migrationstrategy-Consider-portability-security-overall-risk>
18. Computer SecurityInstitute, CSI/FBI Computer Crime and Security Survey, <http://www.gocsi.com/forms/fbi/pdf.html>

Research for Personalized Recommendation of Learning Resource in Mobile Computing Context

Lina Yang^{1,2,*} and Yi Yang²

¹ School of Educational Technology, Beijing Normal University, Beijing, 100875, China

² School of Communication, Tianjin Foreign Studies University, Tianjin, 300204, China
{yang_lina,yangyi_vivian}@163.com

Abstract. A technology for mobile internet of new generation provides seamless support for mobile learning. With the growth of its resources and services, information overload is the main factor disturbing the improvement of mobile learning performance. From the perspective of personalized service, in the light of personalized needs from learning resource, contextualization, and intellectualization of mobile learning, this paper builds a model of resource recommendation for mobile learning, and analyzes key technologies for personalized resource recommendation in mobile computing environment.

Keywords: Mobile Learning, Personalized Service, Resource Recommendation.

1 Introduction

A mobile internet of new generation, especially the popularity of 3G Network, provides actually technical support for establishing mobile learning environment. With the universal application of intelligent mobile devices, it has become possible for users to achieve any information resource at “any time”, “any place”, and in “any way” [1,2]. However, the growth of mobile learning resources will exceed people’s ability in processing information. Besides, ability of mobile devices, such as interface displaying, terminal processing and input/output, is limited, which will cause “mobile information overload” and bring cognitive burden with users in mobile computing environment [3].

Technologies for personalized recommendation, as an effective way solving information overload, have been widely used in e-commerce, e-health and online education. Theories and key technologies for personalized services make some progress in desktop computing, however, compared with traditional internet, the process in mobile computing is much more complicated. Due to mobile terminal’s limitation, needs from mobile users tends to more accurate, personalized, real-time and intelligent, which complicates personalized recommendation process [4]. Strategies for traditionally personalized recommendation service are mainly based on two-dimensional “user-item” and recommendation models are decontextualized, irrespective of other factors influencing users’ preferences identification and recommendation generation,

* Corresponding author.

which is insufficient for the emerging needs of personalized service and relieving mobile information overload[5]. This paper caters to the needs of contextualization, personalization and intelligence for mobile learning. Adapting strategies for personalized service, the paper builds the personalized recommendation model for mobile learning, designs the recommendation strategies, including user preference model, recommendation algorithm, and aims to provide references for personalized service in mobile computing environment.

2 Personalized Recommendation Modeling for Mobile Learning

Following the service-driven trend of mobile internet, aiming to accurately choose and recommend timely and moderate mobile learning resource, there are three aspects of problems to be solved:

First, recognizing and acquiring preferences from mobile learners. It is the base to identify accurately users' preferences mode for personalized recommendation[7]. Different from the case of mobile computing, technologies for preference identification applied in desktop computing cannot be easily transferred to the scenarios of mobile computing. It needs to be redesigned for mobile learners' preferences mode, according to the environmental features of mobile learning, resource organization and mode of transmission.

Second, modeling technologies for mobile learner preferences. Most of existing personalized recommendation systems adopt user modeling technology based on their behaviours[8]. It represents only some aspect of preferences for resources from user's explicit behavior. This paper proposes the context-aware modeling method and identification strategies for user preferences based on multi-data fusion. See the section 3.1 for detailed contents.

Third, personalized recommendation algorithm. Recommendation algorithm is a technical guarantee for the aim of "service provision on-demand and resource recommendation in time"[9]. Many mainstream recommendation algorithms, including collaborative filtering, recommendation based on content and recommendation based on association rules, can't be used simply in the case of mobile learning. In the case of mobile computing, it is a major issue to consider seriously needs change, integrate and redesign recommendation strategies.

According to these three aspects of problems mentioned above, combining the characteristics of mobile learning and personalized needs of users, this paper builds a model of personalized resource recommendation for mobile learning (shown in figure 1). More details about the model are the following aspects:

(1) Identification of User Preference and Learning Context

Identifying accurately learning preference is the key to implementing personalized resource push. In mobile learning context, from the perspective of user behavior, user model can be generated by analyzing user past learning records, learning activities (exercise, test, discussion, etc.) and preferences for resource categories (text, video, audio, etc.). Contextualization is the essence for mobile learning, its influence on learning preferences cannot be ignored in the case of mobile learning[10]. The paper

explores mobile learning context mainly from mobile learner’s learning group, user category and learning process. User’s learning group refers to social cognitive network built through interactive process based-on resources and services in mobile learning. User categories include of new and old users, and different recommendation strategies are needed for them. Learning progress is used to describe the learning rhythm of mobile users. And social context(labeled by learning group), user context(labeled by the categories of user)and state context(labeled by learning progress)are highlighted, which aims to emphasize the significance of learning context.

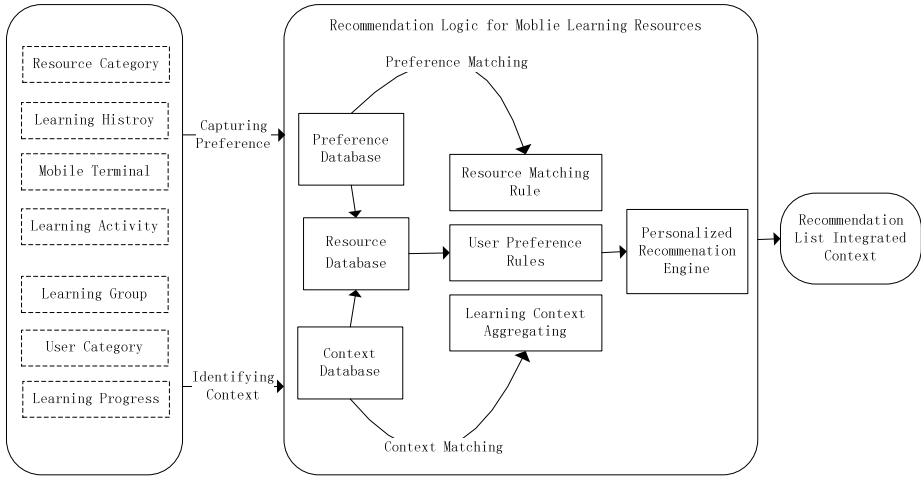


Fig. 1. A model of Personalized Resource Recommendation for Mobile Learning

(2)Implementation of Personalized Recommendation Logic for Mobile Learning Resource

There are three basic problems need to be discussed for resource recommendation in mobile computing context. Firstly, it is the key to build resource database with the tag of context-ware for mobile learning on the basis of preference and context database served the need of personalized recommendation in mobile computing. Learning resource can be labeled basic context by designing resource tags. To generate contextual learning resource database, in the process of identifying user preferences and building contextual database, the resources accessed by learners need to be labeled with context tag, which can provide evidence for contextualized learning resource recommendation. Secondly, the design of resource matching rules for integrating learning preferences and context. To embody the contextualized feature of mobile learning, learning preferences and contextual information needs to be coupled into the matching rules of learning resources, which is distinctively different from the case applied in desktop computing. Thirdly, the design and implementation of personalized recommendation engine. The main function of recommendation engine is to analyze and calculate user’s interest in resource and service, and call relevant recommendation algorithm to implement personalized recommendation by analyzing user preferences model[11].

(3) Personalized Recommendation and Presentation of Mobile Learning Resource

Personalization, contextualization and intelligence are the key to implementing personalized service for mobile learning, so the presentation of source should closely match user's preference as well as have contextual adaptability[12]. Personalized resource recommendation provides users with more choices, including recommending learning resource and service, learning group, and social cognitive network implied in accessing learning resource and service, which all can enrich mobile user's learning experience.

3 Key Technologies for Personalized Resource Recommendation

3.1 Technologies for User Preference Modeling

Identifying user's resource preference is the basis of personalized recommendation. To satisfy user's need for learning resource and service, the following four aspects are involved in the user preference modeling.

1. Preference acquisition based on characteristics of user behavior

In mobile learning context, user's learning behavior may imply potential preference mode. Building user preference model based on behavior characteristics has become a commonly used strategy in mobile computing context. The typical learning behavior includes: resource scanning, collection, study, download and past learning behavior. Identifying and analyzing mobile learning user's various learning behaviors contribute to explore user's potential resource preferences, which is the basic strategy of satisfying user's personalized resource need.

2. Preferences acquisition based on context-awareness

Context-aware recommendation can provide timely and suitable resource services. The mobility of leaning makes learning context change with the swift of position and demand. To realize personalization of mobile learning resource recommendation, key learning contexts need to be identified and analyzed. The key context of mobile learning can be represented by user's terminal category, resource format, learning group and learning progress. Due to dynamic changes of mobile context, all these aspects need to be integrated into user preferences model, which further improves the accuracy and pertinence of resource recommendation of mobile learning.

3. Preference acquisition based on multi-source data fusion

User's learning preferences are constantly implied in the interaction between user's behavior and context. To meet user's all-round needs for mobile learning service, multi-source factors should be considered to identify accurately user's potential preferences, such factors as learning behaviors, learning context and learning interaction. It is essential for personalized recommendation in mobile learning context to explore mechanisms of multi-source data fusion.

4. Evolution mechanism of user preference

Transfer of mobile learning context determines the variability of user preference. The key to achieving the aim of “service provision on-demand and resource recommendation in time” is to identify and analyze the main factors influencing user learning preferences and preference evolution, and study evolution mechanism of user preference according to the relationship between user behavior and its evolution, learning context and user behavior’s evolution.

3.2 Personalized Recommendation Algorithm

To accurately and self-adaptively satisfy user’s need for learning resources, the following strategies can be adopted comprehensively for personalized recommendation in mobile learning context through combining user learning behavior and environmental characteristics.

1. Collaborative Filtering Recommendation

Collaborative filtering(CF) is a typical recommendation technology, which includes of two categories, one is user-based, the other is item-based. For the former, due to similar calculation of user grading data, sufficient user grading data is enough through adopting vector similarity method to calculate neighbour and to recommend neighbour’s preferences to target users, not considering recommendation object’s content. The later makes prediction and recommendation through computing the similarity between programs (product, resource and service). The essence of CF is “guess or prediction”, recommending relevant information about past preferences and novel recommendation, which satisfies potential personalized information need of target users.

2. Content -based Recommendation

Its most prominent characteristic is to recommend information and service similar to past interest and preferences according to target user’s interactive behavior happened in the past, while not providing novel recommendation.

As a learning system, it just explores user’s personalized characteristics, and doesn’t need to match that of other users. To implement this strategy, it is needed to analyze resources accessed by users and form a matching mode base between users and resources, including user characteristics base(e.g. learning style, learning strategy, learning preference, learning state)and resource characteristics base(e.g. subject, , page view, evolution force). As learning resources grows, the system analyzes the bank of target users and pushes the matching resource.

3. Context-aware Recommendation

There are three kinds of strategies for context-aware recommendation. The first is contextual pre-filtration strategy. The strategy leaves out users’ preference data from traditional two-dimensional rating matrix uncorrelated to the service context, and then takes advantage of typical collaborative filtering algorithm to operate the remaining data associated with services context, finally pushes contextual resources recommendation to the target users. The second is contextual post-filtration strategy. The strategy makes use of

traditional two-dimensional rating matrix indicating users' preferences to generate recommendation list, and then uses aggregated contextual information to filter resources irrelevant to current context from the recommendation list, finally pushes resources related to service context. The third is multi-dimensional context-aware strategy. The strategy integrates actual contextual information into existing recommendation algorithm, through reducing dimension, which finally generates contextualized recommendation list.

Umberto Panniello•Michele Gorgoglione examined empirically the effectiveness of these three kinds of strategies[6]. The result is that these strategies are significantly superior to the case used in scenarios without considering service context. And multi-dimensional context-aware strategy has the best performance in pushing sources services. Contextual post-filtration strategy is in the next place. In mobile computing context, much attention should be attached on learning experience of users, among of which, response speed of environment is an important factor affecting user experience. The contextual post-filtration strategy performs well in response speed, To acquire good mobile learning experience, contextual post-filtration strategy is adopted by this paper to carry out contextualized learning resource recommendation.

4 Conclusions and Further Research Directions

Mobile learning, as extension of e-learning, is the result of interaction between information technology and educational need. Personalized recommendation services for mobile learning resources adapts inevitably to the trend of mobile learning: personalization, intelligence, contextualization and self-adaption. This paper carries out a framework study for personalized recommendation model of mobile learning resource, including user modeling technologies and personalized recommendation strategies. Compared to the rapid development of personalized service for mobile learning, much work need to be done for further research for personalization of mobile learning resources, such as evolution mechanism of learning preference, self-adaptive resource recommendation, and improving effectiveness of personalized recommendation, etc..

Acknowledgments. This research is supported both by Humanities and Social Sciences Foundation of Chinese Ministry of Education (Grant No.13YJCZH225) and Chinese Postdoctoral Science Foundation (Grant No.2013M530540).

References

1. Yu, Z., Nakamura, Y., et al.: Content Provisioning for Ubiquitous Learning. *IEEE Journals & Magazines*, 62–70 (2008)
2. Yap, G.E., Tan, A.H., Pang, H.H.: Discovering and Exploiting Causal Dependencies for Robust Mobile Context-aware Recommenders. *IEEE Transactions on Knowledge Data Engineering* 19(7), 977–992 (2007)
3. O'Keeffe, I., et al.: Just-in-Time generation of pedagogically sound, context sensitive personalized learning experiences. *International JI on E-learning* 5(1), 113–127 (2006)

4. Luo, J., Dong, F., Cao, J., Song, A.: A context-aware personalized resource recommendation for pervasive learning. *Cluster Comput.* 13, 213–239 (2010)
5. Yang, S.J.H.: Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning. *Educational Technology & Society* 9(1), 188–201 (2006)
6. Umberto Panniello-Michele Gorgoglione. Incorporating context into recommender systems
7. An empirical comparison of context-based approaches. *Electron. Commer. Re.* 12, 1–30 (2012)
8. Hu, M., Cai, S., Zhang, Y.: Research on Contextualized User Preferences Towards Personalized Recommendation. *Journal of Information* 29(10), 157–163 (2010)
9. Hu, M., Cai, S.: The Extending Study on Personalized Recommendation Method of Loose-coupling Context. *Library and Information Service* (2), 371–377 (2010)
10. Yang, C., Li, H.: Study for Resource Recommendation Based on Personalized Context and Program Category. *Computing Science* 38(10A), 175–179 (2011)
11. Li, P.: Study and Application of Resource Recommendation Based on Pragmatic Context. Chongqing University, Chongqing (2011)
12. Zhang, Q., Zhang, Y.: Study on Collaborative Recommendation Method of Context-aware Scientific and Technical Literature. *Digital Library* 2, 10–18 (2012)

Terrain Classification Using Adaboost Algorithm Based on Co-occurrence and Haar-like Features

Ngoc-Hoa Nguyen and Dong-Min Woo*

Department of Electronics Engineering, Myongji University
San 38-2 Namdong, Cheoin-gu, Yongin, Gyeonggido, 449-728 South Korea
ngochoa9@gmail.com, dmwoo@mju.ac.kr

Abstract. Terrain classification is still a challenging issue in image processing, especially with high resolution satellite images. Specific obstacles include low accuracy in detection of targets, high computation time, and the need for specific algorithms for specific areas. In this paper, we present an approach to classify and detect building footprints, foliage areas, road-grass, and bare ground in a grayscale satellite image (2048x2048). Our contribution is to build a strong classifier using Adaboost based on a combination of co-occurrence and Haar-like features. The Adaboost algorithm selects only critical features and generates an extremely efficient classifier. The combination of two feature extraction decreases the training time and improves the classification accuracy. The accuracy of the proposed method is quite high: over 98.4% for classification and more than 92% for target detection on high resolution images.

Keywords: Texture Feature, Co-occurrence, Haar-like, Adaboost, Terrain Classification, DEM, Satellite Images.

1 Introduction

With the development of remote sensing technologies, automatic object classification from satellite images has become an essential component of many applications. Automatic object extraction is primarily needed in Geographic Information System (GIS) applications, since manual data acquisition for updating the GIS database is costly and time consuming task. GIS applications are widely used in many fields such as the defense industry, transportation systems, emergency management, urban planning, navigation applications, etc. In response to the development trends in the field of satellite image processing, numerous approaches have been proposed to address issues related to accuracy, computation time. However, an optimal algorithm for image processing in general and terrain classification in particular has yet to be proposed. Instead, the choice of algorithm depends on the type of image, what the user is seeking, and the required accuracy. This topic thus warrants further study, and a new combination of feature extraction techniques should be considered.

* Corresponding author.

This paper presents a terrain classification approach based on Adaboost algorithm and the combination of feature extraction to detect multiple objects which include buildings, foliage, bare ground, and grass-roads. Although these objects have some distinguished characteristics such as shape, texture, and color features, the classification task is still difficult due to the complexity, quantity of objects, and quality of the satellite image. For example, buildings and roads have a wide scope of size and shape. It is a difficult issue in non-homogeneous scene. Color feature is very useful information to distinguish among objects of terrain. However, the color feature is omitted because we use a grayscale satellite image. Knowing the above limitations, we seek to design an efficient algorithm which satisfies the criteria such as high accuracy, low computation time, multi-classification. It is also works well with complex data.

We describe a method to extract relevant features from the satellite image so that the Adaboost algorithm can successfully train on a small data set, and classify a larger image. This work utilizes co-occurrence matrix, elevation data to extract 2D and 3D co-occurrence features [1]. Haar-like features are extracted from grayscale image [4]. We then use these features to make a descriptor features training set. A multi-class Adaboost is built based on the idea of Freund and Schapire [5] to classify and detect. The reason why we choose Adaboost and the above feature extraction is discussed in section 3.

The remainder of the paper is organized as follows. Section 2 surveys briefly the related work in the area of terrain classification. Section 3 outlines an overview of our approach. Section 4 presents our results. Lastly, section 5 gives a summary and our future work.

2 Related Work

For decades, many approaches have been proposed and achieved notable successes in solving terrain classification. In this section we present only work relevant to supporting our research. Terrain Classification is a very useful need in life. There has been a tremendous increase in demand for applications to monitor objects related to remote sensing areas. Advances in field remote sensing have promoted robust investment in research in this area. In 1973, Haralick has studied in this area. He proposed an image classification approach using co-occurrence features for descriptor vector of samples, and two classification algorithms (Piecewise Linear and min-max decision rule) [1]. Additionally, Haralick introduced 28 textural features which can be extracted from co-occurrence matrix, and is used so far. In 2003, a terrain classification was presented. The method used neural network algorithm to classify remote sensing data from urban areas [6]. The author uses two feature extraction (decision boundary feature, and discriminant analysis feature) to analyze performance in order to reduce the computational time. A comparison of classifiable performance was also done between co-occurrence features, Gabor filters, and MRFs [11]. It turns out that the co-occurrence features achieved the highest classification rate.

In 2007, a method of terrain classification was proposed [2]. Neural Networks and 3D co-occurrence feature was used to classify multiple objects like grass, road, shadows, and foliage. From the experiments, the 3D co-occurrence feature was evaluated well, and it can effectively reflect physical characteristics of the surface. At the same

time, another approach was introduced to classify vegetation areas from high resolution images in urban areas [7]. This approach uses a linear-kernel SVM with a four dimensional radiometric feature vector to identify vegetation areas.

In recent years, the aspects of terrain classification are still interested. In 2012, Lei Zhang presents a terrain classification approach based on classification ontology to deal with complex multi-class [9], the accuracy of classification around 85.5%. An approach for terrain identification was proposed in 2010 by Abou-Nasr [10]. This approach is a binary classifier using Recurrent Neural Networks.

3 Proposed Approach

The goal of the approach is to find a new solution to improve the accuracy and reduce the computation time of the terrain classification process in the high-resolution satellite images. To address the objective of this research, we focus on two major parts of the image classification issue: feature extraction and learning algorithm. In the feature extraction part, we have a lot of feature extraction. It is very difficult to assess which one best so far. Each feature at least has a good point in a specific problem. Collating the objective of terrain classification, the intention is to classify objects and assign an appropriate label for each pixel in the grayscale satellite image. Texture feature is considered as a feasible choice because it has been extensively used in classification of remote sensing images. From some famous texture feature such as Gabor, Co-occurrence, Fractals, Markov Random Fields (MRFs), we choose co-occurrence feature as a suitable feature for the following reasons: Most of texture analysis methods have been design for homogeneous textures. It is suitable for terrain classification. Co-occurrence feature is assessed better than Gabor, MRFs in the terrain classification [10]. Co-occurrence feature is considered very effective in terrain classification [2]. Due to use the grayscale image, the color feature does not make sense in this case, and many previous works use co-occurrence feature for terrain classification [1, 2, 3, and 11].

In the method developed by Haralick [1], twenty-eight textural features are suggested. These features are computed for each of the four directions: 0, 45, 90, and 135 [2]. The distance parameter Δ can be selected as one or higher. In this paper, we use a unit distance, four directions, and three textural features: Angular Second Moment (ASM), Contrast (CON), and Entropy (ENT) that are introduced to generate co-occurrence features. Among these, ASM gives a strong measure of the uniformity of the data, CON provides evidence of how sharp the structural variations in the image area, and ENT provides a measure of the complexity. We use co-occurrence features to solve our problem with a combination of two dimensional and three dimensional domains with the aim of generating a composite texture. The different between 2D and 3D co-occurrence feature is the input image to calculate the co-occurrence matrix. If the input image is grayscale image, the calculated features are 2D. If the input image is digital elevation modal (DEM), the calculated features are 3D.

Co-occurrence features work well when data is homogeneous and similarity. But building class is non-homogeneous because of a wide scope of size. We need to point out at least one more feature to fix this problem. After reviewing (including testing) several features, we decided to choose Haar-like feature for the following reasons:

First, Haar-like feature includes corner features and rectangle features, and this feature can be used to generate a building descriptor. Second, Haar-like feature can be easily calculated with an integral image.

A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region, and calculates the difference between the sum of dark regions and the sum of bright regions [4]. This difference is then used to categorize subsections of an image. In order to improve the accuracy of the object detection rate, an extended Haar-like method was introduced to increase the number of features [8]. In this paper, we use twenty-one Haar-like features for extraction of each subsection.

In the learning algorithm part, there are many algorithms can be used to classify multiple object such as Neural Networks, SVMs, Adaboost, etc. In this work, we only select a good enough algorithm, which can satisfy the objective high accuracy, and low computation time. The Adaboost learning algorithm is a supervised learning, simple and easy to program, fast, no parameter to tune, no prior knowledge needed about weak learner.

In its original form, the Adaboost learning algorithm [3, 4] is used to boost the classification performance of a simple learning algorithm. It combines a set of weak classification functions to form a stronger classifier. Each weak learner determines the optimal threshold classification function, such that a minimum number of samples are misclassified. In this training process, the error rate of the current weak classifier is reduced based on updating of the previous classifier. The weights of the training set determine the probability of being selected for a feature and they are continuously updated via every weak classifier.

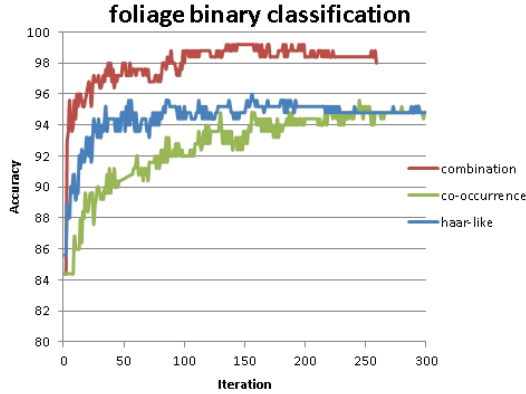
A multi-class problem is a more practical case than binary classification. As we mentioned before, the purpose of this paper is to distinguish between buildings, foliage, bare ground, and grass-roads. Our solution is similar to “One vs. All” classification [12]. The results of the “One vs All” are always a square matrix with size $K \times K$ (K is the number of binary classifier). It looks like the table of results in section 4.

4 Empirical Results

A dataset is manually collected original image with 960 samples (32x32) for training, and 250 samples (32x32) for testing. We use a set of 45 feature values (12 2D co-occurrence features, 12 3D co-occurrence features and 21 Haar-like features) as input pattern for the Adaboost algorithm. The following is results of multi-class classification: The accuracy is calculated by total correct samples divided by total samples:

Confusion Matrix for Multi-Class Adaboost Algorithm						
Accuracy: 98.4%		Predicted				
		Bare ground	Foliage	Buildings	Road-Grass	Total
Ex-pected	Bare ground	49	0	0	1	50
	Foliage	0	49	0	1	50
	Buildings	0	1	48	1	50
	Road-Grass	1	0	0	99	100

The combination of co-occurrence and Haar-like feature, a binary training process of each class achieved higher accuracy with same number of iteration. In other words, the binary training processing just needs a little time to gain high accuracy. Following chart is a testament to show that we can achieve high accuracy with 100 iterations.



To evaluate the results of segmentation, we create a ground-truth image manually from original image that correctly shows the labels of its area. Each area is a rectangle with a label correctly. We then compare the similarity of the label of each rectangle pixel by pixel between the ground-truth image and the segmented image.

Sample	Number of correct pixels	Total number pixels	Accuracy (%)
Buildings	40,373	59,619	67.%
Foliage	67,304	75,052	89.%
Bare ground	173,967	174,050	99.5%
Grass-Road	66,648	66,696	99.%
Total	348,292	375,417	92.8.%

5 Conclusion and Feature Work

In this work, we have presented a combined approach for classification and detection of satellite images. The method begins with the extraction of co-occurrence and Haar-like features. We then manually generated a descriptor vector for each object in data training. Next, we built the multi-class Adaboost algorithm to train and make a prediction function based on the strong classifiers. Finally, the prediction function was used to classify and detects object in the satellite image. Experimental results of our dataset show that, the proposed method improved the accuracy of classification, and reduced the training time. However, the proposed method cannot separate the grass-road class. In future work, we will try to separate grass-road into two classes and increase the accuracy of building detection by using color feature. We will also test the effectiveness among the famous learning algorithms.

Acknowledgment. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2012R1A1A2004950).

References

1. Haralick, D.: Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics* SMC-3(6) (November 1973)
2. Woo, D.-M., Park, D.-C., Nguyen, Q.-D.: Terrain Classification Using 3D Co-occurrence Features and Neural Networks. *International Journal of Information Acquisition* 4(4), 1–9 (2007)
3. Clausi, D.A.: Co-occurrence matrix and its statistical features as a new approach for face recognition. *Turk. Journal Elec. Eng. & Comp. Sci.* 19(1) (2011)
4. Viola, P., Jones, M.J.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57(2), 137–154 (2004)
5. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* 55(1), 119–139 (1997)
6. Benediktsson, J.A., Pesaresi, M., Arnason, K.: Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 1940–1949 (2003)
7. Iovan, C., Boldo, D., Cord, M., Erikson, M.: Automatic extraction and classification of vegetation areas from high resolution images in urban areas. In: Ersbøll, B.K., Pedersen, K.S. (eds.) *SCIA 2007*. LNCS, vol. 4522, pp. 858–867. Springer, Heidelberg (2007)
8. Lienhart, R.: Maydt: An extended set of Haar-like features for rapid object detection. In: *ICIP 2002*, pp. 900–903 (2002)
9. Zhang, L., Wei, H., Zhu, J.: An Ontology-Based Multi-class Terrain Surface Classification System For Aerial Imagery. In: *IEEE International Conference*, pp. 95–95 (January 2012)
10. Abou-Nasr, M.: Terrain Identification in Grayscale Images with Recurrent Neural Networks. In: *International Joint Conference IEEE, Neural Networks (IJCNN)*, pp. 1–5 (July 2010)
11. Clausi, D.A.: Comparison and fusion of co-occurrence, Gabor, and MRF texture features for classification of SAR sea=ice imagery. *Atmos.-Ocean*. 39(3), 183–194 (2001)
12. Rifkin, R., Klautau, A.: In Defense of One-Vs-All Classification. *Journal of Machine Learning Research* 5, 101–141 (2004)

Data Recording System Based on FPGA and DSP

Jun Wang, Xiaoliang Wang, Zhipeng Zhao, Hong Xiang, and Yuxi Zhang

School of Electronic and Information Engineering, Beihang University,
Beijing, China

{wangDSP203, wangxl0720, zhaozhipengBUAA,
xianghong.happy, yxzhang.mr}@gmail.com

Abstract. Data recording system is of great significance in data analysis and scientific researches. According to the development and application of technology, this paper design a data recording system based on FPGA and DSP. DSP is responsible for process data and controlling GPS board. FPGA stores data, communicates with peripheral device with standard HDLC protocol and monitors the system by controlling AD. The data exchange between the system and PC is through Ethernet.

Keywords: Data Recording, FPGA, DSP.

1 Introduction

As the development of modern communication technologies, people increasingly need to store and analyze the real-time data. In many scientific researches, storing and analyzing data is an important and necessary method. Data recording devices, which are also called black-boxes[1,2], are widely applied in daily life. However, data recording devices have some disadvantages such as high cost and demanding of working condition. Based on these reasons, this paper introduces a custom designing data recording system.

The paper introduces the design and implementation of a large capacity data recording system based on FPGA and DSP[3]. As the main controller, FPGA and DSP control the data recording process. The system contains a 8G FLASH and sets aside interface to facilitate future system expansion. Data communication and command transmission between the system and external devices is based on bus and the system adopts a chip with its rate up to 20Mbps. Meanwhile, the system can collect the amount of raw data from external devices. FPGA module monitors the state of system by controlling AD module. The communication between the system and PC is based on Ethernet[4]. PC can set parameters, read data and analyze data. The system implements the function of black-box and has greater compatibility and lower cost.

2 Hardware Design

The system consists of FPGA, DSP, FLASH, bus interface module, Ethernet module, analog data sampling interface module, GPS module and power modules. Power modules convert 5V to the various voltages which are required by chips. The system functional diagram is shown in Figure 1.

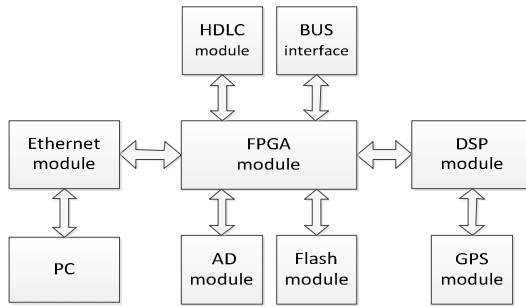


Fig. 1. System architecture

In this system, FPGA controls Ethernet module to communicate with PC, controls AD and flash, communicates with DSP, and implements HDLC protocol when transmits and receives data from bus. DSP implements self-test function in data recording mode, processing data that received from FPGA and communicates with FPGA.

3 FPGA' Function and Design

The functions and modules in FPGA are shown in Figure 2.

Xilinx's XC5VLX50T[5] is chosen in this system. XC5VLX50T has 7200 slices, 480Kb RAM 15 IO Banks and 480 IOs. The clock rate is up to 550MHz and the data rate of IO surmises 1.25Gbps. XC5VLX50T also has 6 DCMs, which provide zero propagation delay, and coarse and fine-grained phase shifting. These features ensure XC5VLX50T to meet the requirements of design and support system's update without wasting resource.

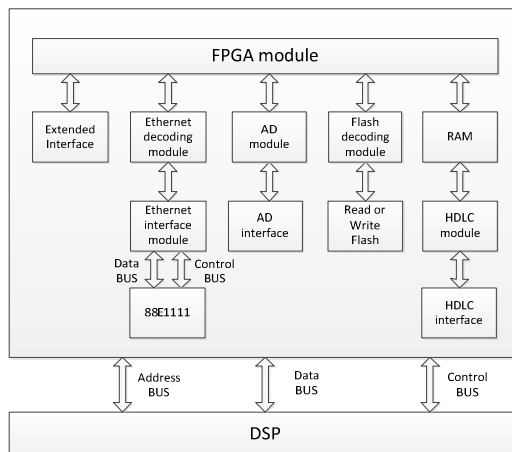


Fig. 2. FPGA architecture

3.1 Flash Module

The function of flash is implemented by PFGA controlling the flash module. In this module, the main function includes function to read, write and erase the flash. The main function performs these functions by calling the corresponding function. The FPGA read, write or erase the flash by Configuring the Register. What else, finding the bad blocks in flash is a very important part. The FPGA realize the function by traversing the bad block logo to find and record the bad block address.

- Normal mode. FPGA communicate with the data machine. When FPGA receive the signal, FPGA notice the DSP by interrupt, and the DSP process and package the data.
- Data transfer mode. DSP receive commands sent by PC through the Ethernet interface to perform different operations. The commands type a total of eight: self-test, reading data, writing data, erasing command, finding bad blocks, marking bad blocks, switching memory and accessing to data distribution.

We choose Samsung's K9F8G08U0M[6] as the system's storage. The memory cell array is 1G×8bit. The system contains two pieces of K9F8G08U0A. One is primary memory and the other is secondary memory. When powered on, the system will detect FLASH. If the primary FLASH is broken or can't be read, the system will use the secondary FLASH in order to prevent data loss.

3.2 Ethernet Module

Users can transmit instructions to the system through the software on PC, which includes loading parameters and inquiring, reading, and erasing recorded data. The design of FPGA contains the decode module, which is used to decode the instructions of PC, and data transfer control module.

MARVELL'S chip 88E1111[7] is chosen as the Ethernet chip which interconnects hardware system and PC. 88E1111 Gigabit Ethernet Transceiver is a physical layer device for Ethernet 1000BASE-T, 100BASE-TX, and 10BASE-T applications. According to the system's requirement, we choose the 1000BASE-X interface, GMII mode. In this mode, data transfer rates can up to 2.5Gbps. The Ethernet chip is controlled by FPGA.

The implement of controlling 88E1111 by FPGA includes PHY interface, EMAC, transmit FIFO, Receive FIFO and user logic interface. The PHY interface connect the FPGA and 88E1111, EMAC package can be implemented by Virtex-5 FPGA embedded tri-mode Ethernet IP core, transmit FIFO and Receive FIFO can be implemented by hardware language VHDL.

Figure 3 shows the functional diagram of the interconnection between FPGA and 88E1111.

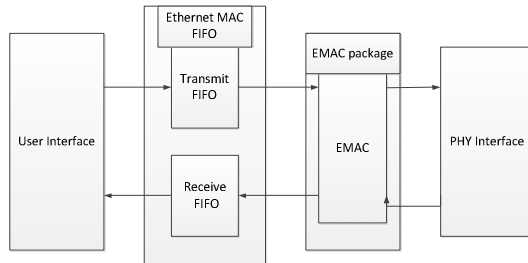


Fig. 3. Ethernet module architecture

3.3 AD Module

In order to detect whether the system is working properly, we choose MAX1270 of MAXIM[8] to monitor the system's voltage. The MAX1270 are multi-range, 12-bit data-acquisition systems that require only a single +5V supply for operation, yet accept signals at their analog input that can span above the power-supply rail and below ground. MAX1270 provide eight analog input channels that are independently software programmable for a variety of ranges: -10 to 10, 0 to 10, -5 to 5 and 0 to 5. We choose the range of -10 to 10 and eight channels. That can totally meet the system's requirement.

3.4 HDLC Protocol

HDLC (High-level Data Link Control) protocol is applied in the communication between the system and peripheral device[9]. HDLC protocol is widely used in digital communication area. As a transfer unit in HDLC, a frame consists of start flag, address, control, data, frame check sequence and final flag. FPGA is used to implement the protocol in this system.

When FPGA transfers data to peripheral device, FPGA will first insert '0' into the data, then process CRC(Cyclic Redundancy Check) and add start flag and final flag to the frame in order to packets data, and finally convert the parallel data to serial data and transmits it. When FPGA receives data from peripheral device, FPGA will process the data by the above steps inversely and stores the unpacked data into buffer RAM. Finally, the unpacked data will be stored in FLASH in the way described above.

4 DSP's Function and Design

In this system, DSP is primarily responsible for writing the main function, controlling the flow of the system. Considering the feature of high computing power and easy development of DSP.

- System self-test: After power on, DSP will initialize the self-test to check whether the system works normally. Only when the system passes the self-test, other functions can be operated properly. To present the state directly, the system has some

indicating LEDs connected with FPGA to detect some functions. If the system fails in self-test, the self-test LED will keep off.

- Controlling GPS board: DSP is responsible for controlling the GPS board. This function is wrapped as a module in DSP. The GPS board is a daughter board that embedded in the main board. The GPS board can provides target's real-time location and time information. The DSP collects these data and sends to FPGA, the latter records them to flash.
- Communicating with FPGA: DSP needs to exchange command and data with FPGA. A communication module is designed in DSP to communicate with FPGA by decoding the command from FPGA and transferring command and data to FPGA.
- Processing Data: As DSP has powerful fixed-point and floating-point capacity and is easy for coding, in this practice, the comparison of data deposited and analysis of receiving data are completed in DSP. This makes the codes more intelligible, facilitates the revise and update of system, and makes the applications broader.

DSP is TI's TMS320F2812[10]. It has following features: high-performance static CMOS technology, high- performance32-bitCPU with Harvard bus architecture and fast interrupt response and processing, up to 128K×16 Flash and 8K×16SARAM. The architecture of DSP is shown in Figure 4.

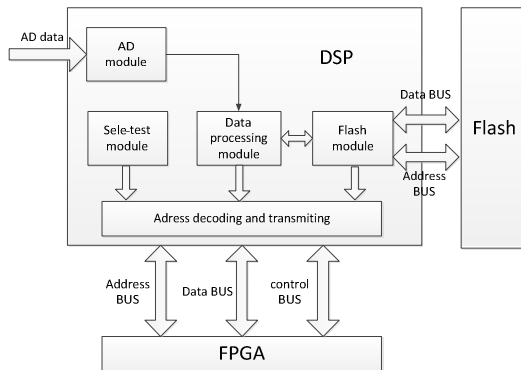


Fig. 4. DSP architecture

5 Implementation of the System

The system has two work modes, which is determined by connection with PC or not. If the system detects that there is no connection with PC, it will start data transfer and recording mode. In this mode, the system reads the loaded parameters, communicates with peripheral device in HDLC protocol, and records the data in FLASH through FPGA and DSP. If the system connects with PC, PC can load parameters to the system, inquire and read data from it, and save the data into EXCEL, which is convenient for analysis. Meanwhile, PC can also simulate the data transfer and recording processes. The flow chart is shown in Figure 5.

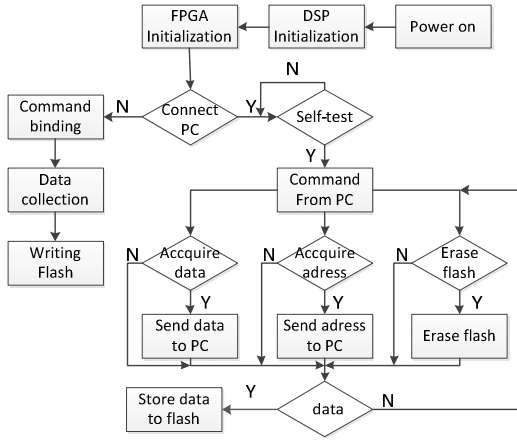


Fig. 5. System flow chart

The situation of testing data transmission of Ethernet module controlled by FPGA is shown in Figure 6 and Figure 7. when the system is connected with computer. The command designed to send can be seen from the signal GMII_TXD. As shown in the figure, the system send 00,11 such orders to computer. These data received can be displayed on the computer through the software wireshark, which is shown in Figure 7. We can see that the computer successfully received the data from the system through Ethernet.

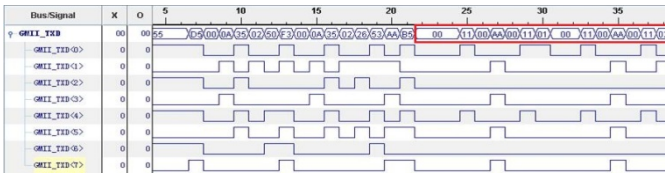


Fig. 6. Ethernet hardware test plans

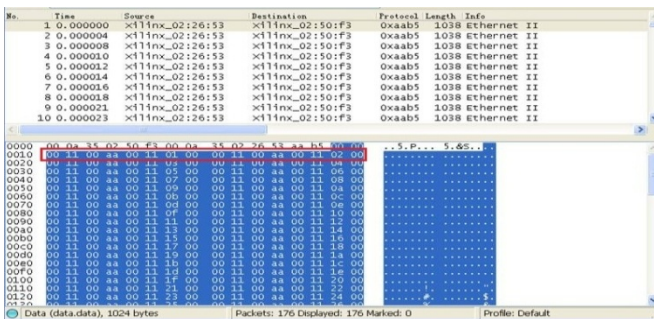


Fig. 7. Computer connect with the system

Figure 8 shows HDLC communication protocol achieved in the FPGA. TX_Reg is the data prepared to send and SPI_Dat_Out_T is the serial data of a changed from TX_Reg on the data line. After sending "7E", flag header the of HDLC protocol, the system will send data "47,52,00,09" in the register. At last, according to the protocol, send "7E" as the end of the frame. The data on the serial data line should be read inverted in the figure because data sent is designed for each byte of the starting low and then high.

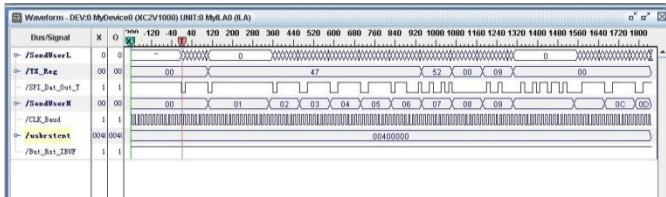


Fig. 8. HDLC protocol hardware test plans

6 Summary

The paper design and implement a data recording system based on FPGA and DSP. HDLC protocol is implemented in FPGA instead of dedicated chip. USB2.0 protocol is used to interconnect the system and PC. FLASH can be controlled flexibly. The system is practical, as the design is modularized so that the program can be debugged and modified easily, and many extend interfaces can facilitate the extension and update.

References

1. Jung, S.M., Lim, M.S.: System on Chip Design of Embedded Controller for Car Black Box. In: 2007 IEEE Intelligent Vehicles Symposium, pp. 1174–1177. IEEE (2007)
2. Zhang, Z., Zhan, J., Li, Y., Wang, L., Meng, D., Sang, B.: Precise request tracing and performance debugging for multi-tier services of black boxes. In: IEEE/IFIP International Conference on Dependable Systems & Networks, DSN 2009, pp. 337–346. IEEE (2009)
3. Baofeng, Z., Ya, W., Junchao, Z.: Design of high speed data acquisition system based on FPGA and DSP. In: 2010 International Conference on Artificial Intelligence and Education (ICAIE), pp. 132–135. IEEE (2010)
4. Lofgren, A., Lodesten, L., Sjöholm, S., et al.: An analysis of FPGA-based UDP/IP stack parallelism for embedded Ethernet connectivity. In: 23rd NORCHIP Conference, pp. 94–97. IEEE (2005)
5. Xilinx Inc. Virtex-5 FPGA User Guide (March 16, 2012)
6. Samsung Semiconductor. K9F2G08U0A Datasheet (March 17, 2006)
7. En, S., Dongming, Y., Jinchun, G., et al.: Design and Optimization of Ethernet Access System for Receiving Frames of Ethernet Based on FPGA. In: 2012 International Conference on Computer Science and Electronics Engineering (ICCSEE), vol. 3, pp. 129–133. IEEE (2012)
8. MAXIM. MAX1270 Datasheet (March 2004)
9. Wang, J., Zhang, W., Zhang, Y., Wu, W.: Weiguang Chang, Design and implementation of HDLC procedures based on FPGA. In: 3rd International Conference on Anti-counterfeiting, Security, and Identification in Communication, ASID 2009, pp. 336–339. IEEE (2009)
10. Instruments T. TMS320F2812 Digital signal processor data manual. SPRS174M

Implementation of High-Speed Serial Interconnects for Multi-Processor Parallel System

Jun Wang, Zhipeng Zhao, Bin Yang, Wengui Fan, and Yuxi Zhang

School of Electronic and Information Engineering, Beihang University
Beijing, China

{wangDSP203, zhaozhipengbuaa, young.being1990,
engineerdotfan, yxzhang.mr}@gmail.com

Abstract. Multi-processor parallel system is widely used to meet the increasing demand of high-performance computing. However, the data interacting between processors is often a bottleneck as it is hard for parallel buses to address the requirement of transmitting mass data in high speed. Based on differential signal and embedded clock technology, high-speed serial interconnect has been the ideal solution. In this paper, we designed a multi-processor parallel system and implemented the high-speed serial interconnects between processors based on Serial Rapid IO (SRIO), PCI Express (PCIe), and Hyperlink protocols. The test results are given to show the throughput performance of these interconnects.

Keywords: Multi-Processor Parallel System, SRIO, PCIe, Hyperlink.

1 Introduction

To meet the increasing requirements of real-time and large data processing in radar, image, and video signal processing, multi-processor parallel system becomes necessary [1,2,3]. The data processing and transmission are two main measurement of performance of multi-processor parallel system. With the rapid development of performance of processors like Digital Signal Processor (DSP) and Field Programmable Gate Array (FPGA), the demand of data processing capability can be satisfied in common. Thus, the capacity of data interacting between processors becomes the key factor of performance of multi-processor parallel system [4].

The traditional interconnect between processors is parallel bus. In order to increase the data rate, parallel bus needs higher clock and wider bus. However, the higher clock rate reduces the skew timing budget and so does a wider bus width [5]. Moreover, a wider bus makes it difficult to layout and route in Printed Circuit Board (PCB). Parallel bus has reached a limit. The ever-increasing throughput requirements have forced conversion of IO buses from multi-load parallel single-end buses to point-to-point serial differential buses [6].

2 High-Speed Serial Interconnects

High-speed serial interconnect takes advantage of differential signal technology and embedded clock technology to avoid clock and data skew, which makes it possible

that the signaling rate can easily get a few Gbps of bandwidth. Furthermore, high-speed serial interconnect can reduce the pin number leading to PCB wiring easier and saving overall cost [4]. The common interconnects are PCI Express (PCIe), Serial Rapid IO (SRIO), Ethernet (GbE), InfiniBand (IB), Fiber Channel (FC). Besides these standard interconnects, there are also some specific interconnects defined by corporations, such as Texas Instrument's Hyperlink. The applications of these interconnects are different. InfiniBand and Fiber Channel are targeted as box-to-box interconnects in System Area Network (SAN). Ethernet is regarded as box-to-box interconnects in both SAN and Local Area Network (LAN). PCIe, SRIO and Hyperlink support chip-to-chip and board-to-board interconnects. In multi-processor parallel system, we are more interested in chip-to-chip interconnects. Thus, in this paper, we focus on SRIO, PCIe and Hyperlink.

SRIO is well suited for control and data operations in communication and embedded systems requiring high-speed I/O with low latency. The signaling rate can be 1.25, 2.5, 3.125, 5 or 6.25 Gbps. The specification defines x1, x2, x4 and x8 lane versions. In our system, SRIO is applied in the interconnect of FPGA-FPGA and FPGA-DSP.

PCIe is a point-to-point and packetized protocol. The supported signaling rates include 2.5, 5 or 8Gbps. This specification describes operations for x1, x2, x4, x8, x12, x16, and x32 lane widths. PCIe is often regarded as one of the main interfaces to connect board to PC [7,8]. Because key PCI attributes, such as its usage model, load-store architecture, and software interfaces, are maintained, whereas its parallel bus implementation is replaced by a highly scalable, fully serial interface. We take advantage of the feature that PCIe supports chip-to-chip interconnect to maximize the bandwidth between FPGA and DSP.

HyperLink provides a high-speed, low-latency, and low-pin-count communication interface between two KeyStone devices. It can run in 1 lane or 4 lanes mode with each lane running at 1.25, 3.125, 6.25 or 12.5Gbps. As Hyperlink is well suitable for the interconnects between two Keystone devices, we choose Hyperlink to connect two TI's Keystone DSPs.

3 Hardware Design

As shown in Fig.1, the multi-processor parallel system consists of two boards. One is interface board and the other is core board. Composed by two FPGAs and one DSP, interface board is responsible for controlling AD and DA. Core board has two FPGAs and two DSPs and accounts for data processing. Every processor in core board has connection with DDR3, which work as external memory. The two boards are interconnected by high-speed cable. In this paper, we focus on the high-speed serial interconnections between processors of core board.

In core board, we use Xilinx's FPGA XC6VVSX315T and TI's DSP TMS320C6678. Xilinx's Virtex-6 series FPGA have GTX transceivers which support various high-speed serial protocols, such as PCIe and SRIO. TMS320C6678 is TI's latest multi-core processor integrated a plethora of high-speed serial interfaces including PCIe, SRIO and Hyperlink. The interconnections of FPGA-FPGA are 24 pairs of

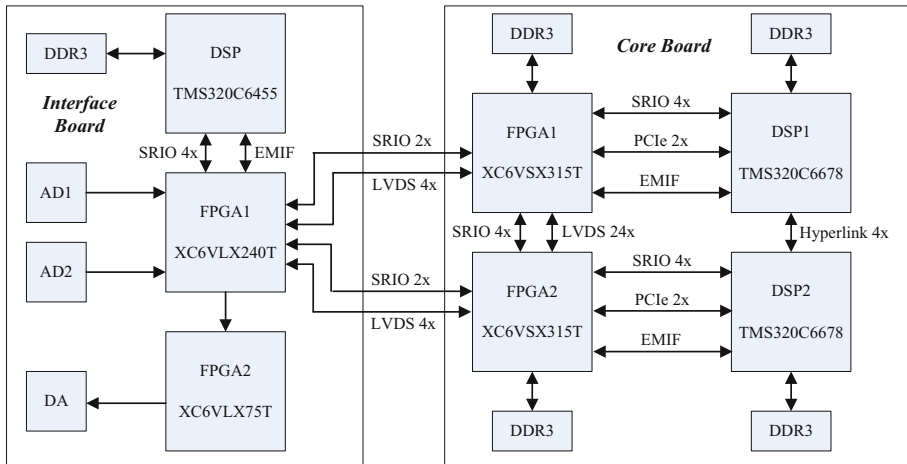


Fig. 1. Architecture of Multi-processor Parallel System

LVDS and 4x width SRIO. FPGA-DSP interconnects are EMIF, 4x SRIO, and 2x PCIe. EMIF is parallel interface and responsible for a small amount of data exchange. PCIe and SRIO are used to transform mass data between FPGA and DSP. Two DSPs are interconnected by 4x Hyperlink. The data rate of SRIO and PCIe is 5Gbps per lane and Hyperlink is 6.25Gbps per lane. Thus, the high-speed serial interconnects bandwidth of FPGA-FPGA, FPGA-DSP, and DSP-DSP is 20Gbps, 30Gbps, and 25Gbps, respectively.

The data rate of high-speed serial interfaces can easily be a few Gbps, which is demanding for hardware design. Thus, signal integrity, power integrity and the quality of the input clock of high-speed serial interface modules must be taken into account when dealing with hardware design of high-speed serial interfaces. Otherwise, the misalignment of lanes or instability of link will occur. When wiring, the length of differential pairs should be equal and the receive pairs and transmit pairs should be wiring in the same layer. AC coupling capacity, used to filter low frequency interference, should be placed near the receiver. It is also suggested that the power supply of high-speed serial interface modules chooses independent power and the power noise should be noticed. The frequency and amplitude of the input clock must meet the requirement of high-speed serial interface modules in processors.

4 Software Design

In software design, we discuss the implementation of SRIO, PCIe and Hyperlink, including the modules design in FPGA, the steps of registers configuration in DSP and the process of data exchange.

4.1 SRIO

SRIO is used as interconnect of FPGA-FPGA and FPGA-DSP. First, we discuss the implementation of SRIO in FPGA. We use Xilinx's Serial RapidIO Endpoint solution

and design Initiator User Design and Target User Design as shown in Fig. 2. Initiator User Design accounts for transmitting the request packets of Local Endpoint and receiving the response packets of Remote Endpoint. Target User Design is responsible for receiving the request packets of Remote Endpoint. The function of Serial RapidIO Endpoint solution is to control the concatenation and parsing of transmit and receive packets, provide buffer, and handle link training, initialization and protocol .

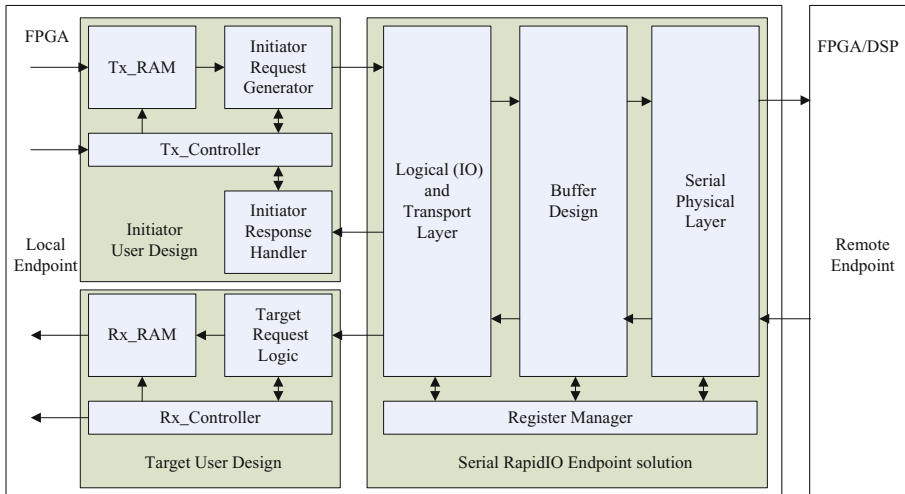


Fig. 2. SRIO Architecture

When Local Endpoint transfers data to Remote Endpoint, data is firstly written to Tx_RAM. After that, Tx_Controller will receive a high pulse to start a SRIO transfer. According to the packet information which is set by user, including packet type, size, number, address and ID, Tx_Controller controls Initiator Request Generator to read data from Tx_RAM and generate packets. After the processing of Serial RapidIO Endpoint solution, these packets are converted to high-speed serial bitstream transferred to Remote Endpoint by differential pairs. When Remote Endpoint transfers data to Local Endpoint, Serial RapidIO Endpoint solution decodes the high-speed serial bitstream to form SRIO packets transferred to Target Request Logic. Under the control of Rx_Controller, Target Request Logic writes the data in packets to Rx_RAM. After that, Rx_Controller send a high pulse to other modules to signal Rx_RAM is ready for reading.

In DSP, the implementation of SRIO can be divided into 4 steps: map address; configure ID, SRIO ports, and interrupt vectors; configure LSU registers; wait for link up. When link is up, DSP is able to transmit or receive SRIO packets. DSP and FPGA should know the destination ID and start address before they initiate a transfer.

4.2 PCIe

Besides SRIO, PCIe is also applied as interconnect of FPGA-DSP. Different from SRIO's peer-peer communication, PCIe's communication is master-slave. DSP is configured as Root Complex (RC); FPGA is configured as Endpoint (EP). The design in FPGA is based on Virtex-6 FPGA Integrated Block for PCI Express core. However, the reference user design provided by the IP core uses PIO transfer which is not suitable for mass data transfer. We designed a Bus Master DMA (BMD) controller based on Xapp1052. The implementation in DSP has three steps: initiation, link training, and configuring remote registers. In initiation, we configure Serdes clock, enable PCIe power and clock, configure PCIe mode, and map address. After initiation, DSP checks Link Training and Status state machine (LTSSM). Only when LTSSM's state is L0, the link is up. After link up, DSP configures FPGA by transmitting Configuration Write/Read packets. The Bus Master Enable bit in command register of FPGA must be enabled.

DSP can control the transfer of FPGA by transmitting several Memory Write/Read packets to configure control and status registers of BMD design including: reset DMA controller, configure address, configure packet size, configure packet number, and start DMA transfer. Then, under the control of BMD controller, FPGA transmits mass Memory Write packets to DSP. When transfers mass data to FPGA, DSP enables EDMA controller to control the transfer of mass Memory Write packets to FPGA.

4.3 Hyperlink

TMS320C6678 is integrated Hyperlink module that accounts for Hyperlink interconnections between two Keystone devices. The configuration of Hyperlink module includes configure system PLL, enable Hyperlink power and clock, choose lane number and clock mode, configure Serdes registers and initiate Hyperlink module. When two devices finish the configuration, lanes will be detected until the link is up. After link up, one device is able to read or write the other one's Hyperlink registers and address should be mapped before transmitting data. Then according to the address map, DSP enables EDMA controller to transfer data to other device.

5 Results

The performance of throughput of SRIO, PCI Express, and Hyperlink is shown in Table 1. It is measured by counting the number of operating clocks until a 16KB transfer is complete. When test SRIO, we choose nwrite packet whose size is 256B. Memory Write Request packet is used in PCIe test and packet size is 128B as the maximum outbound payload size of DSP is 128Bytes. Hyperlink is tested between DSP1's L2 memory to DSP2's L2 memory. Theoretical throughput is decided by lane width, lane speed and encoding efficiency. SRIO and PCIe use 8b/10b encoding while Hyperlink's encoding scheme is equivalent to 8b/9b as HyperLink reduces the encoding overhead. The efficiency is calculated by dividing the practical throughput by theoretical throughput.

Table 1. Performance of High-speed Serial Interconnects

Interconnect Type	SRIO	PCI Express	Hyperlink
Practical Throughput	1793.50MBps	665.48MBps	1960.97MBps
Theoretical Throughput	2048MBps	1024MBps	2844.44MBps
Efficiency	87.57%	64.99%	68.94%

6 Conclusion

Surmounting the bottleneck of bandwidth of parallel buses, high-speed serial interfaces become the trend of interconnects between processors. In this paper, SRIO, PCIe, and Hyperlink, three typical high-speed serial interconnects used in multi-processor parallel system, are introduced. We implement these interconnects based on hardware and software design and gives the performance of throughput.

References

1. Zhan, Z.H., Hao, W., Tian, Y., Yao, D.W., Wang, X.-H.: A Design of Versatile Image Processing Platform Based on the Dual Multi-core DSP and FPGA. In: 5th International Symposium on Computational Intelligence and Design, vol. 2, pp. 236–239 (2012)
2. Bian, M.-M., Gao, L.-N., Xie, Y.-Z., Tan, X.-B.: High-performance System Design of SAR Real-time Signal Processing. In: 2010 International Conference on Computer Application and System Modeling, vol. 12, pp. V12126–V12129 (2010)
3. Fiorucci, F., Verducci, L., Micanti, P., Baruffa, G., Frescura, F.: Implementation of a Re-programmable DSP/FPGA Based Platform for Real-time HD Video Coding. In: European DSP in Education and Research Conference, pp. 185–189 (2010)
4. Huang, X.-Y., Su, H.-B., Wu, Q.-Z., Wu, W.: Multi-Processor Parallel System Based on High-Speed Serial Transceiver. In: 2nd International Workshop on Education Technology and Computer Science, vol. 1, pp. 178–181 (2010)
5. Dhawan, S.K.: Introduction to PCI express - A new high speed serial data bus. In: IEEE Nuclear Science Symposium Conference Record, vol. 2, pp. 687–691 (2005)
6. Lee, B.-T., Mazumder, M., Mellitz, R.: High speed differential I/O overview and design challenges on Intel enterprise server platforms. In: 2011 IEEE International Symposium on Electromagnetic Compatibility (EMC), pp. 779–784 (2011)
7. Kavianipour, H., Bohm, C.: High performance FPGA-based scatter/gather DMA interface for PCIe. In: IEEE Nuclear Science Symposium Conference Record, pp. 1517–1520 (2012)
8. Bittner, R.: Speedy bus mastering PCI express. In: 22nd International Conference on Field Programmable Logic and Applications, pp. 523–526 (2012)

A Twin-Screw Rotor Profile Design Method Based on Computational Fluid Dynamics

He Xueming, Pan Chenglong, Wu Meiping*, and Ji Xiaogang

Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment and Technology(Jiangnan University), Wuxi 214122, China
{hxuem2003,pc10401080707}@163.com,
wmp169@hotmail.com, bhearts@126.com

Abstract. Increasing demands for more efficient screw compressors design method to overcome the problems of rotor profile modification and compressor performance test establishment. A new method for the design of screw compressor shape, size and dimension is described here. As a result, NURBS curves and computational fluid dynamics (CFD) theory being applied to overcome these two problems. NURBS curves are introduced to construct a new rotor profile based on the rotor profile of unilateral asymmetric combined cycloid with pin gear arc. The numerical simulation models of the new profile and the original profile are built to find out the distribution law of pressure and velocity in twin-screw compressors under the operating speed by the means of FLUENT software. Finally, according to the simulation results, the rationality of the designed rotor profile is verified, and new schemes for further optimization are explored.

Keywords: Computational fluid dynamics, Freeform curves, Design, Rotor profile.

1 Introduction

Screw compressors are compact, efficient and reliable. Consequently, they are widely used in industrial application as well as refrigeration systems [1].

Rotor profile is the key factor of twin-screw compressor performance. There are a lot of rotor profile design methods. N. Stosic, et al. [2] proposed an optimization method for screw compressor. He Xueming, et al. [3] presented a positive and reverse design method of screw rotor profiles with freeform curves. Moreover, Yu-Ren Wu, et al. [4] proposed a method for screw compressor rotor profile design based on any arbitrary sealing line. With the development of computational fluid dynamics theory, numerical simulation is becoming increasingly widespread in the design of the screw compressor profile. Wen Jing, et al. [5] used numerical simulation technology in the design of a twin-screw kneader. Shi Wen, et al. [6] proposed an analysis model of twin-screw compressor rotor to provide a theoretical basis for profile design and optimization.

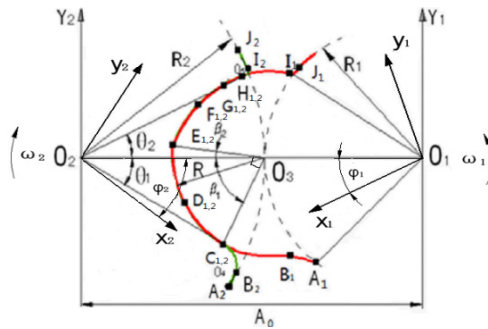
NURBS curves are the most general B-spline curves, which could be adjusted through changes of control points or modification of parametric equation weights [7].

* Corresponding author.

In this paper, rotor profile of unilateral asymmetric combined cycloid with pin gear arc is put as the research object to acquire a new rotor profile. The rotor profile is established by NURBS curves under the screw rotor meshing principle, and the parametric equations of rotor profiles are given. The numerical simulation method is applied to twin-screw rotor profile to improve the design efficiency. Computational fluid dynamics theory [8] is utilized in order to describe the pressure, velocity distribution of twin-screw compressor under the working speed 3000r/min.

2 Twin-Screw Compressor Design

2.1 General Coordinates System



- A_0 —distance between the rotor axes; O_1, O_2 —rotation axis;
- R_1, R_2 —radii of rotor pitches; $Y_2O_2O_3, Y_1O_1O_3$ —static coordinate systems;
- O_3 —intersection point; $O_2x_2y_2, O_1x_1y_1$ —moving coordinate systems;
- O_4, O_5 —centers of arc B_2C_2 and H_2I_2 ; R —radii of arc $C_1D_1E_1, C_2D_2E_2$;
- β_1, β_2 —radians of point $C_{1,2}$ and point $E_{1,2}$; ϕ_1, ϕ_2 —rotation angles;
- ω_2, ω_1 —rotation velocities; θ_1, θ_2 —radians of point C and point H

The female and male rotor profiles are divided into 9 segments respectively, as is shown in Fig. 1. Female or male rotor profile of twin-screw compressor is predefined on each segment; while the other is deduced by the conversion relationship of the coordinate systems and meshing principles of twin-screw compressor in the process of rotor profile design. And the design or optimization operations could be made through redefining the curves in the corresponding segments of the predefined rotor profiles.

2.2 NURBS Curve

The Non-Uniform Rational B-Spline (NURBS) curves are known as the most common freeform curves. Their expressions could be written as follows:

$$P(t) = \sum_{i=0}^n Q_i L_{i,n}(t) \tag{1}$$

$$L_{i,n}(t) = \frac{V_i B_{i,n}(t)}{\sum_{j=0}^n V_j B_{j,n}(t)}$$

$$i=1,2,\dots, nm-1 \tag{2}$$

Where Q_i represent the control points, V_i are weight factors of the parametric equation, m is the number of power, $B_{i,m}(t)$ represent rational basis functions. NURBS curves could be changed through adjusting either control points or weight factors. In this paper, NURBS curves are used to establish rotor profile to shorten profile design time as well as improving the design efficiency.

3 Design and Optimization

As is shown in Fig. 1, the female and male rotor profiles are divided into nine segments, respectively. According to the practical requirements, NURBS curves and other freeform curves are applied to the design of the female rotor profile and optimization. However, the male rotor profile is deduced from the equations of the female rotor profile. Since the freeform curves such as NURBS curves are applied to rotor profile, the twin-screw compressor rotor profile could be modified through either adjusting weight factors or control points easily.

Table 1. Comparison of the rotor profiles

Segments	Female rotor		Segments	Male rotor	
	Origin	New		Origin	New
A ₂ B ₂	Arc	Arc	A ₁ B ₁		Arc
B ₂ C ₂	Line	Arc	B ₁ C ₁	Cycloid	
C ₂ D ₂		Arc	C ₁ D ₁		Arc
D ₂ E ₂	Arc	Arc	D ₁ E ₁	Arc	envelope
E ₂ F ₂		NU	E ₁ F ₁		NURBS
F ₂ G ₂	Cycloid	RBS	F ₁ G ₁	Cycloid	curve
G ₂ H ₂	Line	curve	G ₁ H ₁		envelope
H ₂ I ₂	None	Arc	H ₁ I ₁	Line	Cycloid

The rotor profile of unilateral asymmetric combined cycloid with pin gear arc is shown in Fig.2(a). There’s a long contact line and sealed volume on the profile, which results in loudly operating noise, air leakage, and low working efficiency in practical applications. As is described above, freeform curves such as NURBS curves are introduced in this paper, and the newly designed rotor profile is shown in Fig.2 (b). All the new designs are listed in Table 1.

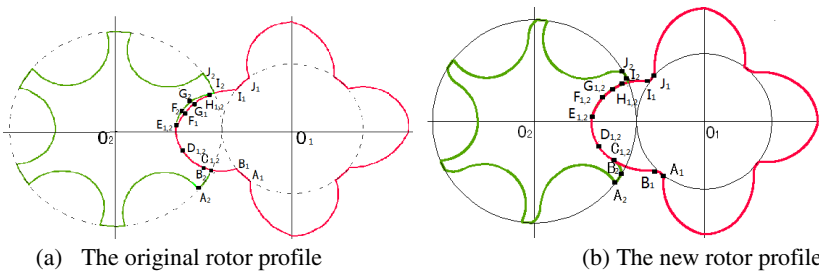


Fig. 2. Rotor profiles

4 Numerical Simulation

4.1 Numerical Simulation Models

Taking into account of the specific twin-screw compressor air conditions and the airflow characteristics, the following assumptions are made: 1) The air flow is stable, isothermal, isotropic. 2) Since the Reynolds number of the air flow is large, the air flow is considered as a kind of turbulent flows. 3) Affects of inertia are ignored as well as gravity. 4) There is no effect of lubricating oil on the nature of the air flow. 5) There is no air leaking outside of the compressor in the compression process. 6) The twin-screw compressor working cycle is theoretical. Consequently, the air flow numerical model is defined as follows:

$$\mu_t = \rho C_\mu \frac{r^2}{\varepsilon} \quad (3)$$

μ_t ——turbulent viscosity; ρ ——gas density;

C_μ ——empirical constant; r , ε ——turbulent kinetic energy; turbulent dissipate rate.

If the profiles of female and male rotor spiral up about their rotation axes on the radial cross section, the corresponding screw rotor three-dimensional model could be formed. Through a Boolean operation, the twin-screw compressor flow models are established. Finally, they are classified into unstructured tetrahedral meshes by dynamic mesh methods, which is shown in Fig.3 (a) (b).

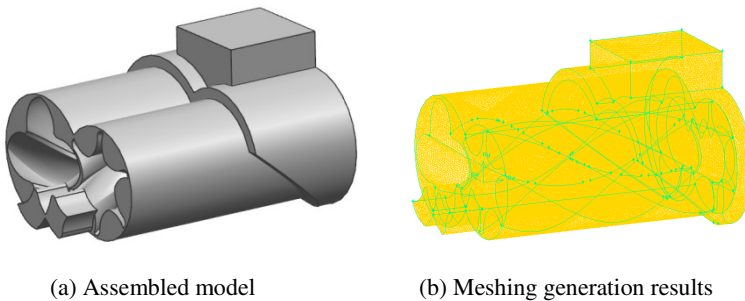


Fig. 3. Three-dimensional models

4.2 Simulation Results and Further Optimization

As is shown in Fig.4, the lowest pressure of the intake port is -0.0922MPa on the original air flow field, and the pressure of the new flow field is -0.0907MPa; the highest pressure of exhaust port is 0.707MPa on original air flow, the new one's is 0.767MPa. Based on the above results, the pressure distribution of the new designed air flow field is improved: the intake pressure increased by 1.6%, the highest exhaust pressure increased by 8.5%, and the maximum pressure difference increased by 7.3 %.

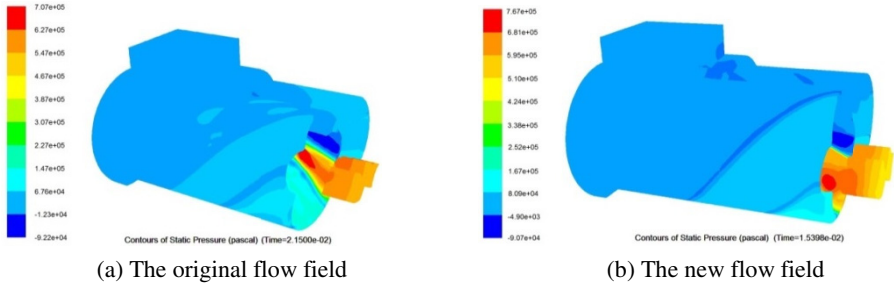


Fig. 4. Dynamic simulation results of pressure

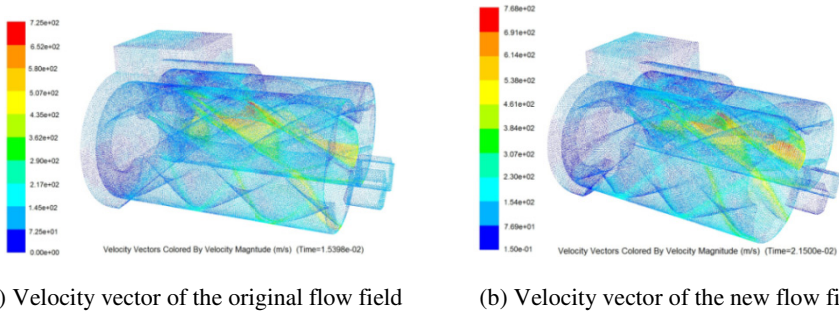


Fig. 5. Dynamic simulation results of velocity

In Fig.5, it is easy to see that the maximum leaking velocity of the original flow is 725m/s. The maximum leaking velocity of new flow field is 768m/s, which is increased by 5.6% compared to the original flow field. The velocity of intake flow field is calculated as the displacement of the compressor. The displacement of the original compressor flow field is 15.5m³/min, and the displacement of the new compressor is 16.1m³/min, which is increased by 3.9%.

According to the dynamic simulation analysis results, the maximum differential pressure of newly designed flow is increased significantly, which indicates a certain rationality of the new rotor profile compared to the original profile. However, the leakage triangle is expanded. Moreover, the further growth in displacement of compressor is affected.

The further improvements of the new designed twin-screw compressor rotor profile are still needed. Freeform curves such as NURBS curves could be used instead of the cycloid and the arcs in segment H₁I₁. The radius of arc H₂I₂ is expanded and points H_{1,2} are moved some distance towards the upper right of their original position. These optimizations could be obtained through adjusting control points or weight factors in parametric equations of curves. Then numerical models would be built to make dynamic simulation again. The rotor profile would be modified until a good performance rotor profile to be obtained based on the simulation results. Finally, a prototype of the twin-screw compressor rotor profile would be produced and a confirmatory experiment is to be made to verify performance of the rotor profile.

5 Conclusion

In this paper, the freeform curves (NURBS curves) are used to establish rotor profiles to achieve a smooth mesh of the rotors, and the modification of rotor profile could be made through adjusting the control points or weight factors of the curves. The CFD theories and freeform curves are introduced to twin-screw rotor profile design. Based on the numerical simulation results, freeform curves are applied to the design or optimization of the rotor profile until a newly good performance profile to be acquired. Through the new design method, the number of physical experiments is reduced as few as possible to achieve the superiorities of time, cost, reliability and agility.

Acknowledgement. This project is funded by Natural Science Foundation of China (51275210), (51105175); Industry-University-Research Foundation of Jiangsu Province (BY2013015-30); and Six-Major-Talent-Summit Project of Jiangsu Province, China(2013-ZBZZ-016).

References

1. Sun, X.: Analysis of Market Development Trends of Screw Compressor for Process Gas. *Compressor Technology* 1, 65–68 (2012)
2. Stosic, N., Smith, I.K., Kovacevic, A.: Optimization of screw compressors. *Applied Thermal Engineering* 23(10), 1177–1195 (2003)
3. He, X., Dai, J., Liu, H.: Positive and Reverse Design of Screw Rotor Profiles Based on Freeform Curve. *China Mechanical Engineering* 23(22), 2752–2756 (2012)
4. Wu, Y.-R., Fong, Z.-H.: Improved rotor profiling based on the arbitrary sealing line for twin-screw compressors. *Mechanism and Machine Theory* 43(6), 695–711 (2008)
5. Wen, J., Sun, X., Sun, W., et al.: Rotor Profiles Design Method and Numerical Simulation for Twin-screw Kneader. *Chinese Journal of Mechanical Engineering* 49(3), 63–73 (2013)
6. Shi, W., Yu, X., Li, J.: Finite Element Modal Analysis of Double Screw Compressor Rotor. *Compressor Technology* 4, 7–9 (2012)
7. Chen, J., Cui, H.: Shape modification of NURBS curves based on weight. *Journal of Naval University of Engineering* 24(4), 108–112 (2012)
8. Han, X., Qian, R.: A study of the basic theory of fluid mechanics. *Spatial Structure* 14(3), 9–12 (2008)

Design of Nanograting Structures for Optoelectronic Devices Based on Rigorous Coupled-Wave Analysis

Nghia Nguyen-Huu^{1,2,*}, Jaromir Pistora², and Michael Cada^{1,2}

¹ Nanotechnology Center, VSB-Technical University of Ostrava,
17. Listopadu 15, Ostrava-Poruba 708 33, Czech Republic

² Department of Electrical and Computer Engineering,
Dalhousie University, Halifax, Nova Scotia, Canada
nghianano@gmail.com, jaromir.pistora@vsb.cz,
michael.cada@dal.ca

Abstract. This paper describes an efficient numerical technique, namely, the Rigorous Coupled-Wave Analysis or Fourier Model Method used to analyze optical properties of diffraction grating structures. Numerical calculations for grating structures featuring simple and complex profiles used for solar cells and an optical filter and sensor are demonstrated. The first application is an analysis of the optical absorbance of an optimized solar absorber in the visible and near infrared regions with a consideration of its different geometric shapes that could be occurred in micro/nanofabrication processes. The solar absorber based on a one-dimensionally complex silicon structure comprises a single-layered Si grating on top of a silicon substrate. The second one represents a plasmonic nanostructure owning the characteristic of optical transmission filtering and refractive indices sensing. The integrated nanodevice is constructed based on a silicon dioxide sandwiched between two metal grating layers.

Keywords: Filters, Grating, Rigorous Coupled-Wave Analysis, Sensors, Silicon, Solar Energy.

1 Introduction

The Rigorous Coupled-Wave Analysis (RCWA) is an efficient tool for analyzing diffraction efficiency of periodic grating structures by solving Maxwell's equations relevant to their geometric shapes [1-5]. Electromagnetic fields in the periodic structures concluding many regions with different refractive indices can be solved by employing continuity boundary conditions, and as a result, diffraction efficiency orders are easily obtained. The RCWA is widely-used for analyzing dielectric, metallic, and semiconductor grating structures with simple and complex geometric profiles. For one-dimensional (1D) gratings, it is well developed with the plane of incidence perpendicular to the grating grooves since its convergence problem was also solved by introducing correct Fourier factorization rules to abrupt discontinuities in permittivity [2]. Moreover, the RCWA is also

* Corresponding author.

applied to 2D and 3D periodic structures with arbitrary permittivity and permeability tensors [6]. In addition, electromagnetic fields of non-periodic structures can be found by applying perfectly matched layered boundary conditions.

In the current paper, along with the numerical descriptions of the RCWA we also present its application to design periodic grating structures working as solar absorbers for solar cells and an integrated device having optical filtering and refractive indices sensing functions. With the fast development of nanotechnologies, photonic nanostructures have emerged as important devices for energy collectors, optical filters, polarized color filters, and biosensors [7-16]. Besides various excellent theoretical investigations, micro/nanofabrication technologies have also developed significantly with great precision [17-19]. Therefore, in this study grating structures comprising a silicon (Si) grating layer deposited on a Si substrate are investigated to collect the solar energy in the visible and infrared regimes. Moreover, to demonstrate efficient calculations of the RCWA the solar absorbers based on the grating structure are analyzed by adding tiny features on grating sidewalls since the actual fabrication process is not completely perfect (photoresist residues or intrinsic limits of a fabrication process) [20-23]. Finally, a novel multifunctional nanodevice (i.e., wavelength filtering and refractive indices sensing) constructed on a single double-sided grating structure is demonstrated. Its optical transmission is studied by varying their geometric parameters and refractive indices surroundings in order to evaluate different applications.

2 Rigorous Coupled-Wave Analysis

The RCWA is an efficient tool for analyzing the diffraction efficiency of periodic gratings bounded by two media I and III with different refractive indices as shown in Fig. 1. The incident wave vector \mathbf{k} lies on the incident plane (i.e. the x-z plane) and its magnitude in regions I and III is given by

$$k_{I,III} = \frac{2\pi n_{I,III}}{\lambda} = n_{I,III}k \quad (1)$$

where n indicates the refractive indices of region I and III, and λ is the wavelength of incidence from free space. The extinction coefficients in regions I and III are given by κ_I and κ_{III} , respectively. The grating region of the structures consists of a periodic arrangement of two different materials. Accordingly, the dielectric function of the grating is a periodic function of x with a period Λ . The RCWA formulations for the transverse electric (TE) and transverse magnetic (TM) waves of this structure are derived in [2]. The following discussions then focus arbitrarily upon the TM wave for short.

The magnetic field in region I is described as the sum of the incident and reflected waves, i.e.

$$H_I(x, z) = \exp[i(k_x x + k_z z)] + \sum_j H_{Ij} \exp[i(k_{xj} x - k_{I,zj} z)] \quad (2)$$

where k_x and k_z are the x - and z -components of \mathbf{k} (with $k_x = kn_I \sin \theta$ and $k_z = kn_I \cos \theta$), j is the number of the diffraction order, H_{Ij} is the amplitude of the j th order reflected wave, and k_{xj} is determined from the Bloch-Floquet condition [1] as

$$k_{xj} = \frac{2\pi}{\lambda} \sin \theta + \frac{2\pi}{\Lambda} j \tag{3}$$

On the other hand, the z-component of \mathbf{k} for the j th order reflected diffraction is as follows:

$$k_{I,zj} = \begin{cases} (k_I^2 - k_{xj}^2)^{1/2}, k_I > k_{xj} \\ i(k_{xj}^2 - k_I^2)^{1/2}, k_{xj} > k_I \end{cases} \tag{4}$$

In (3), $2\pi/\Lambda$ is the magnitude of the grating vector \mathbf{K} . $k_{III,zj}$ is derived as same as $k_{I,zj}$ as Eq. (4), but they have different refractive index.

The magnetic field in region III of the grating structure is the superposition of all the transmitted waves, i.e.

$$H_{III}(x, z) = \sum_j H_{IIIj} \exp(ik_{xj}x + ik_{III,zj}z) \tag{5}$$

where H_{IIIj} is the amplitude of the j th order transmitted wave in region III. The magnetic field in the region II of the grating is given by

$$H_{II}(x, z) = \sum_j S_j(z) \exp(ik_{xj}x) \tag{6}$$

where $S_j(z)$ is the amplitude of the j th space harmonic component. The j th order here is matched with the diffraction order in the other regions of the structure. In the region II, the tangential electric (y-component) and magnetic (x-component) fields are expanded into Fourier series, and the amplitude of the j th order reflected (\mathbf{R}_j) and transmitted (\mathbf{T}_j) waves are then obtained by applying the boundary conditions ($z = 0$, and $z = d$). Finally, the diffraction efficiencies are computed from the time-averaged Poynting vector in the z -direction. The absorptance (A) of the structure is obtained by subtracting the total energy of the reflected (R) and transmitted (T) light from unity based on the law of energy conservation, i.e. $A = 1 - (R+T)$.

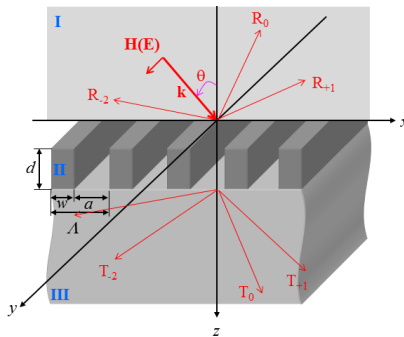


Fig. 1. Schematic illustration of gratings with Λ is the grating period, w and a are the grating width ($w = f \Lambda$ with f being the filling factor) and groove width, respectively, d is the thickness layer, and \mathbf{R} and \mathbf{T} are the reflected and transmitted diffraction orders. The subscript denotes number of diffraction orders.

3 Applications: Solar Absorbers and Optical Filtering Sensors

3.1 Solar Absorbers for Solar Cells

Figure 2 shows a schematic illustration of the complex grating structure studied in this paper. The complex grating is defined as the superposing of grating profiles of two or more one-dimensional grating ones as shown in Fig. 2. The structure consists of a complex Si grating on top of an Si substrate with the period (Λ), the lamella width (b), and the grating thickness (d). Fig. 2(a) shows the optimized grating with smooth grating sidewalls. Figs. 2(b)-(d) illustrate five cases of imperfectly geometric shapes of the optimized absorber: (b) symmetric rectangular features attached at the bottom of grating sidewalls, (c) asymmetric features at the bottom of grating sidewalls, (d) rectangular features with their heights h lengthened to $2h$ vertically at the bottom, (e) symmetric rectangular features attached at the top and the bottom of grating sidewalls, and (f) symmetric rectangular features attached at the top of grating sidewalls. It is note that the optimized parameters of the perfect grating structure are obtained from [7], i.e., the grating period $\Lambda = 130$ nm, the grating thickness $d = 80$ nm, the lamella width $b_1 = 120$ nm and $b_2 = 60$ nm, and the square feature size $h = 10$ nm. In the present analysis, the incident light including TM waves travels through free space with an angle θ between the wavevector \mathbf{k} and the surface normal \mathbf{z} . \mathbf{H} and \mathbf{E} denote oscillation directions of magnetic and electric fields, respectively. In computing optical absorbance of the grating structures, the dielectric function of Si is taken from Ref [24] and is expressed as $\epsilon_\lambda = (n + i*\kappa)^2$ that are obtained by interpolating neighboring data of the extinction index κ and the refractive index n .

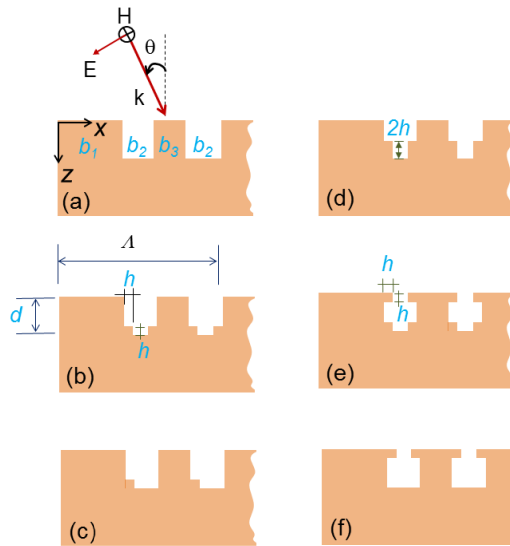


Fig. 2. Schematic illustration of perfect and complex grating structures with the grating period Λ , the lamella widths b_1 and b_3 and the groove width b_2 , the nanoscale size feature h , and the identical grating thickness d

Figure 3 shows the absorbance spectra of the perfect grating and different complex grating structures for the TM normal incidence with wavelengths ranging from 300 ~1100 nm. As shown in Fig. 3, the behavior of the absorbance spectrum of grating structures (a), (b), and (c) in the whole wavelength range is not much different while that of the grating (d), (e), and (f) significantly fluctuates in the wavelength range of 400 and 500 nm, e.g., the grating (f). From an inspection of the right-hand side of Fig. 3, it is seen that the TM absorbance valley C and peak F of the grating (f) at angle of $\theta = 0^\circ$ are equal to 0.88 (at $\lambda = 460$ nm) and 1.0 (at $\lambda = 540$ nm,) which varies much more than those of the others. However, the average absorbance of all types of the complex grating is obtained with 91% while that of the perfect one is of 92%. Overall, the results presented in Fig. 3 demonstrate that the imperfectly geometric shapes of the complex grating also provides a good optical performance for solar absorbers as the optimized structure.

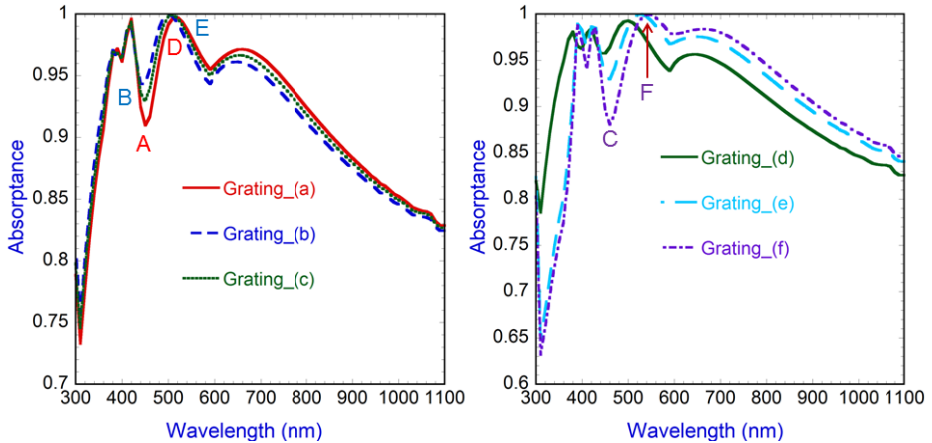


Fig. 3. Absorbance spectra of the studied grating structures for TM waves at normal incidence

3.2 Optical Transmission Filter Sensor

Figure 4 shows a schematic illustration of the one-directional double-sided grating (DSG) structure considered in this study. Its transmission properties are studied in the visible spectrum ($380 \text{ nm} < \lambda < 780 \text{ nm}$). As it is seen, the DSG concludes two identical gold (Au) single-layered gratings (SG) with thicknesses d_1 and d_2 placed in air and separated by a silicon dioxide (SiO_2) layer with a thickness h . Each grating layer has an identical period $\Lambda = 300 \text{ nm}$, the grating widths $w_1 = \Lambda f = 150 \text{ nm}$, the grating thickness $d_1 = d_2 = 150 \text{ nm}$, and the dielectric thickness $h = 50 \text{ nm}$. The coordinate system used to calculate optical transmission of the DSG as same as that shown in Fig. 1, and the dielectric function of materials are taken from [25].

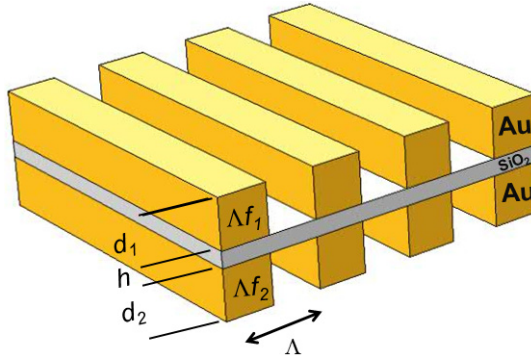


Fig. 4. Schematic illustration of Au DSG structure with grating period Λ , Au grating thicknesses of the first and second layers d_1 and d_2 , respectively, SiO₂ thickness h , and filling factor of the first grating and the second grating f_1 and f_2 , respectively

Figure 5 (a) illustrates the transmission spectra of the Au DSG structures with the same geometric parameters as described above but different both of filling factors of the grating, ranging from 0.5 ~ 1.0 for wavelengths ranging between 380 nm and 780 nm at normal incidence. It is seen that the minimum and maximum transmission peaks of the DSG at points A and B have values of approximately 0.05 and 0.71 at wavelengths of 474 nm and 650 nm, respectively. The maximum transmission peak shifts to longer wavelengths while the minimum one locates at the same wavelength. Moreover, the transmission intensity of the DSG decreases and the transmission spectrum exhibits two separate peaks when grating grooves are completely made of Au. Generally, Fig. 5(a) reveals that the transmission spectrum of the DSG can be tuned by changing the filling factors.

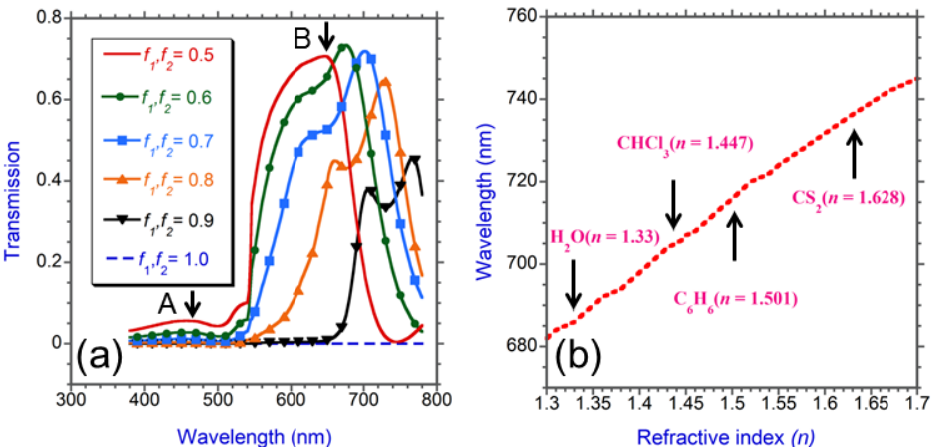


Fig. 5. (a) Transmission spectra of Au DSG structures with changing both filling factors f_1 and f_2 ; (b) Resonant wavelengths with respect to changing refractive indices of surroundings in the visible region

Fig. 5(b) represents behavior of the transmission spectrum of the proposed DSG structure, which can be employed as a sensor. It is observed that the transmission peak B at resonant wavelengths linearly shifting with changing refractive indices of surrounding media ranging from 1.3 to 1.7. In this range of refractive indices, fluids to be detected include water (H_2O), chloroform (CHCl_3), benzene (C_6H_6), and carbon disulfide (CS_2). As shown, the red dashed line shows the ability of detection of the sensor when using two grating sides. Overall, the linear relation of this device confirms that it can be used as a sensor to detect different materials.

4 Conclusion

This study has briefly presented the efficient algorithm namely the RCWA and its applicability to calculate radiative properties of periodic nanostructures featuring a simple and complex grating geometry. Two example of applications of solar absorbers for solar cells and an integrated optical filtering and refractive indices sensing device were demonstrated and showed that the proposed structures exhibit very high optical performance. The solar absorber features high absorbance in the visible and near infrared regions while the plasmonic device works as a tunable optical filter and sensitive sensor. Overall, the RCWA method may be applied to effectively design every periodic nanostructure for specific applications of energy harvesting systems, optical filters, biosensors, or integrated multifunctional nanodevices.

Acknowledgements. This work was supported by the project CZ. 1.05/1.1.00/02.0070 (IT4Innovations) and PostDoc project Opportunity for young researchers No. #CZ.1.07/2.3.00/30.0016. The authors also acknowledge the financial support by NSERC (Natural Sciences and Engineering Council,) CREATE (Collaborative Research and Training Experience program) program, and ASPIRE (Applied Science in Photonics and Innovative Research in Engineering) of Canada. Dr. Nghia would like to appreciate Profs. Yu-Lung Lo and Yu-Bin Chen (National Cheng Kung University-Taiwan) for their useful discussions on the RCWA method.

References

1. Moharam, M.G., Grann, E.B., Pommet, D.A., Gaylord, T.K.: Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *J. Opt. Soc. Am. A* 12, 1068–1076 (1995)
2. Li, L.: Use of Fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A* 13, 1870–1876 (1996)
3. Lalanne, P., Morris, G.M.: Highly improved convergence of the coupled-wave method for TM polarization. *J. Opt. Soc. Am. A* 13, 779–784 (1996)
4. Nguyen-Huu, N., Lo, Y.-L., Chen, Y.-B., Yang, T.-Y.: Realization of integrated polarizer and color filters based on subwavelength metallic gratings using a hybrid numerical scheme. *Appl. Opt.* 50, 415–426 (2011)

5. Chen, Y.-B., Zhang, Z., Timans, P.: Radiative properties of patterned wafers with nanoscale linewidth. *J. Heat Transfer* 129, 79–90 (2007)
6. Li, L.: Fourier modal method for crossed anisotropic gratings with arbitrary permittivity and permeability tensors. *Journal of Optics A: Pure and Applied Optics* 5, 345 (2003)
7. Nguyen-Huu, N., Cada, M., Pistora, J.: Investigation of optical absorptance of one-dimensionally periodic silicon gratings as solar absorbers for solar cells. *Opt. Express* 22, A68–A79 (2014)
8. Mutitu, J.G., Shi, S., Chen, C., Creazzo, T., Barnett, A., Honsberg, C., Prather, D.W.: Thin film solar cell design based on photonic crystal and diffractive grating structures. *Opt. Express* 16, 15238–15248 (2008)
9. Mallick, S.B., Agrawal, M., Peumans, P.: Optimal light trapping in ultra-thin photonic crystal crystalline silicon solar cells. *Opt. Express* 18, 5691–5706 (2010)
10. Nguyen-Huu, N., Lo, Y.-L.: Control of infrared spectral absorptance with one-dimensional subwavelength gratings. *J. Lightw. Technol.* 31 (2013)
11. Lin, S.Y., Moreno, J., Fleming, J.G.: Three-dimensional photonic-crystal emitter for thermal photovoltaic power generation. *Appl. Phys. Lett.* 83, 380–382 (2003)
12. Nguyen-Huu, N., Chen, Y.-B., Lo, Y.-L.: Development of a polarization-Insensitive thermophotovoltaic emitter with a binary grating. *Opt. Express* 20, 5882–5890 (2012)
13. Asano, T., Mochizuki, K., Yamaguchi, M., Chaminda, M., Noda, S.: Spectrally selective thermal radiation based on intersubband transitions and photonic crystals. *Opt. Express* 17, 19190–19203 (2009)
14. Nguyen-Huu, N., Lo, Y.-L., Chen, Y.-B., Yang, T.-Y.: Subwavelength metallic gratings as an integrated device: polarized color filter. In: *SPIE OPTO*, vol. 7934, pp. 79340U–79340U-10 (2011)
15. Dewan, R., Marinkovic, M., Noriega, R., Phadke, S., Salleo, A., Knipp, D.: Light trapping in thin-film silicon solar cells with submicron surface texture. *Opt. Express* 17, 23058–23065 (2009)
16. Nguyen-Huu, N., Lo, Y.-L.: Tailoring the optical transmission spectra of double-layered compound metallic gratings. *IEEE Photon. J.* 5, 2700108 (2013)
17. Ebbesen, T.W., Lezec, H.J., Ghaemi, H.F., Thio, T., Wolff, P.A.: Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* 391, 667–669 (1998)
18. Porto, J.A., García-Vidal, F.J., Pendry, J.B.: Transmission resonances on metallic gratings with very narrow slits. *Phys. Rev. Lett.* 83, 2845–2848 (1999)
19. Cao, Q., Lalanne, P.: Negative role of surface plasmons in the transmission of metallic gratings with very narrow slits. *Phys. Rev. Lett.* 88, 57403–57407 (2002)
20. Gates, B.D., Xu, Q., Stewart, M., Ryan, D., Willson, C.G., Whitesides, G.M.: New approaches to nanofabrication: molding, printing, and other techniques. *Chem. Rev.* 105, 1171–1196 (2005)
21. Barbara, A., Quémerais, P., Bustarret, E., Lopez-Rios, T.: Optical transmission through subwavelength metallic gratings. *Phys. Rev. B* 66, 161403 (2002)
22. Chen, Y.-B., Lee, B.J., Zhang, Z.M.: Infrared radiative properties of submicron metallic slit arrays. *J. Heat Transf.-Trans. ASME* 130 (2008)
23. Nguyen-Huu, N., Cada, M., Pistora, J.: Imperfectly geometric shapes of nanograting structures as solar absorbers with superior performance for solar cells. *Opt. Express* 22, A282–A294 (2014)
24. Green, M.A.: Self-consistent optical parameters of intrinsic silicon at 300K including temperature coefficients. *Sol. Energ. Mat. Sol.* 92, 1305–1310 (2008)
25. Palik, E.D.: *Handbook of Optical Constants of Solids*. Academic (1998)

An Improved Algorithm of Network Maximum Flow Based on Network Flow Matrix

ZhenChao Wang^{1,2}, LiPing Zhang¹, and WeiDong Hao³

¹ College of Electronic and Information Engineering, Hebei University, Baoding, China

² Key Laboratory of Digital Medical Engineering of Hebei Province

³ School of Civil Engineering, Beijing Jiaotong University, Beijing, China

{wangzhenchaohbdx, zhangliping1320}@163.com,

haoweidong094@126.com

Abstract. In this paper, the algorithm of network maximum flow based on network flow matrix is introduced first, then some new properties of network flow matrix are studied according to max-flow min-cut theorem, and on this basis the improved algorithm of network maximum flow based on network flow matrix is given. In this way, the matrix can be reduced order directly when it satisfies certain particular conditions and be translated into several matrices when its order is large, thus simplifying the original algorithm.

Keywords: Network Flow Matrix, Network Maximum Flow, Max-flow Min-cut Theorem.

1 Introduction

The problem of network maximum flow, as an important basis for optimizing a network and evaluating the performance of a network, refers to the maximum flow the network allowing to be transmitted from the source to the destination and how to reach the maximum flow under the situation of constrained capacity. In fact, many practical problems can be converted into maximum flow problem with a slight variation or modification, such as the maximum capacity of the port by cargo, production plan, equipment replacement and investment plan, etc. Therefore, it has great theoretical and practical significance to research the problem of network maximum flow. And the classical algorithms of network maximum flow are mainly based on two strategies. One is augmenting path strategy, such as Ford-Fulkerson algorithm [1], Edmond and Karp's capacity scaling algorithm [2], Dinic algorithm [3], etc. The other is preflow-push strategy, like Goldberg-Tarjan algorithm [4], Park-Pao algorithm [5], etc. On this basis, many scholars improve and optimize one of the algorithms above [6,7] or put forward some new algorithms [8-10] which are different from the two strategies. For instance, SUN et al. [8] regarded the solving process of network maximum flow problem as a dynamic adjustment on the basis of dynamic programming idea. They also compared the performance of the algorithm with that of the classical Ford-Fulkerson algorithm in literature [9]. Literature [10] proposed a new

algorithm of network maximum flow based on network flow matrix with the idea of the node flow balance, problem conversion and reducing the matrix order, which provides a new way to solve the maximum flow problem. However, the order of the network flow matrix is large when a large number of nodes exist in a network, and therefore reducing its order is more complicated.

This paper studies some new properties of network flow matrix according to max-flow min-cut theorem, and with these properties the matrix can be reduced order directly when it satisfies certain particular conditions and be translated into several matrices when its order is large, thus improving the algorithm in literature [10].

2 Algorithm of Network Maximum Flow Based on Network Flow Matrix

In literature [10], the basic idea of getting network maximum flow with network flow matrix is reducing the order of strictly upper triangular matrix initial network flow matrix A without changing its maximum flow. The specific steps of the algorithm based on network flow matrix are as follows:

First step: Establish the directed acyclic graph from the source node to the sink node, and label every node in up-to-down order, so that the labeling of each node is greater than its parent nodes and less than its child nodes.

Second step: Write out the initial network flow matrix $A=(a_{ij})_{n \times n}$, which is also a strictly upper triangular matrix. Then calculate the balance numbers α_i of each row and β_j of each column. If the balance numbers of all the rows (or all the columns) are zero, the maximum flow is $V(A)=\sum_{i=1}^n a_{in}$ (or $V(A)=\sum_{j=1}^n a_{1j}$), else move on to the third step.

Third step: Reduce the order of the matrix $A=(a_{ij})_{n \times n}$ recurrently.

(1) Check whether $\alpha_i=\beta_j=\alpha_i=\beta_j=0$ is satisfied, if it is, reduce the order with the reducing order formula directly, else move on to (2) .

(2) Simplify matrix A with the properties of network flow matrix in literature [10] until $\alpha_i=\beta_j=\alpha_i=\beta_j=0$, and then the matrix $A=(a_{ij})_{n \times n}$ is acquired.

(3) Reduce its order with the reducing order formula, and then the new network flow matrix $A_{n-1}=A=(a_{ij})_{(n-1) \times (n-1)}$ is acquired.

Fourth step: Perform the third step recurrently until the matrix $A_2=\begin{bmatrix} 0 & a_{12} \\ 0 & 0 \end{bmatrix}$ is acquired.

Fifth step: The element a_{12} is the maximum flow of the network.

The algorithm is simple to operate and easy to implement, but before reducing the order all the balance numbers should be calculated, and if the matrix can not be deduced order, the properties of network flow matrix should be used to simplify the matrix. So the process of reducing the order is complicated especially when a large number of nodes exist in a network. Here, we will study some new properties of network flow matrix from the perspective of the relation between network flow matrix and the physical meaning of max-flow min-cut theorem.

3 New Properties of Network Flow Matrix

3.1 Max-flow Min-Cut Theorem

Let $G=(V,E)$ be a communication network, nodes $s,t \in V$ are the source node and the sink node respectively. We divide nodes set V into two sections V_1 and V_2 , such that $s \in V_1, t \in V_2$ and $V_1 \cup V_2 = V, V_1 \cap V_2 = \Phi$. The cut of source node s to sink node t refers to the set of all the edges whose starting point and endpoint are in set V_1 and V_2 respectively, and it is denoted by $C(V_1, V_2)$. The capacity of cut $C(V_1, V_2)$ refers to the combined capacity of all the edges whose starting point and endpoint are in set V_1 and V_2 respectively, and it is denoted by $|C(V_1, V_2)|$. Then the max-flow min-cut theorem is that the maximum flow in G equals the capacity of the minimum cut $C(V_1, V_2)$, namely $\max flow(s,t) = \min |C(V_1, V_2)|$.

3.2 New Properties of Network Flow Matrix and the Proofs

Property 3-1: If the elements of the i th row (or the i th column) in network flow matrix A are all zero ($2 \leq i \leq n-1$), the maximum flow of matrix A remains unchanged after the i th row and the i th column are removed.

Proof: The elements of the i th row in matrix A are all zero means that node i has no outgoing edges, which indicates that there are no paths from node i to sink node t obviously. That is to say, there do not exist the path which traverses node i from source node s to sink node t . So, the maximum flow of sink node t keeps unchanged, after node i and its corresponding edges are removed, namely the i th row and the i th column of matrix A are removed from the perspective of matrix. Similarly, if the elements of the i th column in the matrix A are all zero, namely node i has no incoming edges, the maximum flow of matrix A remains unchanged after the i th row and the i th column are removed.

Property 3-2: If only one non-zero element respectively exists in the i th row and the i th column ($2 \leq i \leq n-1$) of the network flow matrix A , and they are a_{ij} and a_{ki} , the maximum flow of matrix A remains unchanged after a_{ij} is changed to $a_{ij} + \min\{a_{ij}, a_{ki}\}$ as well as the i th row and the i th column are removed.

Proof: The condition of the property is that, the i th row and i th column of the network flow matrix A has only one non-zero element a_{ij} and a_{ki} respectively, which is to say intermediate node i has only one incoming edge (k,i) and one outgoing edge (i,j) .

The conclusion of the property is that removing node i and its corresponding edges, as well as adding $\min\{a_{ij}, a_{ki}\}$ to the flow of edge (k,j) , the maximum flow of the network keeps unchanged. Since the maximum flow of the corresponding network of matrix A is $V(A)$, there exists $V(A)$ edge-disjoint paths from source node s to sink node t . When node i is not on any edge-disjoint paths, it is obvious that the above operation does not change these paths, so the maximum flow of the network keeps unchanged. When node i is on an edge-disjoint path P , as node i has only one incoming edge (k,i) and one outgoing edge (i,j) , node k and node j must be on the edge-disjoint path P , which is equivalent that there existing an edge from nodes k to node j and its flow is $\min\{a_{ij}, a_{ki}\}$. So, removing node i and its corresponding edges as well as adding $\min\{a_{ij}, a_{ki}\}$ to the flow a_{kj} of

edge (k, j) , such that edge-disjoint path P is changed to P' and the other paths stay the same, the maximum flow of the network keeps unchanged.

Lemma 3-1: If the sink node is t , the node set $J = \{v \in V, (v, t) \in E\} = \{v_1, v_2, \dots, v_k\}$, there exists $v_i \in J, 1 \leq i \leq j \leq k$ such that $\forall j \neq i, (v_i, v_j) \notin E$

Proof: The lemma is to say, all the nodes which are connected directly with the sink node t form a set J , and in set J there exists at least one node which has no outgoing edge in the set. Here we use reduction to absurdity to prove the lemma. The negative proposition of the proposition is that all the nodes in set J must have outgoing edges in the set. Assuming that the negative proposition is true, that is for any node $v_i \in J, 1 \leq i \leq k$ there exists $i \neq j$ such that $(v_i, v_j) \in E$. We consider node v_1 , then there must exist $j \neq 1$ such that $(v_1, v_j) \in E$, so we can suppose $j = 2$, then $(v_1, v_2) \in E$. Similarly, we consider node v_2 , then there exists $j = 3$ such that $(v_2, v_3) \in E, \dots, (v_{k-1}, v_k) \in E$. Considering node v_k , there must exist $j \neq k$ such that $(v_k, v_j) \in E$, where j must be between 1 and $k-1$. In this way, there is a circle $v_j v_{j+1}, \dots, v_k v_j$ in the network G , which conflicts with the fact that G is an acyclic graph. So the assumption is not true and the original proposition is true, namely in set J there exists at least one node which has no outgoing in the set.

A is a n -order initial network flow matrix which is also a strictly upper triangular matrix. We consider the $(n-1)$ th node, as the element $a_{n-1,n}$ in the $(n-1)$ th row is a non-zero element, namely the $(n-1)$ th node is connected directly with the sink node t obviously, so it belongs to the set J . Since the other elements in the $(n-1)$ th row of matrix A are all zero, it has no outgoing edges in set J . Therefore, the $(n-1)$ th node is one of the nodes which satisfy the lemma 3-1, and it is denoted by v_{n-1} .

Theorem 3-1: In a single-source single-sink directed graph G , if node v_{n-1} is not in any minimum cut,

- (1) after removing the $(n-1)$ th row and the $(n-1)$ th column of the network flow matrix A , we can get a new matrix A^* , and there exists the formula $V(A) = V(A^*) + a_{n-1,n}$;
- (2) changing a_{1n} to $a_{1n} + a_{n-1,n}$, and then removing the $(n-1)$ th row and the $(n-1)$ th column of the network flow matrix A , we can get a new matrix A° , and there exists the formula $V(A) = V(A^\circ)$.

Proof: (1) Take any minimum cut $C(V_1, V_2)$, as $v_{n-1} \notin V_2, v_{n-1} \in V_1$. Since $v_{n-1} \in J$, where J is a set in lemma 3-1, $(v_{n-1}, t) \in C(V_1, V_2)$. It is known from max-flow min-cut theorem that the capacity of $C(V_1, V_2)$ equals $V(A)$ and the contribution of $(v_{n-1}, t) \in C(V_1, V_2)$ to the capacity of minimum cut $C(V_1, V_2)$ is $a_{n-1,n}$. If the edge (v_{n-1}, t) whose capacity is $a_{n-1,n}$ is removed, $C(V_1, V_2)$ is still a minimum cut and just its capacity changes to $V(A) - a_{n-1,n}$. As the $(n-1)$ th row of matrix A has only one non-zero element $a_{n-1,n}$, the elements of the $(n-1)$ th row in matrix A are all zero after the edge (v_{n-1}, t) is removed. According to the property 3-1, after removing the $(n-1)$ th row and the $(n-1)$ th column we get a new matrix A^* , and the maximum flow remains unchanged, so $V(A^*) = V(A) - a_{n-1,n}$.

(2) Take any minimum cut $C(V_1, V_2)$, and label the source node and sink node 1 and n respectively, then $s \in V_1, t \in V_2$. So the maximum flow of matrix A increases by $a_{n-1,n}$ after a_{1n} is changed to $a_{1n} + a_{n-1,n}$. And if the new matrix is A' , $V(A') = V(A) + a_{n-1,n}$. After removing the $(n-1)$ th row and the $(n-1)$ th column of matrix A' , we can get a new matrix A° , then $V(A^\circ) = V(A') + a_{n-1,n}$ according to the theorem 3-1(1), so $V(A) = V(A^\circ)$.

Theorem 3-2: In a single-source single-sink directed graph G , if the node v_{n-1} is in a minimum cut $C(V_1, V_2)$, namely $v_{n-1} \in V_2$, after performing the following operations on the network flow matrix A , we can get a new matrix A° , then $V(A^\circ) = V(A)$.

- (1) change $a_{n-1,n}$ to 0 ; (2) add each element of the n th column to the $(n-1)$ th column ;
- (3) remove the n th row and the n th column.

Proof: The theory is to say, the maximum flow of the network G keeps unchanged, after the sink node t is removed and the edges connected with sink node t are pointed to node v_{n-1} . Since node v_{n-1} points to node t only, and it has no outgoing edges after sink node t and its edges are removed, so node v_{n-1} can be regard as the new sink node in the new network. We discuss two cases:

1) If $v_i \in V_2 (i=1,2,3,\dots,k)$, where v_i is an element of J in lemma 3-1, that is v_i is connected directly with sink node t , the maximum flow of the network keeps unchanged after the sink node t is removed and the edges connected with sink node t are pointed to node v_{n-1} . Because the edges contained by minimum cut $C(V_1, V_2)$ stay the same, and the capacity of the minimum cut $C(V_1, V_2)$ keeps unchanged.

2) If there exists $v_j \notin V_2 (1 \leq j \leq k)$, after the sink node t is removed and the edges connected with sink node t are pointed to node v_{n-1} , the contribution of (v_j, t) to the capacity of minimum cut $C(V_1, V_2)$ equals that of (v_j, v_{n-1}) to the capacity of the new cut $C(V_1, V_2^*)$, where $V_2^* = V_2 - \{t\}$. And the contributions of the other edges contained by cut $C(V_1, V_2)$ stay the same, so $V(A) = C(V_1, V_2) = C(V_1, V_2^*) = V(A^\circ)$.

It can be easily observed that the matrices A and A° got from theorem 3-1 and theorem 3-2 respectively are both $(n-1)$ -order. As we can not judge whether node v_{n-1} is in a minimum cut in advance, we take the less, namely $V(A) = \min\{V(A^\circ), V(A^*)\}$. It is obviously that the reduced-order method is different from the method in literature [10], which provides a new train of thought for reducing the order of network flow matrix.

4 The Improved Algorithm of Network Maximum Flow Based on Network Flow Matrix

As the balance numbers need not to be calculated when we reduce the order with the theorems 3-1 and 3-2, they can be used in the early stage of reducing the order as the matrix is large in most cases. In the late stage of reducing the order, the method of balance reducing order in literature [10] can be used, which is beneficial to overcome the shortcoming that the matrix need to be translated into two matrices when using the theorems 3-1 and 3-2.

Below we give the improved algorithm's specific steps of network maximum flow based on network flow matrix.

First step: Same with that of the original algorithm.

Second step: Same with that of the original algorithm.

Third step: Reduce the order of matrix A by k times with the theorem 3-1 and theorem 3-2, then we can get 2^k matrices whose order are $(n-k)$.

Fourth step: Calculate the maximum flows $\omega_i (i=1,2,3,\dots,2^k)$ of these matrices by reducing the order of each matrix with the method of balance reduced order given in the section 2.2.

Fifth step: Calculate the network maximum flow with the formula $\omega = \min_i \omega_i$.

We should notice that (1) Considering whether the property 3-1 and property 3-2 can be used to reduce the order directly before reducing the order of every matrix. (2) The choice of k should depend on the size of the network flow matrix and the property of the computer.

5 Conclusion

This paper mainly studies some new properties and theorems about network flow matrix are according to max-flow min-cut theorem, and on this basis giving the improved algorithm of network maximum flow based on network flow matrix. This paper establishes strict theoretical basis by theoretical analysis and proof, and gives the specific steps of the improved algorithm according to the theory. Since the algorithm is simple in operation and rigorous in theory, and its validation work can be implemented manually, therefore the simulation is omitted here.

References

1. Ford, L.R., Fulkerson, D.R.: Maximal flow through a network. *Canadian Journal of Mathematics* 8(5), 399–404 (1956)
2. Edmonds, J., Karp, R.M.: Theoretical improvements in algorithmic efficiency of network flow problems. *Journal of ACM* 19(2), 248–264 (1972)
3. Dinic, E.A.: Algorithms for solution of a problem of maximum flow in networks with power estimation. *Soviet Mathematical Doklady* 11(8), 1277–1280 (1970)
4. Goldberg, A.V., Tarjan, R.E.: A New Approach to the Maximum Flow Problem. *J. Assoc. Comput. Mach.* 35, 921–940 (1988)
5. Park, G.H., Pao, Y.H.: Neural-net Approach to the Maximum Flow Problem. *Electronics Letters* 34(8), 786–787 (1998)
6. Zhang, J., Qiu, X.-S.: Algorithm and Its Implementation for the Network Maximal Flow. *Journal of Chongqing University: Natural Science Edition* 29(5), 132–134 (2006)
7. Du, D.L., Chandrasekaran, R.: The multiroute maximum flow problem revisited. *Network: An International Journal* 47(2), 81–92 (2006)
8. Sun, X.-J., Wang, Z.-Q.: A New Algorithm for Solving the Maximum Flow Problem in the Network without Circle. *Journal of Jiangxi Normal University (Natural Science)* 33(3), 370–372 (2009)
9. Sun, X.-J., Wang, Z.-Q., Liu, S.-Y.: Performance Analysis of Algorithms for Solving the Maximum Flow Problem. *Mathematics in Practice Theory* 43(17), 120–124 (2013)
10. Wu, Y., Yang, Y.-L., Liu, S.-Y.: The Network Maximum flow Based on the Flow Matrix. *Systems Engineering* 25(10), 122–125 (2010)

Author Index

- Aziz, Abdul 211
- Bilek, Ondrej 79
- Byun, HwiRim 123
- Cada, Michael 343
- Chang, Jae-Woo 31, 67
- Chen, Wu-Lin 105
- Chen, Yung-Hui 53
- Cheng, Chao 183
- Cheng, Jingde 1, 7, 15, 23
- Chenglong, Pan 337
- Cho, Hyunhun 149
- Cho, Kyungeun 157, 165
- Cho, Oh-hoon 289
- Cho, So Ra 99
- Cho, Sungryung 39
- Cho, Yong-Hyun 177
- Choi, In-Ho 45
- Choi, Sunguk 273
- Choi, Yeonyi 205, 211, 217
- Chu, Phuong 157
- Chung, Sungmoon 39
- Du, Junping 239, 249
- Du, Tianming 239, 249
- Duc-Tan, Tran 191
- Ekiba, Takeo 15
- Fan, Wengui 331
- Gao, Hongbiao 1, 23
- Gil, Joon-Min 149, 177
- Goto, Yuichi 1, 7, 15, 23
- Hanh, Phan Trong 191
- Hao, WeiDong 351
- Hong, Seokjoon 217
- Hong, Sung-Hwa 177
- Hu, Yu-Chen 105
- Huang, Chi-Fu 143
- Ikram, Warda 165
- Jang, Miyoung 67
- Jang, Unsoo 129
- Jeong, Chan-Hee 45
- Jeong, Yoonji 165
- Jeong, Young-Sik 123, 231
- Joe, Inwhhee 39, 205, 211, 217
- Jung, Kyeong-Hoon 199
- Jung, Seung-Won 111, 117
- Jung, Wonjin 61
- Kan, Yantao 205
- Kang, Dong-Wook 199
- Kim, Byoung Cheul 129
- Kim, Changhee 111
- Kim, Dongho 231
- Kim, Eung-gon 281
- Kim, Eung-Kon 257, 265, 273, 289
- Kim, Hyeong-II 31
- Kim, Hyun-Woo 123
- Kim, Jee-In 297
- Kim, Sung-Hoon 199
- Kim, Taehwan 61
- Kim, Woohyun 149
- Kim, Yong-Guk 45
- Kim, Yoonkyoung 129, 135

- Kim, Yun-Jung 135
 Ko, Daejune 135
 Ko, Hye-Kyeong 73
 Krejcar, Ondrej 79

 Le, Thanh-Ha 117
 Lee, Byeonggwon 157, 165
 Lee, Eui Chul 129, 135
 Lee, Jaewon 211
 Lee, Kyungrak 205
 Lee, Seongjo 117
 Lee, Wooyeob 39
 Lee, Young-Jae 257
 Li, Yan 171
 Li, Yong 205
 Li, Zhiyong 183
 Lim, Sooyeon 223
 Lin, Wei-Chen 143
 Lin, Yuxi 183
 Lyu, Jaeha 223

 Ma, Suoning 211
 Maresova, Petra 85, 93
 Meiping, Wu 337

 Nam, Gi Pyo 99
 Nanaumi, Shunsuke 1
 Nguyen, Dat Tien 99
 Nguyen, Ngoc-Hoa 317
 Nguyen-Huu, Nghia 343

 Oh, Yeon-Jae 265

 Park, Hyun-A. 297
 Park, JaeHyun 297
 Park, Jeman 39
 Park, Jeongwook 149
 Park, Jinhyung 149
 Park, JongSung 297
 Park, Kang Ryoung 99
 Park, Kyoung-Wook 265
 Park, Kyung-Wook 257
 Park, Mi-jeong 281
 Park, Min Woo 129
 Pistora, Jaromir 343

 Shin, Byeong-Seok 171
 Shin, Kwang Yong 99
 Shin, YoungSung 31
 Song, Eun-Ha 123
 Song, Youngho 67
 Suh, Kun Ha 135

 Thuy, Nguyen Thanh 191
 Tsai, PiYu 105

 Um, Kyhyun 157, 165

 Van Tam, Vu 191
 Van Thanh, Tran 231
 Vu, Hoang 157

 Wagatsuma, Kazunori 1
 Wang, Bo 7, 15
 Wang, Jun 323, 331
 Wang, Jyu-Wei 53
 Wang, Xiaoliang 323
 Wang, Xiaoru 239, 249
 Wang, Xingjun 183
 Wang, Yuanyou 239, 249
 Wang, Zhe 7
 Wang, ZhenChao 351
 Won, Chee Sun 111, 117
 Woo, Dong-Min 317

 Xiang, Hong 323
 Xiaogang, Ji 337
 Xueming, He 337

 Yang, Bin 331
 Yang, Lina 309
 Yang, Yi 309
 Yeh, Jieh-Shan 105
 Yeo, Donghyeon 157
 Yoon, Min 31, 67
 Yun, Seokmin 111

 Zhang, LiPing 351
 Zhang, Yuxi 323, 331
 Zhao, Zhipeng 323, 331
 Zhou, Yuan 7