Natarajan Meghanathan
David C. Wyld (Eds)

Computer Science & Information Technology

7th International Conference on Signal Image Processing and Multimedia (SIPM 2019) March 23-24, 2019, Sydney, Australia

AIRCC Publishing Corporation

Volume Editors

Natarajan Meghanathan,
Jackson State University, USA
E-mail: nmeghanathan@jsums.edu

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

ISSN: 2231 - 5403
ISBN: 978-1-921987-99-1
DOI : 10.5121/csit.2019.90301- 10.5121/csit.2019.90314

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The 7th International Conference on Signal Image Processing and Multimedia (SIPM 2019) March 23-24, 2019, Sydney, Australia. The 7th International Conference of Advanced Computer Science & Information Technology (ACSIT 2019), 7th International Conference of Information Technology, Control and Automation (ITCA 2019), 7th International Conference on Information Technology in Education (ICITE 2019) was collocated with 7th International Conference on Signal Image Processing and Multimedia (SIPM 2019). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The SIPM 2019, ACSIT 2019, ITCA 2019, ICITE 2019 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, SIPM 2019, ACSIT 2019, ITCA 2019, ICITE 2019 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SIPM 2019, ACSIT 2019, ITCA 2019, ICITE 2019

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Natarajan Meghanathan
David C. Wyld

Organization

General Chair

Natarajan Meghanathan	Jackson State University, USA
David C. Wyld	Southeastern Louisiana University, USA

Program Committee Members

Almir Pereira Guimarães,	Federal University of Alagoas, Brazil
Aberto Magrenan,	International University of La Rioja (UNITE), Spain
Ahmad T. Al-Taani,	Yarmouk University, Jordan
Ahmed Mohamed Khedr,	sharjah university, UAE
Ali Javadi,	Iran University of Science and Technology, Iran
Ali Salem,	University of Sfax,Tunisia
Ameera Saleh. Jaradat,	Yarmouk University, Jordan
Amel B.H.Adamou-Mitiche,	University of Djelfa, Algeria
Ammar Al-Masri,	Albalqa Applied University, Jordan
Amir Salarpour,	Bu-Ali Sina University, Iran
Andre F.Dantas,	UnP - Universidade Potiguar, Brasil
Ankit Chaudhary,	Truman State University, USA
Ava Clare Marie O. Robles,	Mindanao State University, Philippines
Abdulwahid Abdullah Al Abdulwahid,	Jubail University College, KSA
Aref Tahmasb,	Shahid Bahonar University, Iran
Christian Mancas,	Ovidius University, Romania
Chang-Hyun,	Korea Marine Equipment Research Institute, Korea
Chin-Chih Chang,	Chung-Hua University, Taiwan
Daniel Gomes,	Estacio de Sa, Brasil
Djekoune A. Oualid,	Advanced Technologies Development Centre, Algeria
Duta Cristina-Loredana,	University Politehnica of Bucharest, Romania
Emad Al-Shawakfa,	Yarmouk University, Jordan
El Miloud Ar Reyouchi,	Abdelmalekessaadi University, Morocco
Elena Battini Sonmez,	Istanbul Bilgi University, Istanbul
Essam Halim Houssein,	Minia University Egypt
Florence SEDES,	Toulouse University, France
Gabriel Badescu,	University of Craiova, Romania
Gebeyehu Belay,	Bahir Dar University, Institute of Technology, Ethiopia
Hamid Ali Abed AL-Asadi,	Basra University, Iraq
Hamed Al-Rubaiee,	University of Bedfordshire, United Kingdom
Hamdi Yalin Yalic,	Hacettepe University, Turkey
Isa Maleki,	Islamic Azad University, Iran
Islam Atef,	Alexandria university, Egypt
Ivo Bukovsky,	Czech Technical University, Czech Republic
Ivan Izonin,	Lviv Polytechnic National University, Ukraine
Jafar Mansouri,	Ferdowsi University of Mashhad, Iran

Jalel Akaichi,	University of Tunis, Tunisia
Jamal El Abbadi,	Mohammadia V University Rabat, Morocco
Jun-Cheol Jeon,	Kumoh National Institute of Technology, South Korea
Jun Zhang,	South China University of Technology, China
Keneilwe Zuva,	University of Botswana, Botswana
Klimis Ntalianis,	Athens University of Applied Sciences, Greece
Mohamed HAMLICH,	Hassan II University, MOROCCO
Marcelo Seido Nagano,	University of Sao Paulo, Brazil
Mario Henrique Souza Pardo,	University of Sao Paulo, Brazil
Marius CIOCA,	Lucian Blaga University of SIbiu, Romania
Maryam hajakbari,	Islamic Azad University, Iran
Mamy alain Rakotomalala,	Universite d'Antananarivo, Madagascar
Madya Dr. Mohammad Bin Ismail,	Universiti Malaysia Kelantan, Malaysia
Marco Furini,	Universita Di Modena E Reggio Emilia, Italy
Masoud Ziabari,	Mehr Aeen University, Iran
Miguel A. Rodriguez-Hernandez,	ITACA Universitat Politecnica de Valencia, Spain
Mohiy Mohamed Hadhoud,	Menoufia university, Egypt
Mohamed Khayet,	University Complutense of Madrid, Spain
Mohamed SENOUCI,	Universite d'Oran 1 Ahmed Ben Bella, Algeria
Murat Tolga OZKAn,	Gazi University Fakulty of Technology, Turkey
Noura Taleb,	Badji Mokhtar University, Algeria
Narges Shafieian,	Azad University, Iran
Nicolas H. Younan,	Mississippi State University, USA
Noureddine Idboufker,	Cadi Ayyad University, Morocco
Omar Boussaid,	University of Lyon, France
Olufade F. W. Onifade,	University of Ibadan, Nigeria
Ouafa Mah,	Ouargla university, Algeria
Paulo Roberto Martins de Andrade,	University of Regina, Canada
Prakash Duraisamy,	University of Central Missouri, United States
Pradnya Kulkarni,	Federation University, Australia
Quang Hung Do,	University of Transport Technology, Vietnam
Rafael Stubs Parpinelli,	State University of Santa Catarina, Brazil
Ruchi Tuli,	Jubail University College, Kingdom of Saudi Arabia
Rougang Zhou,	Huazhong University of Science & Technology, China
Roberto De Virgilio,	Roma Tre University, Italy
Shoeib Faraj,	Institute of Higher Education of Miaad, Iran
Shirish Patil,	Independent/Industry, USA
Subarna Shakya,	Tribhuvan University , Nepal
Saadat Pourmozafari,	Tehran Poly Technique, Iran
Savita Wali,	Basaveshwar Engineering College, Bagalkot
Sriharee	King Mongkut's University of Technology, Thailand
Seyyed Mohammad Reza Farshchi,	Ehran University, Iran
Siuly Siuly,	Victoria University, Australia
Smain Femmam,	UHA University France, France
Saban Gulcu,	Necmettin Erbakan University, Turkey
Terumasa AOKI,	Tohoku University, Japan
Wahiba Ben Abdessalem,	High Institute of Management of Tunis, Tunisia
Wenwu Wang,	University of Surrey, United Kingdom
Willie K Ofosu,	Penn State Wilkes-Barre, USA

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

7th International Conference on Signal Image Processing and Multimedia (SIPM 2019)

7th International Conference of Advanced Computer Science & Information Technology (ACSIT 2019)

7th International Conference on Information Technology in Education (ICITE 2019)

7th International Conference of Information Technology, Control and Automation (ITCA 2019)

SIMILARITY BASED CLASSIFICATION AND DETECTION OF RESPIRATORY STATUS IN FREQUENCY DOMAIN

Suyeol Kim[1], Chaehwan Hwang[1], Jisu Kim[2], Cheolhyeong Park[2] and Deokwoo Lee[2]

[1]Department of Biomedical Engineering, Keimyung University, Daegu, Republic of Korea
[2]Department of Computer Engineering, Keimyung University, Daegu, Republic of Korea

ABSTRACT

Sleep apnea is considered one of the most critical problems of human health, and it is also considered one of the most important bio-signals in the area of medicine. In this paper, we propose the approach to detection and classification of respiratory status based on cross correlation between normal respiration and apnea, and on the characteristics of respiratory signals. The characteristics of the signals are extracted by analyzing frequency analysis. The proposed method is simple and straightforward so that it can be workable in practice. To substantiate the proposed algorithm, the experimental results are provided.

KEYWORDS

Respiration, Apnea, Fourier transform, Detection, Classification

1. INTRODUCTION

Obstructive breathing disease can lead to severe health problems such as abnormal lung conditions (asthma, emphysema, etc.), heart disease, allergies, etc. Extensive research has been carried out to find methods for diagnosis, treatment or prevention. In the areas of signal analysis and processing, detection and monitoring of respiratory signal have long been of interest, and to achieve efficient measurement of respiration also has been extensively investigated [1]. Major methods of signal acquisition can be categorized in invasive or non-invasive medical instruments [2, 3]. In general, vital signs (also called bio signal) have been able to play a key role to indicate the status of human health. Among the other signs, body temperature, heart rate, blood pressure, electrocardiogram and respiratory rate are considered the most primary vital signs [4]. Furthermore, recording the signs above is also considered the standard for monitoring the health status of patients in hospital ward. However, it has been reported that respiratory signal has gained less attention than the other three signals even though respiratory status can be both an important indicator and a predictor of severe illness [4]. Since a few years, there have been extensive research on detecting sleep apnea one of the breathing disorders. Sleep apnea can cause cardiovascular problem, so there have been research on detection and monitoring of the breathing status using invasive or non-invasive methods (or contact based or non-contact based methods).

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 01–09, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90301

Contact based monitoring methods use devices that are attached to human body, but the methods undergo difficulties when the patients suffer from generating respiratory signal. Besides, the monitoring equipment is usually expensive, gives discomfort to testee people, and they need to be guided how to generate their respiratory signal that can sufficiently provide meaningful analysis results. However, the contact based method, if it is stable setup for use, acquired signal is reliable, leading to accurate measurement and analysis. On the other hand, non-contact or non-invasive method can be alternatives to the contact based methods. They can be audio based, temperature based or vision based or distance based (e.g., UWB radar) methods [5, 6]. In general, non-contact based devices are, in general, low cost and easy to use compared to the contact based ones. This paper chiefly focuses on classification of respiratory signal that is acquired using a contact-based medical instrument. This work aims at detecting a signal of an abnormal respiration. Abnormal respiration in this work is apnea. Respiratory signal is defined in 1-dimensional domain, and the classification is basically performed using correlation coefficient between the signals. The proposed approach is expected to maximize inter-class distance and to minimize intra-class distance. The overall description of the present work is shown in Figure. 1.

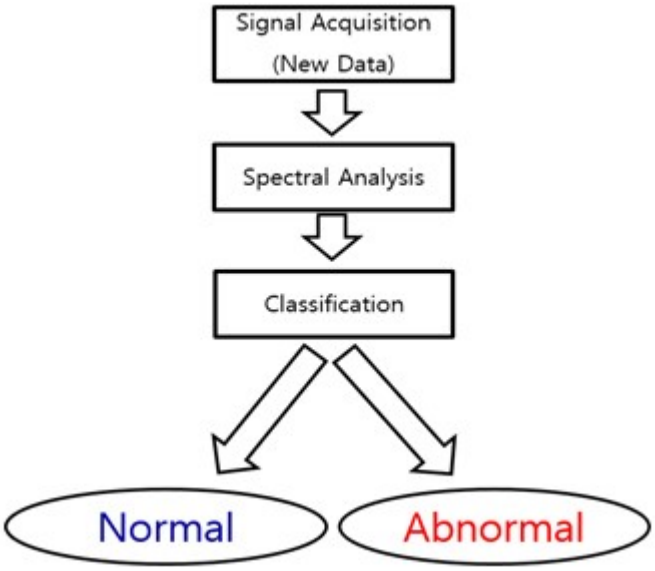


Figure 1. Overall architecture of a classification based monitoring and detection of respiratory status

Rest of the paper is organized as follows. In section II, preprocessing, noise reduction using Savitzky-Golay and median filter are introduced, and the filtering results are also briefly provided. In section III, similarity between the different respiratory signals are estimated so that classification of normal and abnormal status can be performed. Experimental results are provided in section IV, and we conclude this paper and suggest future direction.

2. NOISE REDUCTION

Respiratory signal is acquired using a UWB radar sensor, NOVELLDA X4, and the device is not only used for respiration measurement, but also for detection of dynamic activities. The measurement is based on the calculating the distance by estimating duration of reflection of radar

signal. The device can provide denoised signal so that users do not have to endeavor to reduce or remove noise component. In case of existence of noise component, we have used a filter that combines Savitzky-Golay filter and median filter [7]. The filter not only reduces noise component but also maintains high frequency component of a signal. Savitzky-Golay (SG) filter has been used in the areas of signal denoising by smoothing the signal with noise of high frequency components. Let be an original signal (this signals is ideally noiseless signal), random noise is added to the signal. Let and be the denoised signal and the signal with noise, respectively, is represented as follows.

$$x_n(t) = x(t) + n(t) \qquad (1)$$

and is random noise. The noise model is not discussed in this paper because it is beyond the scope of the present work. SG filtering process can be represented as an average of convoluted noised signal, and written as follows.

$$x_d(t_j) = k \cdot \sum_{i=-M}^{M} C_i \cdot x_n(t_{j+i}) \qquad (2)$$

where k is $\frac{1}{2M+1}$, a $\cdot$ b is an inner product of a and b, $2M+1$ is the smoothing window, C_i is the smoothing coefficient (or convolution coefficient), respectively [7]. The essential idea behind SG filter is signal smoothing based on polynomial approximation using least-squares method. Due to the limitation of the SG filter, as explained in the next section, median filter is applied to the signal. Median filter, nonlinear filter, has been popularly used in image processing. Median filtering can be simply performed by extracting the median value in moving window. In the present work, since the signal of the interest is 1-dimension, the running window is simply a row vector (or column vector). The size (or length) of a vector can be adjusted according to the areas of applications. While the basic idea of median filter is very simple, it preserves high frequency components of a signal. Intuitively, median filter still show limitation in that it achieves successful preservation of high frequency components in certain conditions and loses information in boundary areas of a signal. Concerning all above, this paper proposes the new filter that combines SG filter, $h_{sg}(t)$ and median filter, $h_m(t)$. SG filter shows significant performance in smoothing signal, and median filter shows strength in preservation of edge information. This paper, hence, fully exploits both properties of the filters. The proposed filter, $h_c(t)$ is applied to the noised signal, and it reduces the noise components while preserves edge information, and $h_c(t)$ is written as

$$x_d(t) = h_c(t) * x_n(t) \qquad (3)$$

where $h_c(t) * x_n(t)$ represents the convolution between $h(t)$ and $x_n(t)$, and $h_c(t) = h_m(t) * h_{sg}(t)$. $h_c(t)$ is a composite filter of SG filter and median filter, and can be represented as

$$x_d(t_j) = k \cdot \sum_{i=-M}^{M} C_i \cdot [h_m(t_{j+i}) * x_n(t_{j+i})] \qquad (4)$$

Filtered respiration results are shown in Fig.s 5 and 6 [9].

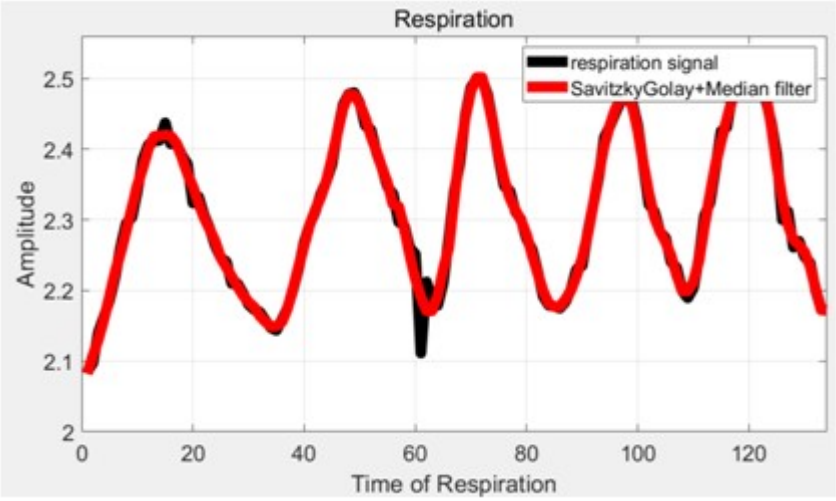


Figure 2. Filtered normal respiration signal

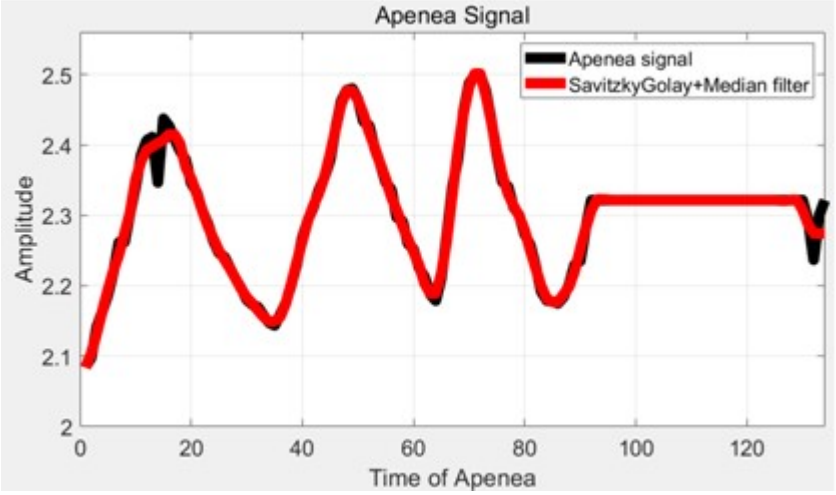


Figure 3. Filtered apnea signal

3. DECISION OF RESPIRATORY STATUS

In this paper, respiratory signal is analyzed in frequency domain. Since the acquired signal is defined in discrete time domain, discrete Fourier Transform (DFT) is used in this work.

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi kn}{N}}, \; k = 0,1,\ldots\ldots,N-1 \qquad (5)$$

Once DFT is calculated, the result is composed of magnitude response and phase response. This paper chiefly deals with magnitude response. In particular, decision of respiratory status is based on the frequency component in which the magnitude response is maximized.

$$\hat{f} = argmax_f |X[k]| \qquad (6)$$

where $\hat{f}$ is a frequency component that corresponds to the maximum magnitude response. Once $\hat{f}$ is estimated, the function $o(\hat{f})$ determines the status of respiration.

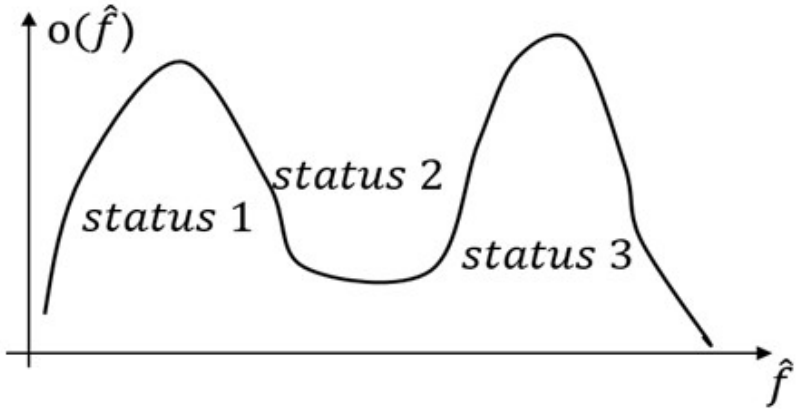


Figure 4. Decision function to estimate respiratory status

4. EXPERIMENTAL RESULTS

This section details the experimental results that substantiate the proposed approach. UWB radar sensor, NOVELLDA X4 is used. Respiration is categorized as normal and abnormal status. Normal signal is categorized as totally normal one and the normal status while speaking. Abnormal signal is an apnea which is partially contained in normal signal. Each person provides respiration signal composed of three status stated above. Each person provides 50 respiration signals each of which can be one of three statuses. Each respiration is acquired during about 60seconds, and ages are between 23 ~ 25. Apnea is contained in the normal signal, and it is generated during about 15seconds. Sampling frequency of the signal is 10Hz. Examples of respiration signals are shown in Fig.s 5 - 7. Results of spectral analysis of respiratory statuses are shown in Fig.s 8 – 10

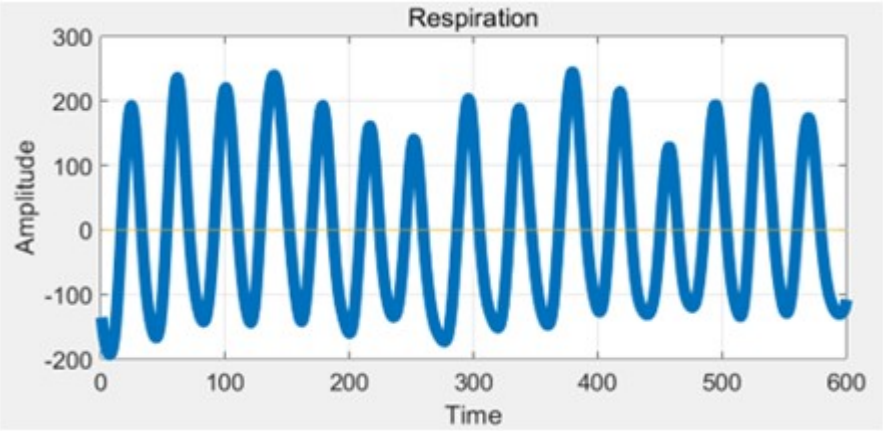


Figure 5. Totally normal respiration

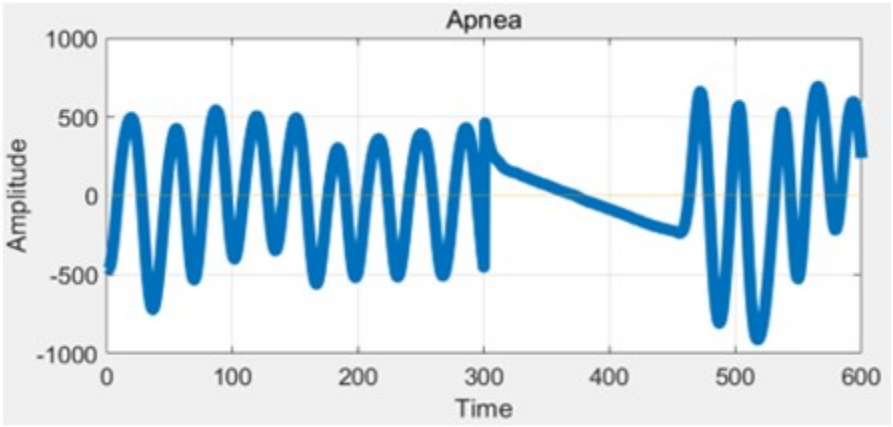


Figure. 6. Partial apnea contained in normal respiration

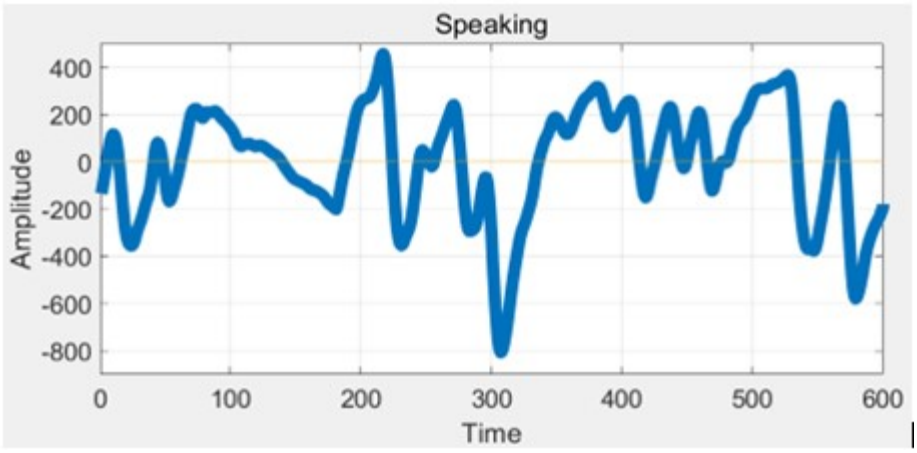


Figure. 7. Normal respiration during speaking activity

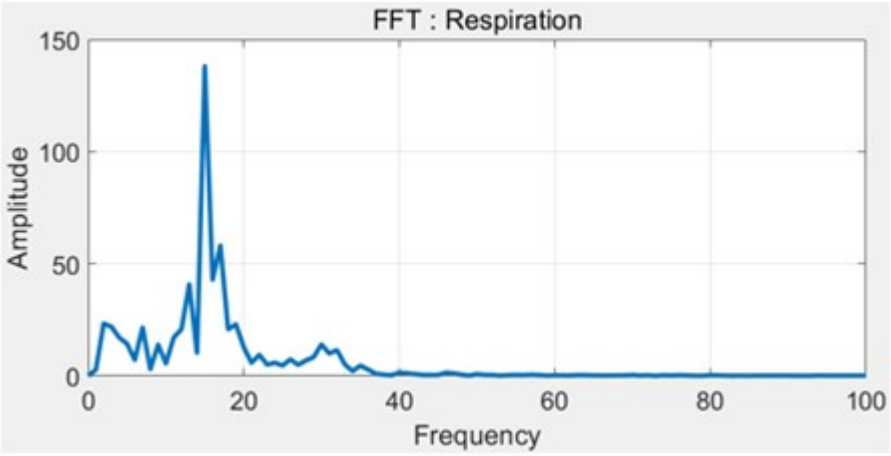


Figure. 8. Magnitude response of DFT of normal respiration

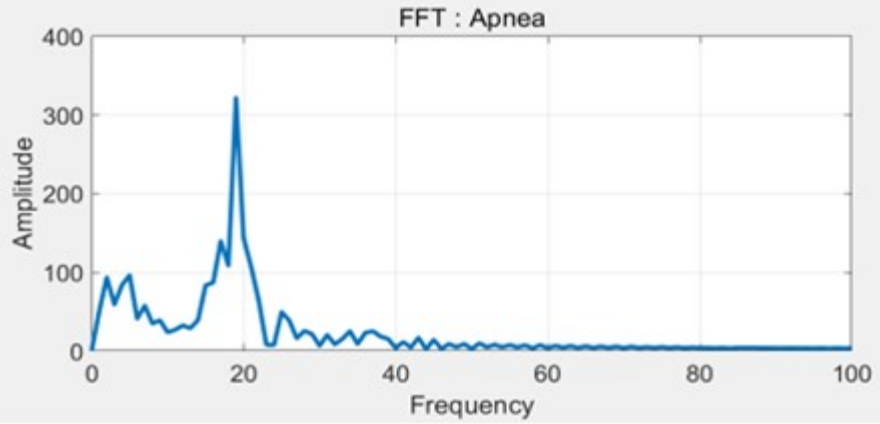


Figure. 9. Magnitude response of DFT of apnea signal

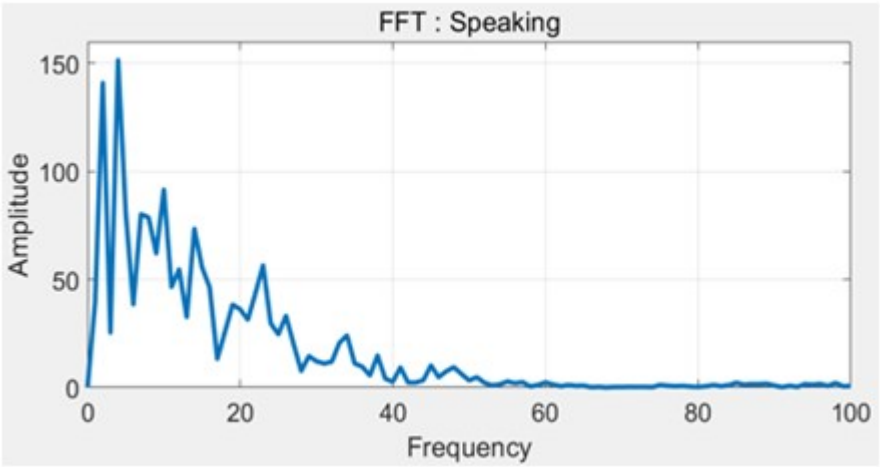


Figure. 10. Magnitude response of DFT of normal respiration with speaking activity

Frequency component that corresponds to the maximum magnitude response is estimated to classify the respiratory status. As shown in Fig.s 8-10, can be used to classify respiratory status, e.g., is in the range of 11 – 16Hz (average is 13.2Hz), 17 – 27Hz (average is 20.9Hz) and 2 – 4Hz (average is 3Hz), in case of normal respiration, apnea and normal respiration during speaking activity, respectively. Intuitively, apnea has a relatively high frequency component due to the discontinuity in the signal, hence, of apnea is higher than the one of the normal respiration. The experiments have been carried out using 250 respiration signals, and the status is picked randomly so that the proposed method can be reliably evaluated in quantitative way.

5. CONCLUSION

In this paper, respiratory status is classified in frequency domain. DFT is simply calculated and the procedures of spectral analysis is straightforward so that the proposed approach can be applied to estimation of vital signal using simple and low-cost device. The experimental results show that the proposed method can reliably classify the respiratory status, and it can be extended

to development of the intelligent sensing system for the respiratory status. In the future, we will try to propos the approach to sensing respiratory status with short-time measurement in non-invasive way so that fast and efficient classification and abnormality detection can be accomplished.

ACKNOWLEDGEMENTS

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding)

REFERENCES

[1] F. Q. AL-Khalidi, R. Saatchi, D. Burke, (2011) H. Elphick and S. Tan, "Respiration rate monitoring methods: A review", Pediatric Pulmonology, 46(6), 523-529.

[2] K. Watanabe, T. Watanabe, H. Watanabe, H. Ando, T. Ishikawa and K. Kobayashi, (2005) "Noninvasive measurement of heartbeat, respiration, snoring and body movements of a subject in bed via a pneumatic method", IEEE Transactions on Biomedical Engineering, 52(12), 2100-2107.

[3] A. D. Droitcour, T. B. Seto, B-K. Park, S. Yamada, A. Vergara, C. Hourani, T. Shing, A. Yuen, V. Lubecke and O. Boric-Lubecke, (2009) "Non-contact respiratory rate measurement validation for hospitalized patients", 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4812-4815. Minneapolis, MN, USA.

[4] M. A. Cretikos, R. Bellomo, K. Hillman, J. Chen, S. Finfer and A. Flabouris, (2008) "Respiratory rate: the neglected vital sign", Medical Journal of Australia, 188(11),657-659.

[5] G. Ossberger, T. Buchegger, E. Schimback, A. Stelzer and R. Weigel, (2004) "Non-invasive respiratory movement detection and monitoring of hidden humans using ultra wideband pulse radar", 2004 International Workshop on Ultra Wideband Systems Joint with Conference on Ultra Wideband Systems and Technologies. Joint UWBST & IWUWBS 2004 (IEEE Cat. No.04EX812), Kyoto, Japan, pp. 395-399.

[6] K. S. Tan, R. Saatchi, H. Elphick and D. Burke, (2010) "Real-time vision based respiration monitoring system", 7th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2010), Newcastle upon Tyne, UK, pp. 770-774.

[7] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita and L. Eklundh, (2004) "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter", Remote Sensing Environment, 91(3-4), 332-344.

[8] S. Kim, C. Hwang, J. Kim, C. Park and D. Lee, (2018) "Application to Detection and Classification of Respiratory Status based on a Signal Correlation", to appear In: Annual Fall Conference of the Korean Institute of Electronics and Information Engineers (IEIE).

[9] S. Kim, C. Hwang, J. Kim and D. Lee, (2019) "Signal Classification based System for Monitoring Abnormal Respiratory Status", submitted to Journal of the Institute of Electronics and Information Engineers.

AUTHORS

Suyeol Kim

Suyeol Kim is in department of biomedical engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on signal processing, particularly on respiration signal analysis. He is currently pursuing his bachelor degree in engineering.

Chaehwan Hwang

Chaehwan Hwang is in department of biomedical engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on signal processing, particularly on respiration signal analysis in frequency domain. He is currently pursuing his bachelor degree in engineering.

Jisu Kim

Jisu Kim is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on image processing, computer vision, signal processing and machine learning. He is currently pursuing his M.S degree in computer engineering.

Cheolhyeong Park

Cheolhyeong Park is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on geometric image analysis, computer vision, computer graphics and machine learning. He is in the course of integrated B.S and M.S degree in computer engineering.

Deokwoo Lee

Dr. Deokwoo Lee is an Assistant Professor in the department of computer engineering at Keimyung University. Dr. Lee has received B.S degree in electrical engineering from Kyungpook National University, Daegu, Republic of Korea, and M.S and Ph.D degree from North Carolina State University, Raleigh, NC, USA, respectively. He has been working on the areas of computer vision, image processing, signal processing and machine learning. In particular, he has been conducting camera calibration, bio-signal analysis and image denoising

INTENTIONAL BLANK

A SURVEY OF VISIBLE IRIS RECOGNITION

Yali Song[1], Yongzhong He[1,2] and Jin Zhang[1]

[1]Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, China, Beijing
[2]Science and Technology on Electronic Information Control Laboratory, Chengdu, China

ABSTRACT

In recent years, research on iris recognition in near-infrared has made great progress and achievements. However in many devices, such as most of the mobile phones, there is no near-infrared device embedded. In order to use iris recognition in these devices, iris recognition in visible light is needed, but there are many problems to use visible iris recognition, including low recognition rate, poor robustness and so on. In this paper, we first clarified the challenges in visible iris recognition. We evaluate the effectiveness of three traditional iris recognition on iris collected from smart phones in visible light. The results show that traditional methods achieve accuracy not exceeding 60% at best. Then we summarize the recent advances in visible iris recognition in three aspects: iris image acquisition, iris preprocessing and iris feature extraction methods. In the end, we list future research directions in visible iris recognition.

KEYWORDS

visible iris recognition, mobile phones, iris image acquisition, feature extraction

1. INTRODUCTION

In recent years, smart phones have been widely used in various fields due to their portability and light weight. From the initial simple call, texting development to later receiving verification code, notepad, payment management and other applications, the hidden danger of user security caused by the excessive amount of data has become an important problem. Biometric recognition provides a feasible solution for user identity security authentication due to its stability and not easy to lose.

In all biometrics characteristics including fingerprint, palm print, gesture, iris and face, iris recognition is deemed to be the most trustworthy biometric identification technology because of its high stability, high recognition rate, and not easy to counterfeit., it has been widely used in finance, medical, security and other fields. With the gradual maturity of iris recognition technology, user acceptance has gradually improved, and the market space is huge.

Iris recognition based on smart phones embeds traditional iris recognition technology into smart phone devices, which can be divided into two categories, namely near-infrared iris image recognition and visible light iris image recognition. The former has a high cost due to needing to embed a near-infrared acquisition device in a mobile phone. However, because of its high recognition rate, it has been released on the market. For example, in May 2015, Fujitsu and

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 11–23, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90302

Japanese operators jointly launched the first new Smartphone with iris recognition, which can unlock the phone through blinking; in August 2016, Samsung released the flagship machine Note7, equipped with iris recognition function, and later rushed out of the market due to the explosion; in March 2017, Samsung continued to release the new flagship S8/S8+ with iris recognition, which quickly unlocks the phone through the iris. The latter has obvious cost advantages because it does not require embedding near-infrared devices in mobile phones., but the recognition rate is low due to the lack of rich texture features of the visible iris image acquired by the ordinary camera.

This paper investigates and evaluates visible iris recognition, and its contributions are as follows:

1. Discussing the purpose and significance of visible iris recognition.
2. Clarifying the current challenges of visible iris recognition
3. Summarizing the recent advances in visible iris recognition
4. Discussing the future research directions in visible iris recognition

The rest of this paper is organized as follows. Section 2 clarified the challenges in visible iris recognition. Section 3 introduces some current iris databases briefly. In Section 4, we summarize some iris preprocessing methods. Section 5 describes several iris feature extraction techniques. Finally, section 6 briefly discusses the conclusions and future work.

2. PRELIMINARY OF VISIBLE IRIS RECOGNITION

This section will briefly describe the basis of iris recognition, mainly from the overall structure and challenges

2.1. THE OVERALL ARCHITECTURE OF VISIBLE IRIS RECOGNITION

Iris recognition is to determine people's identity by comparing the similarities between iris image features. The process of iris recognition technology generally involves the following four steps:

1. Iris image acquisition: The entire eye of a person is photographed using a specific camera device, and the captured image is transmitted to an image preprocessing software of the iris recognition system.
2. Image preprocessing: The acquired iris image is processed to meet the requirements of extracting iris features. It consists of three parts: iris positioning, iris image normalization and image enhancement.
3. Feature extraction: The feature points required for iris recognition are extracted from the iris image by a specific algorithm and then encoded.
4. The feature encoding of the iris texture is compared with the data in the database to achieve the purpose of identity recognition.

2.2. THE CHALLENGES OF VISIBLE IRIS RECOGNITION

The iris can be divided into a light iris and a dark iris according to the color. Since the texture of the dark iris is difficult to be resolved under visible light, the traditional iris recognition is studied for the iris under the near infrared. However, visible iris recognition has gradually attracted

people's interest due to the variety of imaging sensors and the advancement of recognition algorithms. Next, we will explain the challenges of visible iris recognition from two aspects: data acquisition and recognition algorithms.

2.2.1. Data Acquisition

For data collection, we use smartphone as a stand-alone collection device. Using the rear and front cameras of the mobile phone to collect iris images in indoor and outdoor environments. Figure 1 shows a small sample iris. From this we can see that the iris collected by the smartphone has problems such as low image resolution, more noise factors and not obvious texture features.

Figure 1. Sample images of iris

2.2.2. Recognition Algorithms

Due to the low quality of the visible iris collected by the smartphone, its iris texture information is less and difficult to extract. We use three traditional methods to evaluate the performance of visible iris captured by smartphones. According to the experiment, Table 1 lists the performance evaluation of the three methods. Figures 2 and Figures 3 depict the ROC curves for the three methods under outdoor conditions and indoor conditions. The results show that the recognition accuracy of the three traditional methods is the highest 60%, which indicates that the traditional visible iris recognition method cannot extract the iris features well, and further research is needed.

Table 1. Performance evaluation

Method	RR	AUC	EER
K-median clustering	0.60	0.79	0.32
PCA	0.56	0.76	0.40
Log-Gabor transformation	0.43	0.72	0.64

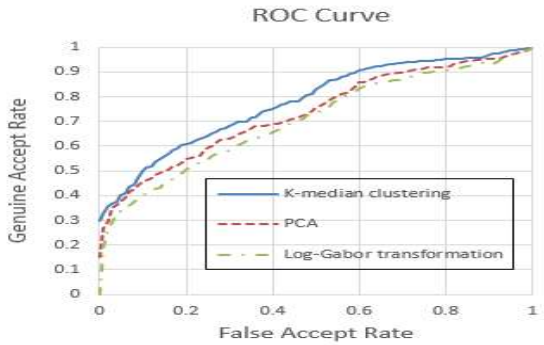


Figure 2. ROC curve in outdoor environment

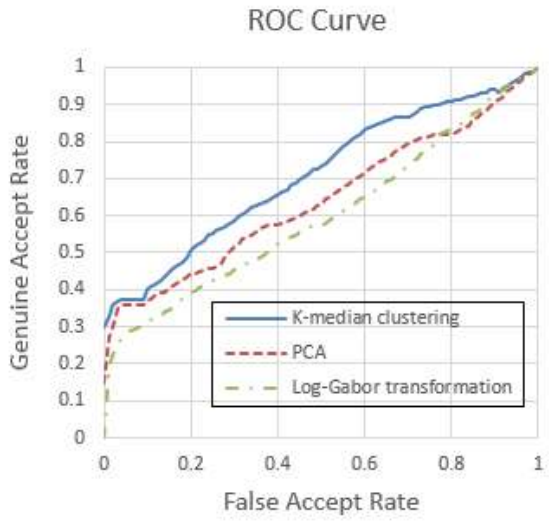


Figure 3. ROC curve in indoor environment

In view of some of the challenges listed above, we summarize the latest developments in visible iris recognition from three aspects: iris image acquisition, iris preprocessing and iris feature extraction. The specific content is shown in the following chapters.

3. VISIBLE IRIS DATABASES

Iris recognition performance is generally determined by the iris image quality and iris recognition methods. Thus, it is quite important to collect iris with high resolution. In this section, we will introduce some publicly available visible iris databases and some other visible iris databases collected by smartphones or light-field cameras.

3.1. OPEN VISIBLE IRIS DATA SETS

There are some publicly available visible iris databases for iris images under visible light. UBIRIS.V2 [22] iris database is an iris image that is acquired over long distances in an indoor environment of natural light sources and artificial light sources. The subject is between 3-10 meters away from the collection device. The database includes 522 irises of 261 people, and a total of 11102 iris images are collected, which can be used for long-distance iris recognition

under visible light. BIP Lab [4] also provides a visible iris dataset of 75 unique irises collected separately from indoor and outdoor using the front and rear cameras of the iPhone 5 and Samsung Galaxy S4 under two different illuminations. In addition, MICHE I and MICHE II also provide visible iris data sets for 4 samples under 8 different conditions under uncontrolled and different illumination using different mobile devices [5], Among them, MICHE I has 1600 pictures from 50 volunteers, while MICHE II has 3120 pictures from 75 volunteers.

3.2. OTHER VISIBLE IRIS DATA SETS

In addition to the above public datasets, some of the literature used their own iris datasets for experiments and performance comparisons with published datasets. Trokileewicz [1] used a mobile device to capture an image of 70 people's irises. The data was acquired by the iPhone 5S's rear camera and the flash was turned on indoors to ensure sufficient light when capturing images. The image quality exceeded the iris image of the near-infrared illumination, and finally 3192 images of different irises were collected. Kiran et al. [2] used two different smartphones, the iPhone 5S and Nokia Lumia, to collect iris images of different colors of volunteers from most Nordic countries under mixed illumination as data sets. The data was collected under the condition of semi-cooperative and unconstrained conditions, and 5 images of 28 people were collected indoors and outdoors, respectively, and a total of 560 iris images were collected. In addition, there is a patent that describes how iris collection should be performed and the issues that need to be noted during the acquisition process. In [3], it is mentioned that the light source and the eye's line of sight are usually adjusted to an angle of at least 30 degrees to direct the visible light onto the iris surface for iris feature collection. Table 2 summarizes some of the current visible iris data sets.

Table 2. Iris data sets

Database	Subjects	Collection environment	Images
UBIRIS .v1	241	1)camera : Nikon E5700 2)focal length :8.9-71.2mm 3)exposure time:1/30s 4)ISO:200	1877
UBIRIS .v2	261	1)camera : Canon EOS 2)focal length :400mm 3)exposure time:1/200 s 4)ISO:1600	11102
UPOL	62		384
UBIPr	344	1)different postures 2)different gaze 3)different lighting	10252
MICHE I	50	1)phone: iPhone5, Samsung Galaxy S4 2)camera: front, rear 3)environment: indoor, outdoor	1600
MICHE II	75	1)phone: iPhone5, Nokia1020 2)camera: front, rear 3)environment: indoor, outdoor	3120
VSSIRIS	28	1)phone: iPhone5, Samsung Galaxy S4 2)environment: Semi-cooperation, unconstrained	560

BDCP	99	camera :LG4000,CFAIRS	1737
VISOB	550	1)phone: iPhone5, Oppo N1, Samsung Note 4, 2) resolution: iPhone 720px, Samsung and Oppo 1080px 3) environment: conventional office, dim office, natural light office	

3.3. EVALUATION OF IRIS DATA SETS

Based on the iris data sets listed above, we evaluated them as follows:

1. The current visible iris is mainly collected by cameras and smart phones. Different acquisition devices have different iris quality due to different camera configurations. For example, UBIRIS database and MICHE, VSSIRIS and VISOB databases have different iris resolutions. The number of irises collected by the same device is also different. The number of irises in the iris database MICHE and VSSIRIS collected for smartphones is quite different, and the scale is generally small.

2. Most mobile phone models currently collecting irises are iPhone or Samsung, and there is no attempt to collect them with other models and other resolutions.

3. Most of the iris database classifications are classified according to the collection equipment, and there is no classification based on other factors such as race, iris color and so on.

4. The current visible iris database can be used for long-distance iris recognition and low-quality iris recognition, in addition to visible iris recognition.

4. IRIS PREPROCESSING

After obtaining the iris image, the iris needs to be pretreated. Iris positioning is a key step in iris recognition, and its accuracy, robustness and positioning speed are extremely important. The purpose of iris localization is to determine the position of the inner and outer boundaries of the iris and the upper and lower eyelid boundaries in the image. In this section, we will discuss the iris localization scheme in iris recognition. The specific algorithms are as follows.

4.1. INTEGRAL DIFFERENTIAL OPERATOR METHOD

Daugman [6-9] proposed using the integral differential operator method to globally search the center and radius to achieve iris localization. It is a geometric feature similar to the circle of the iris and the pupil, which can be achieved by detecting the circle. Its integral differential operator is defined as follows:

$$\max_{(r,x_p,y_o)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)ds}{2\pi r} \right| \tag{1}$$

Where I(x, y) is the acquired image, is the Gaussian function, and r is the search radius. It uses the first derivative to search and find the appropriate three parameters, that is, by changing the center and radius to search for the circle with the largest change of pixel value, iterating sequentially, and gradually reducing the amount of smoothing to achieve positioning. However, if the captured image contains more noise factors, the performance of the algorithm is greatly reduced.

4.2. HOUGH TRANSFORM

In 1996, Wildes [10] proposed the use of edge detection algorithm to detect the edge points of the image, and the hough transform was used to fit the position of the iris contour, which greatly improved its accuracy. Firstly, the edge map is generated by Canny edge detection. This method is implemented based on the edge detection of the gradient. And the set of inner and outer edge points of the iris is obtained by the threshold. Second, the hough transform is used to vote on all possible parameter sets (center and radius), and the parameter set corresponding to the largest vote is the detected circle. Its voting function is as follows:

$$H(x_c, y_c, r) = \sum_{j=1}^{n} h_1(x_j, y_j, x_c, y_c, r) \tag{2}$$

$$h_1(x_j, y_j, x_c, y_c, r) = \begin{cases} 1, \ if \ g(x_j, y_j, x_c, y_c, r) = 0 \\ 0, \ if \ g(x_j, y_j, x_c, y_c, r) \neq 0 \end{cases} \tag{3}$$

$$g(x_j, y_j, x_c, y_c, r) = (x_j - x_c)^2 + (y_j - y_c)^2 - r^2 \tag{4}$$

Where (x_j, y_j) is all edge points, $h_1(x_j, y_j, x_C, y_C, r)$ is a function of whether or not to vote, $g(x_j, y_j, x_C, y_C, r)$ is a discriminant function of the circle. The parameter set that finally obtains the maximum value of $H(x_C, y_C, r)$ is the parameter sought.

4.3. OTHER METHODS

Jan [11] proposed an efficient non-circular iris contouring scheme that uses an integral differential operator, image grayscale intensity, pupil/iris geometry and adaptive threshold mixing to pinpoint the iris contour. Proença [12] proposed the use of color component analysis to determine the iris boundary by performing color channel analysis within the iris region of the image. Wang et al. [17] proposed using a repair method based on the Navier-Stokes (NS) equation to fill the reflection points, and using the possible boundary (Pb) edge detection operator to initially detect the pupil edge to provide accurate positioning for low-quality iris images. Baek et al. [19] proposed an iris center localization algorithm based on the eyeball model to facilitate accurate positioning of the iris center under different head postures. Table 3 lists some of the positioning segmentation algorithms in the literature.

Table 3. Iris localization and segmentation algorithm

Reference	Method	Data sets	Accuracy
Yingzi et al. [23]	ellipse fitting based on DLS	IUPUI	EER: 1.79%
Daugman[6]	integral differential operator		EER: 0.08%
Wildes [10]	hough transform		EER: 1.76%
Jan [11]	integral differential operator, image gray intensity, pupil/iris geometry adaptive threshold	UBIRIS V1.0	RR: 93.50%
		MMU V1.0	RR: 99.25%
		IITD V1.0	RR: 99.46%
Proença [12]	color component analysis	UBIRIS.v2	EER: 5.02%
Wang et al. [17]	Pb edge detection operator	CASIA-Iris-Thousand	EER: 1.82%
Baek et al.[19]	iris center positioning based on eyeball model	Gi4e	RR: 81.4%
		HPEG	RR:88.6%
Chen et al. [24]	merge active contour model	THU Iris V1.0	EER : 0.43
Sahmoud et al.[25]	K-means clustering	UBIRIS V1.0	RR: 98.76%

4.4. Evaluation Of Iris Positioning

For the above positioning method, the following evaluation was made:

1) They are mostly based on the circular geometry of the iris or combined with iris color analysis. As with the method of using only the iris shape or only the color analysis, the method of combining the two has a significantly better positioning effect. For example, the method in the Jan literature is better positioned than Proença.
2) The iris positioning effect is different on different datasets due to the different iris quality. Relatively speaking, the near-infrared iris dataset is better than the other datasets because of its better iris quality. Such as the CASIA-Iris-Thousand near-infrared data set and the UBIRIS.v2 visible data set.
3) Due to the long calculation time of the iris localization algorithm, the speed of positioning using the above method is not increased by an order of magnitude. Therefore, some faster methods can be studied for iris positioning, and positioning speed can be improved based on ensuring accuracy.
4) The above positioning method can be used for both near-infrared iris recognition and visible iris recognition. However, the positioning effect of the two must be different depending on the quality of the iris.

5. Iris Feature Extraction

Feature extraction refers to extracting unique feature points from the separated iris images by a certain algorithm and encoding them. A lot of literature has been introduced on the method of iris feature extraction. The specific method is as follows.

5.1. ORDINAL FEATURES AND CNN MODELS

Zhang et al. [14] proposed to measure the local iris texture by extracting the optimized ordinal measurement features, and then use the convolutional neural network to automatically learn the pairwise features to measure the correlation between the two irises. First, the ordinal feature selection is performed, and the normalized iris picture is divided into different areas using two-leaf and three-leaf ordinal filters, and different ordinal filters are applied to different regions. Select 15 regional features that can express texture information, reduce feature dimension and reduce processing time. After obtaining the OMs characteristics of the iris, the Hamming distance of the two feature templates is calculated. Secondly, using the paired CNN model to measure the correlation between the two irises, input two iris images of size 128 × 128, the intra-class pairs and the inter-class pairs are labeled as 1,0. The 64 pairs of filters in the first layer of convolutional layer are zero-filled the second and third layers are processed equally, and the dropout rate is set to 0.5, using the dropout technique to prevent overfitting. The experiment finally achieved a better recognition effect of 0.8% EER.

5.2. GEOMETRIC KEY

Tan et al. [18] proposed feature extraction of long-distance iris images using geometric-key-based iris coding, and experiments in three published iris databases, compared with several other feature extraction methods, and finally proved that the proposed method has improved the recognition performance. First, the Log-Gabor transform is used to encode the global iris feature, which has the iris intensity of the iris region with less matching noise. Second, geometric information is used to provide efficient coding from local iris area pixels. Geometric keys (a set of coordinate pairs) are randomly generated and assigned to each topic, which uniquely defines the way the iris is coded, and the local iris feature is encoded based on the binary at the geometric key position. Then, configuring the geometry keys to resolve the proportional and rotational variations of the local iris area, and combining global iris coding with local iris coding to accommodate higher intra-injection imaging transformations. Finally, an effective match is made using the Hamming distance. The method uses three different databases to conduct experiments, and the error rates were 36.3%, 32.7% and 29.6%.

5.3. LOG-GABOR TRANSFORMATION

Texture features are extracted from the radial and angular directions using a two-dimensional Log-Gabor filter. The Gabor function is as follows:

$$g(x,y)=\frac{1}{2\pi\delta_x\delta_y}exp[-(\frac{x^2}{\delta_x{}^2}+\frac{x^2}{\delta_y{}^2})]\times cos[2\pi f(xcos\theta+ycos\theta)] \quad (5)$$

Where f is the center frequency of the filter, θ is the filter direction, and (δ_x,δ_y) is the Gaussian function standard deviation.

To achieve coverage of the filter in multiple directions, multiple filters are needed to show the discrepancy between different texture features. Its feature extraction formula is as follows:

$$F_{kj}(x,y) = H_{kj} * I(x,y) \tag{6}$$

Where * is a convolution operation, I (x, y) is the processed iris, k is the scale, and j is the direction.

5.4. OTHER TECHNOLOGIES

A lot of literature has been introduced on the method of iris feature extraction. Tan et al. [13] proposed using the phase information of the Zernike moment to encode the iris image and fuse it with the Gabor filtered result, and weight the code according to the information of the vulnerable bits. Kiran et al. [15] proposed a method for iris recognition by extracting iris features by multi-segment depth sparse histograms combined with color channels. Raja et al. [2] proposed a new feature extraction method based on depth sparse filtering to obtain robust iris features. In [16], the strategy of iris and periocular information fusion is proposed to extract their respective features and fuse them, which can achieve cross-sensor iris recognition. Table 4 lists the feature extraction algorithms and their recognition effects in some literatures.

Table 4. Iris feature extraction algorithm

Reference	Method	Data sets	Result
Tan et al. [18]	geometric key	UBIRIS V2.0	RR: 36.3%
		FRGC	RR:32.7%
		CASIA.v4-distance	RR:29.6%
Kiran. et al.[15]	multi-segment depth sparse histogram	MICHE-I	EER: 0.37%
		MICHE-II	GMR: 95%
Raja et al. [2]	depth sparse filtering	VSSIRIS	EER:1.62%
Tan et al.[13]	zernike moment phase characteristics	UBIRIS V2.0	RR:54.3%
		FRGC	RR: 32.7%
		CASIA.v4-distance	RR: 42.6%
Zhang et al.[14]	ordinal feature and CNN model		EER:1.2%
Radu et al, [26]	2D Gabor wavelets	UBIRIS.v1	EER:4.5%
Tsai [27]	LBP	UBIRIS V2.0	ACC:92.5%
Raja et al.[28]	BSIF		EER:1.05%

5.5. EVALUATION OF IRIS FEATURE EXTRACTION

For the above feature extraction method, the following evaluation was made:

1) The above feature extraction method can be used for both near-infrared iris recognition and visible iris recognition. However, due to the difference in the amount of iris texture information under near-infrared and visible light, the recognition effect under visible light is not as good as that under near-infrared (except for the iris acquired at a long distance).
2) The recognition effect of iris on different data sets is very different due to the difference in iris quality. For example, the recognition effect under UBIRIS and MICHE data sets is better than other data sets.
3) At present, the average recognition rate of visible iris recognition is generally not high. Compared with the traditional recognition method, the use of depth sparse filtering or

ordinal features and CNN model recognition accuracy is relatively good. Therefore, some more advanced feature extraction methods can be studied in order to better extract iris features.

6. DISCUSSION AND FUTURE RESEARCH DIRECTIONS

In recent years, due to the advantages of visible iris recognition, such as the use of convenient mobile phones, long-distance recognition, and iris obvious color discrimination, visible iris recognition based on smartphones is considered an important research topic. However, there are still some challenges for visible iris recognition.

1) The lower resolution of the iris image captured by the smartphone makes the texture information difficult to obtain.
2) Dark irises in Asians' visible light cause less texture information due to more melanin
3) Due to the low quality of the iris image, the accuracy of iris recognition using some traditional methods is low, and some more advanced methods must be studied to make it better to extract iris features.

In this paper, we investigated some databases on the identification of visible iris recognition, preprocessing methods and feature extraction methods. According to the survey, most of the visible iris recognition uses some open iris databases for experiments, but some of the literature use their own internal data sets for performance evaluation. The data set is small and not convincing. Secondly, most of the iris localization methods of the investigated literature are based on the circular geometry of the iris or combined with iris color analysis, and the average positioning effect is over 80%. For the feature extraction method, most of the literatures have low recognition rate and high error rate, indicating that the current method still needs further improvement

In order to better promote the further development of visible iris recognition, in the future, our work can be carried out in the following aspects:

1) Large-scale visible iris database: The current visible iris datasets are few and the classification is not clear. In the future, irises can be collected to form new iris databases for different races and colors in a semi-cooperative environment.
2) Incorporating eye information: Since there is less information on the iris texture of the Chinese under visible light, some eye contour information can be added in the future to obtain more texture features.
3) Improvement of recognition method: Since the traditional recognition method cannot extract the characteristics of visible iris, some deep learning methods such as convolutional neural network can be used in the future.

ACKNOWLEDGMENT

This work was supported in part by the following grants: National key R&D program of China No. 2017YFC0820100

REFERENCES

[1] M. Trokielewicz, "Iris recognition with a database of iris images obtained in visible light using smartphone camera", IEEE International Conference on Identity Security and Behavior Analysis, pp. 1-6, 2016.

[2] KiranB. Raja, R. Raghavendra, V.K. Vemuri, C. Busch. Smartphone based visible iris recognition using deep sparse filtering. Pattern Recognit. Lett., 57 (2015), pp. 33-42

[3] Keith J. Hanna. Systems and methods for illuminating an iris with visible light for biometric acquisition. US9124798B2. 2011-05-17.

[4] http：//biplab.unisa.it/MICHE/database/

[5] M. De Marsico, M. Nappi, D. Riccio, H. Wechsler, "Mobile Iris Challenge Evaluation (MICHE)-I biometric iris dataset and protocols", Pattern Recognition Letters, vol. 57, pp. 17-23, 2015.

[6] J. Daugman. How iris recognition works. Proceedings of 2002 International Conference on Image Processing, Vol. 1, 2002

[7] J. Daugman. Biometric personal identification system based on iris analysis. United States Patent, Patent Number: 5,291,560, 1994.

[8] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 11, 1993.

[9] J. Daugman. Biometric decision landscapes. Technical Report No. TR482, University of Cambridge Computer Laboratory, 2000.

[10] R. Wildes. Iris recognition: an emerging biometric technology. Proceedings of the IEEE, Vol. 85, No. 9, 1997.

[11] Jan, F. (2016). Non-circular iris contours localization in the visible wavelength eye images. Computers and Electrical Engineering Journal, 62, 1–12.

[12] H. Proença. Iris recognition: on the segmentation of degraded images acquired in the visible wavelength. IEEE Trans. Pattern Anal. Mach. Intell., 32 (8) (2010), pp. 1502-1516

[13] Tan C W,Kumar A. Accurate iris recognition at a distance using stabilized iris encoding and Zernike moments phase features[J]. Image Processing, IEEE Transactions on, 2014, 23(9): 3962-3974.

[14] Zhang, Q., Li, H., Sun, Z., He, Z., Tan, T.: Exploring complementary features for iris recognition on mobile devices. In: International Conference on Biometrics (2016, to appear)

[15] K. Raja, R. Raghavendra, S. Venkatesh, C. Busch, Multi-patch deep sparse features for iris recognition in visible spectrum using collaborative subspace for robust verification, Pattern Recognit. Lett. Same Volume TBA.

[16] Xiao, L., Sun, Z., Tan, T.: Fusion of iris and periocular biometrics for cross-sensor identification. In: Zheng, W.-S., Sun, Z., Wang, Y., Chen, X., Yuen, P.C., Lai, J. (eds.) CCBR 2012. LNCS, vol. 7701, pp. 202–209.

[17] N. Wang, Q. Li, A.A.A. El-Latif, T. Zhang, X. Niu. Toward accurate localization and high recognition performance for noisy iris images. Multimed. Tools Appl. (2012), pp. 1-20

[18] C.-W. Tan, A. Kumar. Efficient and Accurate At-a-Distance Iris Recognition Using Geometric Key-Based Iris Encoding. IEEE Transactions on Information Forensics and Security, 9 (9) (2014), pp. 1518-1526

[19] Baek, S.J., Choi, K.A., Ma, C., et al: 'Eyeball model-based iris center localization for visible image-based eye-gaze tracking systems', IEEE Trans. Consum. Electron., 2013, 59, (2), pp. 415–421.

[20] Radman A, Zainal N, Ismail M. Efficient iris segmentation based on eyelid detection[J]. Journal of Engineering Science and Technology, 2013, 8(4): 399-405.

[21] Radu P, Sirlantzis K, Howells G, et al. A Versatile Iris Segmentation Algorithm[C]//BIOSIG.2011: 137-150.

[22] H. Proenca, S. Filipe, R. Santos, J. Oliveira, L. A. Alexandre, "The UBIRIS.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance", IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 8, pp. 1529-1535, Aug. 2009.

[23] Y. Du, E. Arslanturk, Z. Zhou, C. Belcher.Video-based noncooperative iris image segmentation. IEEE Trans. Syst. Man Cybern. B, 41 (2011), pp. 64-74

[24] R. Chen, X.R. Lin, T.H. Ding. Iris segmentation for non-cooperative recognition systems. IET Image Process., 5 (2011), pp. 448-456.

[25] S.A. Sahmoud, I.S. Abuhaiba. Efficient iris segmentation method in unconstrained environments. Pattern Recogn., 46 (2013), pp. 3174-3185.

[26] P. Radu, K. Sirlantzis, W. Howells, F. Deravi, S. Hoque. Information fusion for unconstrained iris recognition. Int. J. Hybrid Inf. Technol., 4 (4) (2011), pp. 1-12

[27] Y. Tsai. A weighted approach to unconstrained iris recognition. Intl J. Comput. Inf. Syst. Control. Eng. (1) (2014), pp. 30-33

[28] K.B. Raja, R. Raghavendra, C. Busch. Binarized statistical features for improved iris and periocular recognition in visible spectrum. International Workshop on Biometrics and Forensics (IWBF) (2014), pp. 1-6

INTENTIONAL BLANK

OCCLUSION HANDLED BLOCK-BASED STEREO MATCHING WITH IMAGE SEGMENTATION

Jisu Kim, Cheolhyeong Park, Ju O Kim and Deokwoo Lee

Department of Computer Engineering, Keimyung University, Daegu 42601, Republic of Korea

ABSTRACT

This paper chiefly deals with techniques of stereo vision, particularly focuses on the procedure of stereo matching. In addition, the proposed approach deals with detection of the regions of occlusion. Prior to carrying out stereo matching, image segmentation is conducted in order to achieve precise matching results. In practice, in stereo vision, matching algorithm sometimes suffers from insufficient accuracy if occlusion is inherent with the scene of interest. Searching the matching regions is conducted based on cross correlation and based on finding a region of the minimum mean square error of the difference between the areas of interest defined in matching window. Middlebury dataset is used for experiments, comparison with the existed results, and the proposed algorithm shows better performance than the existed matching algorithms. To evaluate the proposed algorithm, we compare the result of disparity to the existed ones.

KEYWORDS

Occlusion, Stereo vision, Segmentation, Matching.

1. INTRODUCTION

One can understand, recognize and analyze real world three dimensional scenes using parallax generated from both eyes [1]. Computer vision, especially stereo vision, utilizes the parallax principle in order to make a machine achieve understanding real world scenes. The technique, called computer vision, has brought tremendous attention in diverse areas such as recognition, machine learning, deep learning, etc. Stereo matching is one of the procedures for depth estimation from multiple-view camera systems (usually stereo camera system). To achieve accurate depth values, disparity should be calculated with high accuracy. Accurate stereo matching can be achieved if one can provide sufficiently accurate rectification result that is from camera calibration. Disparity is a difference of horizontal pixel locations (vertical location is identical if rectification is successfully done) between a pair of images each of which is projection of 3D real world scene of target [2]. The results of stereo matching are reliable if the vertical pixel locations are identical between a pair of images. Unfortunately, in practice, it is difficult to find the perfect algorithm that can be applied to all of the scenes due to the existence of ambiguity inherent with images, e.g., discontinuities in boundaries or depth, occlusion or texture-less regions. Therefore, numerous methods for stereo matching have been proposed so that various cases stated above can be solved [3,4]. In general, the most popularly employed techniques for the stereo matching can be categorized into two methods, the one is global matching and the other one is local one [5]. Global matching method usually deals with the cases of existence of texture-less region or occlusions. Global method provides accurate disparities

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 25–31, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90303

whereas it has a limitation of higher computational complexity compared to the local one. Graph-cut, belief propagation or dynamic programming based stereo matching are the widely used techniques of the global matching method. Graph-cut theoretic method borrows the concept of computer network, and disparity values are assigned based on using max-flow / min-cut algorithm that has shown superior performances of disparity estimation [6]. Belief propagation also has shown the competitive result of matching accuracy, however, it consumes high computation time to achieve convergence of the optimal solution [7]. To alleviate the limitation of computational complexity, dynamic programming(DP) based stereo matching and disparity estimation has been proposed. DP based disparity estimation finds the optimal path for disparity that is defined in disparity space image (DSI) generated by the relationship between a reference and a test image (or a pair of images) [8]. Local method for stereo matching, contrary to the global one, is simple to implement, shows lower computational complexity. However, it is rather vulnerable to have errors in the cases of occlusion, texture-less scenes, and to provide low accuracy of matching in boundary regions. Thus local method is popularly used in the areas that require real-time processing and high speed applications [9]. Local method can be categorized into two techniques, the one is feature based method and the other one is area-based method. The former one fully exploits edge, corner or cave information instead of only using pixel values of an image [10]. SIFT(Scale invariant feature transform) and SURF(Speed-up robust feature) are well known algorithms of the feature based technique [11,12]. As well known, SIFT is invariant to light, noise, view-points, rotation, scale, etc. SURF, based on multi-scale space theory, can alleviate the limitation of computational complexity of SIFT algorithm. SURF carries out feature extraction using Hessian matrix [12]. Block based matching algorithm is one of the local based matching techniques. It searches the matched areas by calculating cross correlation and by searching the region that results in the minimum difference of pixel values between the regions each of which belongs to the reference and the test image, respectively. Block based matching algorithm shows competitiveness in that it is fast and simple to be implemented.

In this paper, we propose the approach to stereo matching based on block based techniques and segmentation of sub-regions of images. The proposed method improves the existed block based matching method. This work maintains the performance of block-based method while increasing the accuracy of matching result by adding segmentation technique. To substantiate the propose approach, experimental results are provided, and the results outperform the existed result of block based matching. The rest of this paper is organized as follows. Section 2 briefly introduces the current block based matching followed by section 3 that introduces the proposed approach in this work. Section 4 provides the experimental results followed by concluding remark in section 5.

2. BLOCK BASED STEREO MATCHING

Block based stereo matching is the area-based matching technique. Prior to searching task, searching window is established followed by cross correlation is calculated. Correlation quantifies the similarity between sub-regions belonging to a pair of images, and the most similar areas result in the least square difference. Determination of size of searching window is also important because the size affects to performance of matching task, e.g., accuracy of matching, error sensitivity and computational complexity.

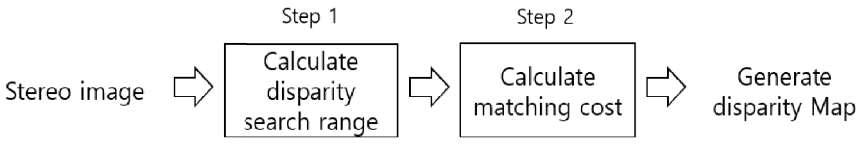


Figure 1. Overall flow of stereo matching and disparity map estimation

Figure 1 shows the overall flow of stereo matching. Matching is carried out within specific range, called search range. In practice, search range also needs to be optimized in order to achieve system efficiency, but this is not discussed in this paper (Figure 2).

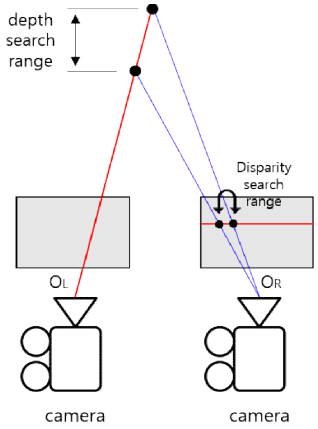


Figure 2. Search range for stereo matching

The similarity is also quantified by calculating mean square error between a pair of sub-regions contained in the matching windows, and the error is called matching cost. Ultimately, stereo matching is carried out so that the algorithm finds the point that generates the minimum matching cost. Numerous method is used to calculate matching cost, and one of the most widely used methods is SAD (sum of absolute difference). The point of the least SAD is decided to be the matching area, and the disparity is defined at this point.

3. Proposed Approach To Stereo Matching

This section is the most contribution of this paper and details the proposed method. Stereo matching is carried out based on the block based techniques. In addition, image segmentation is combined with the block based method in order to increase matching accuracy. This work also deals with the occlusion. Figure 3 shows overall flow diagram of the proposed approach in this paper.

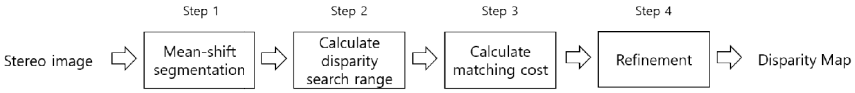


Figure 3. Overall flow of the proposed algorithm for stereo matching

Image segmentation is performed using mean-shift algorithm, and the clustered sub-regions are used for stereo matching. Mean-shift algorithm, one of the clustering methods, is based on the greedy algorithm, and the clustering is performed by searching the area of highest density of the data points [13]. Searching range of mean shift algorithm, called searching radius, needs to be established prior to carrying out clustering. Eq. 1 is mathematical modeling of mean-shift algorithm that updates the mean area (or the area of the highest density).

$$m_h(x) = \frac{\sum_{i=1}^{n} x_i k\left[\left\|\frac{x-x_i}{h}\right\|^2\right]}{\sum_{i=1}^{n} k\left[\left\|\frac{x-x_i}{h}\right\|^2\right]} - x, \quad (1)$$

where $m_h(x)$ is the updated mean value at the location of data x (x is also the center of the searching window). n is a number of the points in searching window, i.e., data points are represented as x_1, x_2, …… x_n. h is a radius of a searching window, and $k(\bullet)$ represents kernel function. According to Eq. 1, $m_h(x)$ is updated to the location of higher data density, and it converges to the highest density point. The result of segmentation using mean shift algorithm is shown in Figure 4

.

Figure 4. Original image (left) and segmentation and clustering result (right)

As shown in Figure 4, mean-shift algorithm provides the clustering result that preserves edge information while normalizing the pixel values in the same cluster. In this paper, to evaluate matching cost (or dissimilarity) C_{SAD} is employed, and C_{SAD} is written as

$$C_{SAD}(x_L, x_R) = \sum_{m=-W}^{+W} \sum_{n=-W}^{+W} |I_R(x_L + m, s + n) - I_R(x_R + m, s + n)| \quad (2)$$

where x_L, x_R are real values, horizontal coordinates of the pixels that are contained in the left and the right matching windows of images, respectively. Size of the matching window is W X W. I_R is a sub-region of an image and s represents the vertical coordinate of a pixel (since a pair of images is already rectified, vertical coordinate of the matched pixels are identical). By using SAD, matching cost is calculated, and the sample result is provided in Figure 5.

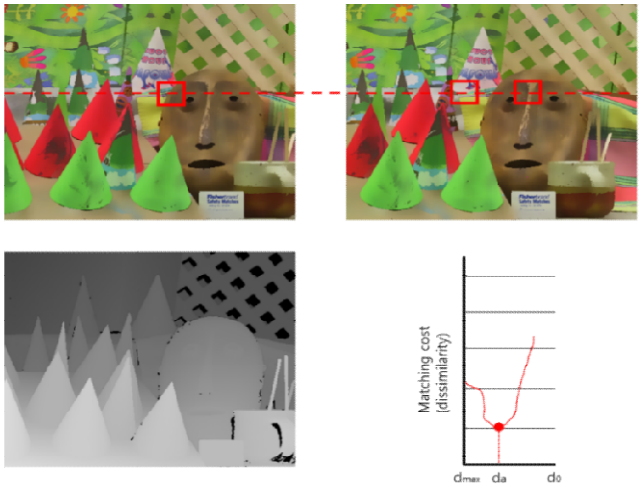


Figure 5. Matching cost calculated on the scanline. The minimum matching cost corresponds to the matched area.

Stereo matching sometimes shows weakness in the case of occlusion. To avoid error due to occlusion, prior to matching, detection of the occlusion area is required. This paper also deals with detecting regions of occlusion. By carrying out occlusion detection, one can provide high quality disparity map. In this work, to detect occlusion, gradient of an image is calculated. The

calculation is performed in a pair of images, especially at the same region of both images. If the edge is detected at the same region (i.e., gradient is identical at the same region), the region is considered no occlusion. If not, in other words, gradient is radically increased in only one image, it is considered the occlusion is existed. If occlusion is existed, disparity value or gradient has high standard variation. Based on the standard variation, the areas of false matching are detected. The area of false matching is filled with the neighboring values and the disparity map is refined. Figure 6 shows detected occluded regions and the refined disparity map.

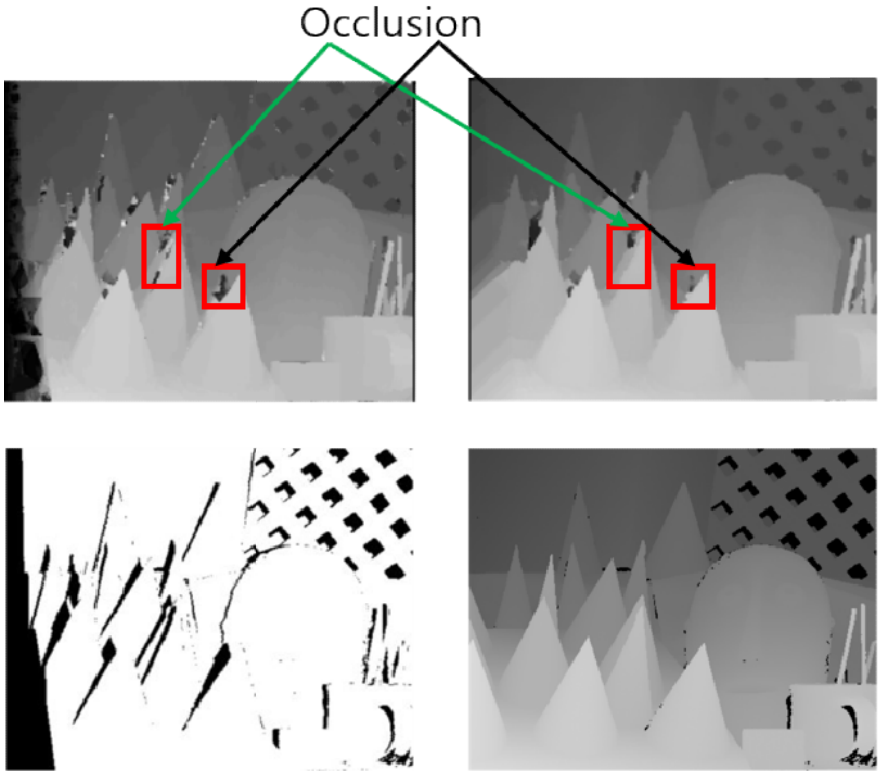


Figure 6. Detection of occlusion based on gradient of image (left-top and right-top), Occluded regions (left-bottom) and refined disparity map(right-bottom)

4. EXPERIMENTS

In the experiments, middlebury dataset is used. Comparison and evaluation are also carried out using the results provided in the website of middlebury stereo vision [14]. To evaluate the proposed approach in this paper, MSE(mean square error) of the disparity map results between the methods and the ground-truth is calculated [15]. Ground truth of disparity value is provided *a priori* in the page of middlebury stereo dataset. The comparison is performed between the proposed method and block-based matching technique. Although only the results from two images are presented, the experiments have been conducted using all of the images in *2003 dataset* contained in middlebury dataset.

As shown in Table 1, the performance is affected by the size of matching window. However, in overall, the proposed method outperforms the conventional block based matching algorithm.

Table 1. Mean square error of the disparity values

Method	Image 1				Image 2			
	3 X 3 (window size)		9 X 9		3 X 3		9 X 9	
	Left	Right	Left	Right	Left	Right	Left	Right
Block based matching	11516	11259	11770	1000	7832	7442	8082	7573
Proposed	11130	10396	11456	10722	7728	7380	8065	7477
Improvement(%)	3.35	7.66	2.66	2.52	1.32	0.83	0.02	1.26

5. CONCLUSION

This paper presents the stereo matching algorithm that combines block-based matching and image segmentation while occlusion is handled. Prior to performing stereo matching, image segmentation is carried out so that the feature is accurately extracted and boundary information is preserved. While preserving strength of the conventional matching algorithm the proposed method could increase the accuracy of matching and the result of disparity estimation. However, there still exists matching errors due to the various environments of the images. In the future work, learning based feature extraction and matching will be used in order to achieve adaptive matching can be performed.

ACKNOWLEDGEMENTS

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding)

REFERENCES

[1] Hartely, Richard. & Zisserman, Andrew (2003) *Multiple View Geometry in Computer Vision*, Computer graphics, image processing and robotics, Cambridge University Press.

[2] Mühlmann, Karsten & Maier, Dennis & Hesser, Jürgen & Männer, Reinhard, (2002) "Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation", *International Journal of Computer Vision*, Vol. 47, No. 1, pp.79-88.

[3] Xu, Jintao & Yang, Qingxiong & Feng, Zuren, (2016) "Occlusion-Aware Stereo Matching", *International Journal of Computer Vision*, Vol. 120, No. 3, pp.256-271.

[4] Kim Kyung Rae & Kim Chang Su, (2016) "Adaptive smoothness constraints for efficient stereo matching using texture and edge information*", 2016 IEEE International Conference on Image Processing (ICIP)*, pp.3429-3433.

[5] Brown, Myron Z & Burschka, Darius & Hager, Gregory D, (2003) "Advances in computational stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 25, No. 8, pp.993-1008.

[6] Huang, Xiaoshui & Yuan, Chun & Zhang Jian, (2015) "Graph Cuts Stereo Matching Based on Patch-Match and Ground Control Points Constraint", Advances in Multimedia Information Processing – PCM, Vol. 9315, pp14-23.

[7] Mozerov, Mikhail G & Weijer, Joost van de, (2015) "Accurate Stereo Matching by Two-Step Energy Minimization", *IEEE Transactions on Image Processing,* Vol. 24*,* No. 3, pp.1153-163.

[8] Salehian, Behzad & Fotouhi, Ali M & Raie, Abolghasem A, (2018) "Dynamic programming-based dense stereo matching improvement using an efficient search space reduction technique", *Optik*, Vol. 160, pp.1-12.

[9] Zhu, Shiping & Yan, Lina, (2017) "Local stereo matching algorithm with efficient matching cost and adaptive guided image filter". *The Visual Computer,* Vol. 33, No. 9, pp. 1087-1102.

[10] Kang, C & Kim, J & Lee, S & Nam, K, (1997) "Stereo Matching Using Dynamic Programming with Region Partition". *Journal of the Institute of Electronics and Information Engineers,* Vol. 20, No. 1, pp.479-482.

[11] Lowe, David G, (1999) "Object recognition from local scale-invariant features", *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pp.1-8.

[12] Bay, Herbert & Tuytelaars, Tinne & Gool, Luc V, (2008) "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp.345-359.

[13] Lee, K-M. and Lin, C-H, (2017) "And Image Segmentation and Merge Hierarchical Region using Mean-Shift Tracking Algorithm", *Proceedings of Annual Conference of IEIE*, pp.704-706.

[14] Scharstein, D & Szeliski, R, (2002) "A taxonomy and evaluation of dense two-frame stereo corresponding algorithms", *International Journal of Computer Vision,* Vol. 47, No. 1, pp.7-42.

[15] Scharstein, D & Szeliski, R, (2003) "High-Accuracy Stereo Depth Maps Using Structured Light", *IIEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.195-202.

Author

Jisu Kim is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on image processing, computer vision, signal processing and machine learning. He is currently pursuing his M.S degree in computer engineering.

Cheolhyeong Park is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on geometric image analysis, computer vision, computer graphics and machine learning. He is in the course of integrated B.S and M.S degree in computer engineering.

Ju O Kim is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on image analysis and Processing. He is pursuing B.S degree in computer engineering.

Dr. Deokwoo Lee is an Assistant Professor in the department of computer engineering at Keimyung University. Dr. Lee has received B.S degree in electrical engineering from Kyungpook National University, Daegu, Republic of Korea, and M.S and Ph.D degree from North Carolina State University, Raleigh, NC, USA, respectively. He has been working on the areas of computer vision, image processing, signal processing and machine learning. In particular, he has been conducting camera calibration, bio-signal analysis and image denoising.

INTENTIONAL BLANK

IMAGE SEGMENTATION BASED ON MULTIPLEX NETWORKS AND SUPER PIXELS

Ivo S. M. de Oliveira[1,2], Oscar A. C. Linares[1], Ary H. M. de Oliveira[3], Glenda M. Botelho[3] and João Batista Neto[1]

[1]Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo,
Caixa Postal 668, 13560-970, São Carlos, Brazil
[2]Instituto Federal do Tocantins, Campus de Paraíso do Tocantins,
Caixa Postal 151, 77600-000, Paraíso do Tocatins, Brazil.
[3]Universidade Federal do Tocantins, Palmas, Brazil

ABSTRACT

Despite the large number of techniques and applications in the field of image segmentation, it is still an open research field. A recent trend in image segmentation is the usage of graph theory. This work proposes an approach which combines community detection in multiplex networks, in which a layer represents a certain image feature, with super pixels. There are approaches for the segmentation of images of good quality that use a single feature or the combination of several features of the image forming a single graph for the detection of communities and the segmentation. However, with the use of multiplex networks it is possible to use more than one image feature without the need for mathematical operations that can lead to the loss of information of the image features during the generation of the graphs. Through the related experiments, presented in this work, it is possible to identify that such method can offer quality and robust segmentations.

KEYWORDS

community detection; complex networks; image segmentation; multiplex networks; super pixels

1. INTRODUCTION

Image segmentation is one of the most important techniques used in Digital Image Processing to extract information from an image. It allows to decompose the image into two or more parts (regions or objects) [1], and it allows to extract information that can be used in subsequent processes of computational vision. Among the various applications we can mention: diseases diagnoses, allowing the identification of anomalies and the measurement of tissues, recognition of individuals, identification of several objects and industrial analysis.

Despite the large number of existing techniques, digital image segmentation is still an open field, due to the subjectivity and complexity of subdividing of the image into the most diverse scenarios and purposes. Techniques of image segmentation can use a variety of strategies and features, or a combination of them to achieve the goal, which is a correct segmentation.

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 33–42, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90304

Some common features in carrying out the segmentation are: intensity values of luminance and chrominance, textures and morphological forms. The known segmentation strategies come mainly from the methods of point, line and edge detection, thresholding, and region-based techniques [1]. The segmentation of images may be slow, depending on the computational resources, method and resolution used.

Several works, [2], [3], [4], [5], [6], [7], use graphs to represent the elements of an image to perform the segmentation process. Once the graphs can represent any discrete data, model relationships and use several functions with combinatorial optimization, making feasible the construction of widely efficient solutions [8]. An approach that has emerged in this context is the modelling of an image as a graph, for further segmentation through community detection algorithms, which generally use algorithms such as Efficient Label Propagation [9] [10] [11], Fast Greedy [12] [13] and Louvain Method [14]. Such an approach is directly related to the area of complex networks.

Through more detailed studies of the concepts of complex networks it is possible to identify a modality that has been explored to improve the analysis of the different types of networks, which are the networks of multiple layers. They have become promising because they allow the analysis of networks with multiple resolutions that vary over time, multiplex networks and hybrid networks with the characteristics presented [15] [16]. These networks allow to assign a greater number of relevant features to the context in the formation of the communities, seeking to offer greater precision in the various applications.

The studies of the multilayer networks are not recent, but they were stagnant and have now been resumed becoming one of the main research topics of the network science. Over the years, monoplex networks have been extensively explored and today researchers are seeking to extend and generalize their various features and functions for multilayer networks. However, such extensions and generalizations are not simple given that new phenomena arise given the new peculiarities.

The detection of communities in multilayer networks is already something more recent and does not yet have several concrete concepts, even the concept of degree that is somewhat trivial in single-layer networks, does not yet have a well-defined generalization in the multilayer structure. There are extremely rare researches that address segmentation of digital images in the context of multilayer networks or multiplex networks, reflecting the difficulty of generalizing certain trivial concepts in the area of complex networks.

Hu et al. in [7] present a proposal of image segmentation using community detection in multilayer networks based on the modularity measure presented by Mucha et al. in [15]. However, the authors have identified that the computational cost and the analysis and characterization of the performance of the modularity optimization for such application is still a challenge.

This work seeks to broaden the segmentation quality observed in previous research, extending the proposal for the use of multilayer networks of multiplex characteristics for the segmentation of images, in which each layer will use different features, being able to relate values of intensity of luminance, chrominance and simple characteristics of the regions of the images. Finally, an image segmentation method is presented, with pre-segmented images through super pixels, through the decomposition of the image into several features (luminance, chrominance and histograms),

modelling it through algorithms of multiplex networks to images represented as graphs, with reasonable processing time and with high precision.

As a main contribution of this work it is possible to quote the organization of an approach that initially contemplates the use of several features of the images in specific networks to allow the generation of specific communities for the accomplishment of segmentation of images through the detection of communities. Where such an approach provides a performance that enables its application in large synthetic and real images.

This article is organized as follows: Related work is presented in Section 2. The methodology is presented in Section 3. Results and discussion are in Section 4. Finally, conclusions are addressed in Section 5.

2. RELATED WORK

Browet et al. [14] present a technique based on graphs to detect segments or contours of objects in a given image. The algorithm is based on an approximation of the Louvain method (community detection algorithm) that deploys the community structures in a large graph. However, the optimal definition of a contour requires adjustment of parameters depending on the user or the application. The obtained communities are hierarchical, allowing to find sub-regions within an object. Through the realized experiments it is possible to observe that the technique avoids over segmentation through a weighted modularity scheme. The technique has a good performance of time and quality, but initially it only acts in grayscale images and it has clear dependence of parameters for the diverse applications.

However, Linares et al. [17] presented a new approach to segmentation of high-resolution images through the use of community detection and super pixels, reaching, through this combination, precision and a low processing time in the segmentation. For the detection of communities, the Fast Greedy algorithm was used and for the generation of super pixels, the Speeded-up Turbo Pixels. However, the graph used to generate community detection was obtained by sum of the Euclidean distance of the three CIELAB channels.

As presented in the introduction, Hu et al. in [7] present a technique based on multilayer networks for the segmentation of images. The algorithm of community detection has the modularity based on the generalization presented by Mucha et al. in [15] and is presented in equation (1), below:

$$Q_{multi} = \frac{1}{2\mu} \sum_{ijsr} \left[\left(A_{ijs} - \gamma_s \frac{k_{is} k_{js}}{2m_s} \right) \delta_{rs} + \delta_{ij} C_{jsr} \right] \delta(g_{is}, g_{jr}), \quad (1)$$

However, for the authors, an obstacle identified in the application of the technique was the computational performance, since each pixel was represented in the graph as a vertex.

3. METHODOLOGY

The diagram of Figure 1 generally illustrates the proposed method of segmentation through community detection algorithms in Multiplex Network and Super Pixels. Given an input image,

the pre-processing is applied by super pixels, considering several different features (intensity, color and others). Each of these features, or the super pixels that represent them, will form a layer of the network. After the generation of multiple graphs, the vertices are connected between the graphs, forming a multiplex network. The final segmentation is obtained by means of a community detection algorithm on the multilayer network that represents the image

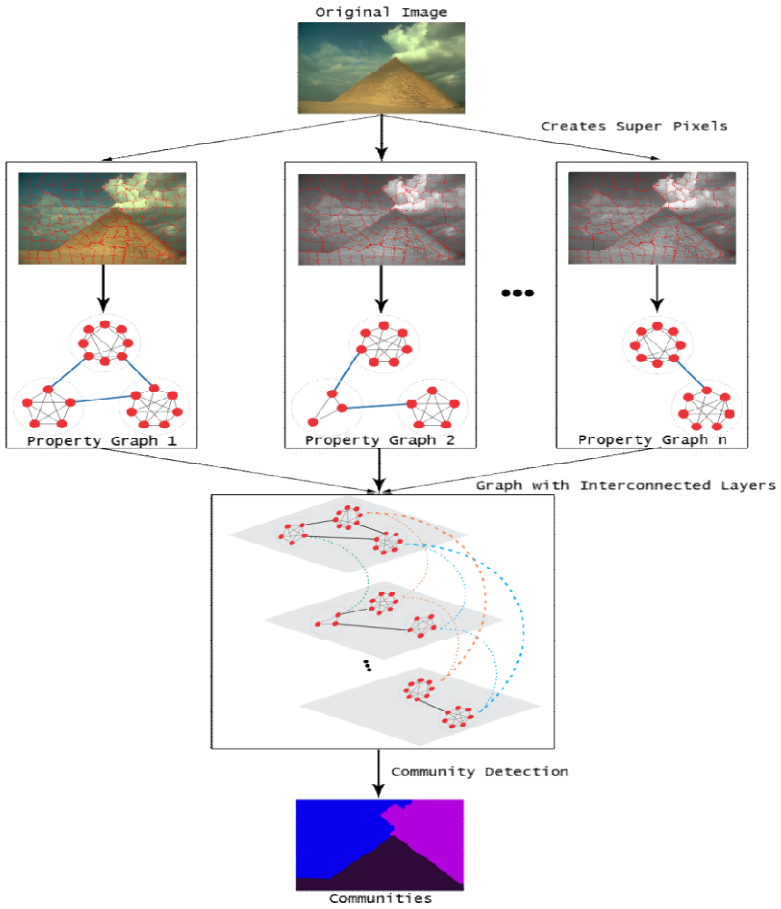


Figure 1. Proposed method of Image Segmentation by Means of Community Detection Algorithms in Multiplex Network and Super Pixels

3.1. PRE-PROCESSING - CREATION OF SUPER PIXELS

When receiving an original image this step consists of obtaining the super pixels of an image based on the various features of the image, in this step the main challenge is to obtain an algorithm extremely fast and effective in relation to the quality of pre-segmentation. The algorithm of generation of super pixels used was the Speeded-up Turbo Pixels [6], because it presents a good performance of time and quality in the pre-segmentation defined. The parameters used were the same ones observed in other studies [6] [17], among them: the size of super pixels, 10 X 10; number of iterations = 6 and the parameters $\lambda 1 = 1$ and $\lambda 2 = 0.9$.

However, it should be emphasized that the size of the super pixel influences the execution time of this stage and the quality of the segmentation, for images with different resolutions it is necessary to revise these parameters. Based on the super pixels, it is possible to generate the graphs that will compose the layers of the multilayer network, the methodology of connecting the vertices of such graphs is presented in the next section.

3.2. GRAPH GENERATION

In the creation of the multiple graphs it is essential to emphasize that it is composed of vertices obtained through the super pixels, which can be represented, for example, through the medium intensity, the average of each channel for the various LBP (Local Binary Pattern) color models and histograms for textures, among others. Already the edges are obtained by calculating the weight of a given property of two vertices of the graph. Therefore, if the weight between two vertices is less than a threshold t, a connection is established, where t can be modified based on the similarities of the pixels.

For all features the weight function can be defined by equation (2):

$$W_{i,j} = 1 - | I_i - I_j| \leq t, \qquad (2)$$

in which, Ii offers for example the average intensity of the super pixel i.

The threshold t is relative to the characteristics of the images. When it has a very high value, such a threshold can generate communities with mixed image regions. When it has a very low value, such threshold can generate over-segmentation in the image. Therefore, an adaptive threshold will be useful and adequate, consisting of the following steps for each super pixel:

1. Initially, a low threshold is displayed; if the super pixel has no connection, go to step 2;
2. The threshold is incremented, if the super pixel has no connection, repeat step 2 until the vertex has at least one connection.

To avoid the connection of super pixels very far, a radius = 5 (super pixels) will be applied, as observed in literature [17]. The Figure 2 illustrates the use of the radius in the proposed method. For each defined feature will be generated a graph with vertices based on their respective super pixels. Since a multiplex network is created, all the vertices of one layer will be connected with the corresponding vertices of the other layers.

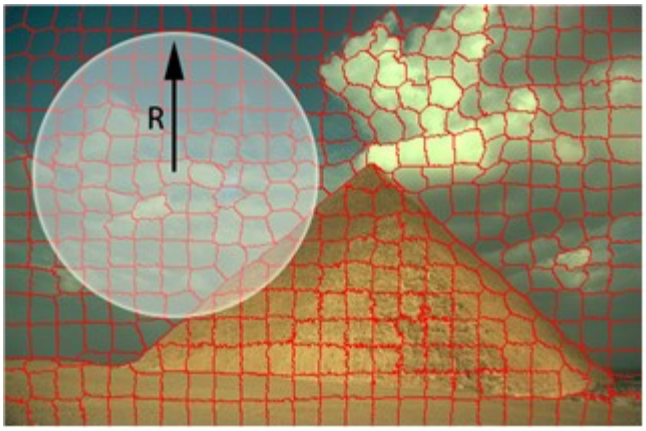


Figure 2. Pre-segmented image with super pixels illustrating a radius in which the super pixels will make connections in the creation of a graph

The creation of the graphs was implemented using the Igraph library in Language C. The process of image segmentation by community detection is presented in the next section and uses the graph for community detection.

3.3. IMAGE SEGMENTATION BY DETECTION OF COMMUNITIES

The proposed community detection process is based on the exploration of the most consolidated multilayer network algorithms in the literature and implemented in the MuxViz tool [18]. Among the algorithms used are the multiplex Infomap and the Louvain. The Louvain algorithms used adjacency sums to allow analysis of a multiplex network. Based on the communities generated by the MuxViz tool using the mentioned algorithms, the super pixels of the same community are united generating the segmentation of the image. These experiments aim to present the characteristics of the segmentation based on the proposed method exploring two algorithms of detection of different communities, Infomap and Louvain

4. EXPERIMENTAL RESULTS AND COMPARISONS

The results of two experiments for the analysis of the proposed method are presented below. All experiments were performed on an Intel Core i7 2.9 GHz computer with 4 GB RAM and the Ubuntu/Linux Desktop Operating Systems 64 bits 12.10 and Linux Mint 18.2 64 bits (Detection Communities).

4.1. EXPERIMENT A: SEGMENTATION OF REAL IMAGES WITH INFOMAP (MULTIPLEX NETWORKS) AND LOUVAIN (MULTIPLEX NETWORKS)

In this experiment, image segmentation based on a real image randomly selected from the Berkeley image database is being compared. In order to observe the segmentation characteristics in each algorithm for real images. The time of realization of each algorithm is also observed. It is important to emphasize that the generation of the detection of the communities was using the MuxViz tool and the features used as each layer of the graph were the CIELAB color channels and the histogram of the image.

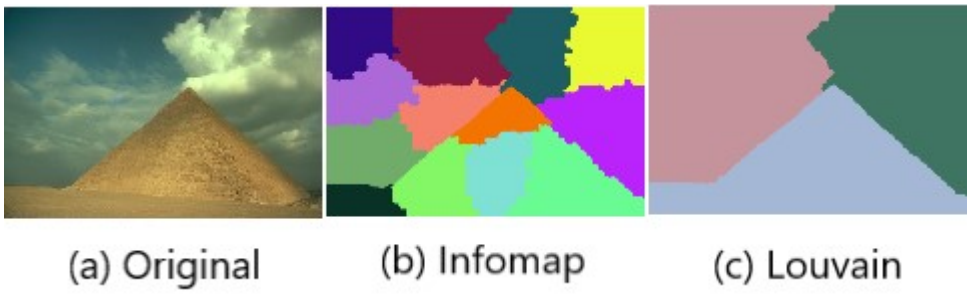


Figure 3. Original image (a), segmentation using Infomap (Multiplex Network) (b) and segmentation using Louvain (Multiplex Network) (c)

As can be observed in Figure 3, the segmentations performed by the Informap method have several over segmentations and the changes are very sensitive, since the method based on the Louvain algorithm is more robust and does not present over segmentation and is also faster to execute. The Infomap algorithm takes about 0.5 seconds longer than the Louvain method, which consumes about 2 seconds, the time was not accounted for more precisely because it was not listed as an improvement factor proposed in this work. However, it is expected that the more communities the algorithm generates, the longer the processing time of the method, since the creation of more segments takes longer to process the output, the segmented image.

4.2. EXPERIMENT B: QUALITATIVE ANALYSIS OF IMAGE SEGMENTATION

Considering the Berkeley image database and the segmentations presented in the human segmentation benchmark of this database, a comparison of the segmentations performed for a quick quality comparison is presented, as for each image there are several segmentations in this work, we present only the segmentation with the number of segments similar to the one returned in experiment A.

The segmentation that most approached the segmentation based on the Infomap algorithm was the segmentation of the image in 11 communities. The segmentation that most approached the Louvain method was the segmentation of image with 3 communities.

As can be observed in Figures 4 and 5, where the green lines represent the segmented regions, although both segmentations did not obtain the same results, both segmentations are close to the benchmark used, especially the segmentation performed by the Louvain method. However, common subjectivity must be considered in segmentation tasks that vary widely by observer. But it should be added that rapid post-processing actions can still be implemented to minimize over segmentations in algorithms that provide outputs with these characteristics.

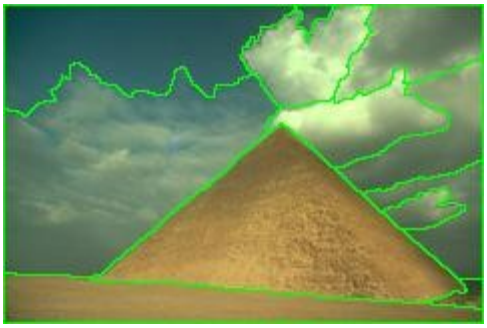


Figure 4. Human segmentation of image 299091.jpg of Berkeley Segmentation Dataset with 11 Segments

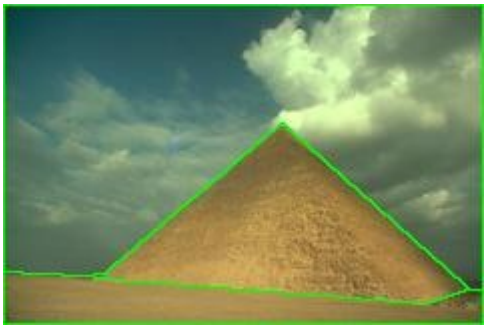


Figure 5. Human segmentation of image 299091.jpg of Berkeley Segmentation Dataset with 3 Segments

5. CONCLUSIONS

This work presents a useful method based on multiplex network and super pixels for the segmentation of images. Initial results provide quality segmentation. This method is also robust because it depends on specific parameters only in the pre-segmentation phase. It is interesting to note that such a method can act with several image features without the need of performing mathematical operations that can lead to the loss of feature information during the generation of graphs. However, issues related to the other properties that can be used deserve further investigation.

With the emergence of new generalizations in the area of multilayer networks it is interesting to review the algorithms used, since they can allow greater versatility and new adaptations in the method of segmentation of images. As future works, it is expected that multiplex networks will be replaced by multilayer networks with different super pixel sizes for each layer (property).

REFERENCES

[1] Gonzalez, Rafael C. and Woods, R. E. (2018). Digital Image Processing. 4th Edition. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[2] Boykov, Y. and Funka-Lea, G. (2006). Graph cuts and efficient n-d image segmentation. Int. J. Comput Vision, 70(2):109–131.

[3] Eriksson, A., Barr, O., and Astrom, K. (2006). Image segmentation using minimal graph cuts. In SSBA Symposium on Image Analysis.

[4] Vicente, S., Kolmogorov, V., and Rother, C. (2008). Graph cut based image segmentation with connectivity priors. In CVPR.

[5] Peng, B., Zhang, L., and Yang, J. (2010). Iterated graph cuts for image segmentation. In Proceedings of the 9th Asian Conference on Computer Vision - Volume Part II, ACCV'09, pp. 677–686, Berlin, Heidelberg. Springer-Verlag.

[6] Çigla, C. and Alatan, A. (2010). Efficient graph-based image segmentation via speeded-up turbo pixels. In Image Processing (ICIP), 2010 17th IEEE International Conference on, pp. 3013–3016.

[7] Hu, H., van Gennip, Y., Hunter, B., Bertozzi, A. L., and Porter, M. A. (2012). Multislice modularity optimization in community detection and image segmentation. In Vreeken, J., Ling, C., Zaki, M. J., Siebes, A., Yu, J. X., Goethals, B., Webb, G. I., e Wu, X., editors, ICDM Workshops, pp. 934–936. IEEE Computer Society.

[8] Lézoray, O. and Grady, L. J. (2012). Image Processing and Analysis with Graphs, Theory and Practice. CRC Press.

[9] Raghavan, U. N., Albert, R., and Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. Physical Review E, 76(3):036106+.

[10] Wang, F., Wang, X., and Li, T. (2007). Efficient label propagation for interactive image segmentation. In Machine Learning and Applications, 2007. ICMLA 2007. Sixth International Conference on, pp. 136–141.

[11] Elqursh, A. and Elgammal, A. (2013). Online motion segmentation using dynamic label propagation. In Computer Vision (ICCV), 2013 IEEE International Conference on, pp. 2008–2015.

[12] Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. Phys. Rev. E, 69:026113.

[13] Abin, A. A., Mahdisoltani, F., and Beigy, H. (2014). Wisecode: wise image segmentation based on community detection. The Imaging Science Journal, 62(6):327–336.

[14] Browet, Arnaud & Absil, Pierre-Antoine & Van Dooren, Paul. (2011). Community Detection for Hierarchical Image Segmentation. 358-371.

[15] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. Science, 328(5980):876–878.

[16] Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. Journal of Complex Networks.

[17] Linares, O. A., Botelho, G. M., Rodrigues, F. A., & Neto, J. B. (2017). Segmentation of large images based on super-pixels and community detection in graphs. IET Image Processing, 11(12), 1219-1228.

[18] Domenico, M. De, Porter, M. A., Arenas, A. (2015). MuxViz: a tool for multilayer analysis and visualization of networks, Journal of Complex Networks, Volume 3, Issue 2, 1 June 2015, Pages 159–176.

AUTHORS

Ivo Socrates M. de Oliveira holds a degree in Information Systems from the Luterano University of Palmas (2007), a master's degree in Informatics from the University of Brasília (2013) and currently holds a doctorate in Computer Science and Computational Mathematics from the University of São Paulo, São Paulo. He is currently Professor at the Federal Institute of Tocantins. He has experience in the area of Computer Science, with emphasis on Image Processing, working mainly on the following themes: image segmentation and pattern recognition.

Oscar A. C. Linares holds a degree in Computer Science from Universidad Católica San Pablo, Peru (2008). Master's degree (2013) and PhD in Computer Science and Computational Mathematics from the University of São Paulo, Brazil. Did internship as scholarship researcher (2017) at the Institute for Data Analysis and Visualization (IDAV) at the University of California, Davis, United States. He has experience in the area of Computer Science, with emphasis on Image Processing, Complex Networks, Computer Graphics and Big Data.

Ary Henrique M. de Oliveira holds a degree in Information Systems from the Luterano University of Palmas (2002), a Master's Degree in Computer Science from the Fluminense Federal University (2006) and PhD in Systems and Computer Engineering from the Federal University of Rio de Janeiro (2015). He is currently Adjunct Professor II of the Federal University of Tocantins. Has experience in the area of Computer Science, with emphasis on Methodology and Computer Techniques. Acting mainly on the following topics: Cloud Computing, Reproducibility of Scientific Experiments, Electronic Science, Scientific Workflow, Scientific Workflow System and e-Science.

Glenda M. Botelho holds a degree in Computer Science from the Federal University of Goiás (2007), a master's degree (February 2011) and PhD (September 2014) in Computer Science from the Institute of Mathematical and Computer Sciences (ICMC) São Paulo (USP). She is currently Adjunct Professor II of the Computer Science course at the Federal University of Tocantins (UFT) and coordinator of the Nucleus of Applied Computing (NCA) -UFT. She works in the areas of image processing and artificial intelligence, focusing mainly on data/image analysis and machine learning.

João Batista Neto holds a degree in Computer Science from the Federal University of São Carlos (1988), a master's degree in Computer Science and Computational Mathematics from the University of São Paulo (1991) and PhD in Biomedical Engineering - Imperial College - University of London (1996). He is currently an MS-3 professor at the University of São Paulo, in the city of São Carlos. He has experience in Computer Science, with emphasis on Image Processing, working mainly on the following topics: segmentation, feature extraction, and pattern recognition.

GEOMETRIC DEEP LEARNED FEATURE CLASSIFICATION BASED CAMERA CALIBRATION

Cheolhyeong Park, Jisu Kim and Deokwoo Lee

Department of Computer Engineering, Keimyung University, Daegu, 42601, Republic of Korea

ABSTRACT

This paper chiefly focuses on calibration of depth camera system, particularly on stereo camera. Owing to complexity of parameter estimation of camera, i.e., it is an inverse problem the calibration is still challenging problem in computer vision. As similar to the previous method of the calibration, checkerboard is used in this work. However, corner detection is carried out by employing the concept of neural network. Since the corner detection of the previous work depends on the exterior environment such as ambient light, quality of the checkerboard itself, etc., learning of the geometric characteristics of the corners are conducted. The pro-posed method detects a region of checkboard from the captured images (a pair of images), and the corners are detected. Detection accuracy is increased by calculating the weights of the deep neural network. The procedure of the detection is de-tailed in this paper. The quantitative evaluation of the method is shown by calculating the re-projection error. Comparison is performed with the most popular method, Zhang's calibration one. The experimental results not only validate the accuracy of the calibration, but also shows the efficiency of the calibration.

KEYWORDS

Calibration, Neural network, Deep learning, Re-projection error, Depth camera

1. INTRODUCTION

This document describes, and is written to conform to, author guidelines for the journals of AIRCC series. It is prepared in Microsoft Word as a .doc document. Although other means of preparation are acceptable, final, camera-ready versions must conform to this layout. Microsoft Word terminology is used where appropriate in this document. Although formatting instructions may often appear daunting, the simplest approach is to use this template and insert headings and text into it as appropriate. In the past a few years, practical applications of computer vision have been increase in the areas of automotive engineering, medicine, security, etc[1]. Estimation of 3D coordinates of the 3D real world scene from multiple 2D images is the central part of computer vision technology. Quality of the estimation of 3D coordinates (or depth) highly depends on establishing the relationship between 3D coordinates and 2D coordinates and estimating distortion parameters of lenses of cameras. These procedures are called camera calibration. This paper proposes the approach to increasing accuracy of the calibration. In particular, the procedure of corner detection of checkerborard, applies deep neural network. In the past decades, one of the most popular calibration algorithms is Zhang's method, and the crucial part is calculating homography matrix[2]. Homography matrix can be defined as four categories. The first one is that homography is established by calculating a relationship between 2D and 3D coordinates. The second one is that the homography is established by calculating a relationship between 2D coordinates of images. The third and the fourth one is established using the relationships of straight lines contained in a pair of images, and these are detailed in Zhang's

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 43–49, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90305

paper [2]. The most popularly used method of calibration is using objects that have known patterns, e.g., grid pattern, circular pattern, etc. In general, multiple captures of the objects from different locations are carried out followed by fully exploiting geometric relationship between patterns in images. Contrary to the method explained above, calibration can be performed without specially designed objects and the method is called self-calibration. Since the calibration using the patterned objects shows limitations in terms of computation time. Experimental environment also affects to the calibration accuracy. However, self-calibration can alleviate the limitations of the existed calibration methods because it does not depend on the experimental conditions. Concerning all above, calibration can be categorized into two methods, conventional calibration and self-calibration. Self-calibration has brought great attection in the past a few years because many practical applications such as vehicles, robotics and medicines desire real-time and efficient calibration system. Conventional calibration has limitations as follows. In the first, multiple captures in good experimental condition are required. Cameras capture patterned objects at least 5 times from different locations (usually more than 10 captures). In the second, calibration needs iterative optimization procedures. Since the optimization procedure is basically inverse problems, the result can be overfitting leading to unexpected wrong results. In general, self-calibration chiefly deals with epipolar constraints defined from geometric relationship between 3D objects and 2D images. From the geometric relationship, fundamental matrix is estimated, and the matrix is one of the criterion decides the accuracy of calibration accuracy [3,4,5,6]. The conventional calibration still shows higher accuracy because the algorithm has been already stabilized and the calibration is usually carried out in almost ideal conditions with the known patterns such as square, circular (periodic and aperiodic) ones [7].

In this paper, we propose the approach to increasing the accuracy of camera calibration. Conventional approach, using patterned objects, with a concept of deep neural network (deep learning) is employed. Deep learning is applied in the course of corner detection. Although checkerboard is used in this work, by applying deep neural network, successful trained parameters of the network lead to fewer number of captures of checkerboard images. In particular, this paper shows that only two captures maintained or outperforms the existed calibration method in terms of re-projection error. Training of the corners are carried out in 3 hidden layers each of which contains 5 steps, Prepare, FAPL, PAMFG, Reconstruction and PADCROP [8]. Evaluation result is compared to the Zhang's method and the criterion for comparison is re-projection error. The rest of the paper is organized as follows. Section 2 introduces deep learning based calibration, the most contribution of this paper, followed by substantiating the proposed algorithm by showing experimental results (including comparison results) in section 3. Section 4 concludes this paper.

2. Introduction To Projection Matrix

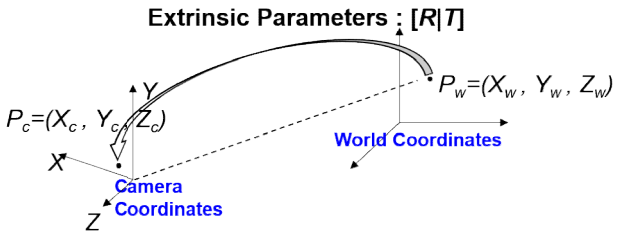


Figure 1. Extrinsic parameters define relationship between camera and real world coordinates

Camera calibration provides geometric relationship between coordinates of real world and ones of camera domain, or between the cameras themselves, and it also provides internal characteristics of the cameras. The former ones are called the extrinsic parameters and the latter ones are called

the intrinsic parameters. The extrinsic parameters are composed of a rotation matrix(R) and a translation vector(T). R and T are represented using 3×3 matrix and 3×1 vector, respectively (Figure. 1). The intrinsic parameters are composed of focal length, skew factor and principal point. Intrinsic parameters also deal with lens distortion. Concerning the extrinsic and the intrinsic parameters, the relationship between 2D and 3D coordinates can be written as follows. Please note that the first paragraph of a section or subsection is not indented. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either. Subsequent paragraphs, however, are indented.

$$s\mathbf{m} = \mathbf{PM}, \qquad (1)$$

where s is a scaling factor, and **m** and **M** are 2D and 3D coordinates (homogeneous), respectively. P is composed of the intrinsic parameters A and the extrinsic parameters [R|T], i.e., P=A[R|T]. The intrinsic parameters are composed of the focal length, skew factor and principal point's coordinates in image plane. Skew factor describes geometric distortion of the axes of sensors, i.e., if there is no distortion skew factor is zero. Once we have the extrinsic parameters, relative distances of real world scene of interest can be estimated. If the absolute distance value is required, the intrinsic parameters are needed in addition to the extrinsic parameters. The absolute value of depth is written as

$$z = bf/d, \qquad (2)$$

where z is a distance between the optical center and the 3D point of interest, b is a distance between optical centers of cameras, d is a disparity value resulted from parallax of two view points and f is the focal length of cameras (focal length of cameras are assumed to be equivalent here). In other words, camera calibration is a process of estimating the values in P. Cameras are assumed to be identical, but normalized coordinates are sometimes needed if different cameras are used [9].

3. Introduction To Projection Matrix

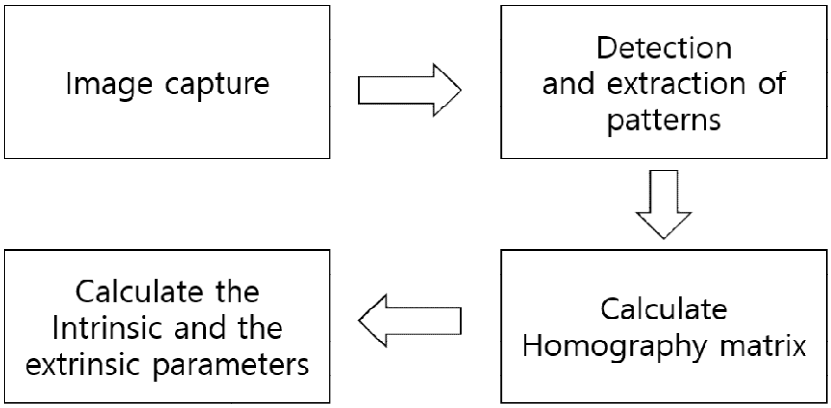


Figure 2. Conventional approach for camera calibration.

Figure 2 describes overall flow of conventional calibration of depth camera. More de-tails about conventional calibration methods are explained in [2, 7, 10, 11]. Among the methods proposed in the past few decades, Zhang's method is considered the most reliable in practice. In the method, Homography estimation is also one of the most important procedures. Homography matrix defines the geometric relationship between a pair of images. To achieve successful estimation of Homography matrix, feature extraction from the images is important. Homography between two images can be represented as follows.

$$\begin{pmatrix} \mathbf{x_L} \\ \mathbf{y_L} \\ \mathbf{w_L} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} \mathbf{x_R} \\ \mathbf{y_R} \\ \mathbf{w_R} \end{pmatrix}, \qquad (3)$$

where $(x_L \;\; y_L \;\; w_L)^T$ and $(x_R \;\; y_R \;\; w_R)^T$ are coordinates of 2D images from left and right camera, respectively. Since homogeneous coordinates system is used, w_L and w_R can be omitted or set as a constant number. Once homography is estimated, nonlinear optimization is applied to calculate calibration parameters. Calibration accuracy is evaluated using re-projection error. In this section, corner detection from checkerboard is performed based on deep neural network. In the first, checkerboard from whole image is extracted and the corner is detected from learned and weighted parameters.

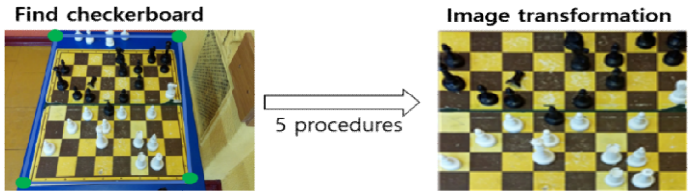


Figure 3. Learning based detection of chessboard and geometric transformation of an image.

Detection of a chessboard and of patterns lead to learning of the detection of corners in checkerboard for calibration. In Figure 3, geometric image transformation is carried out to accurately find corners. Prior to the transformation, a number of procedures is conducted. The procedures (5 procedures in Figure 3) are depicted in Figure 4.

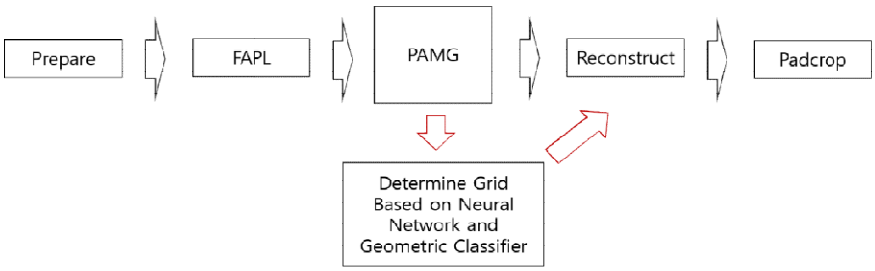


Figure 4. Learning procedures for corner detection.

As shown in Figure 3, 4 detection and extraction of corners from images employ deep learning algorithm. In the neural network in this paper, 3 hidden layers are used. Each layer processes Prepare, FAPL, PAMG, Reconstruct and Padcrop. Preapre is a preprocess to find straight lines in an image using retinex algorithm by subtracting unnecessary parts from the image [12]. FAPL extracts straight lines using retinex from the preprocessed image. In this procedure, extraction of the lines is carried out using image processing techniques such as “overexpose”, “blur” and “tint”. PAMG detects corners from the image, and the geometric image transformation can be done using Padcrop. In PAMG, neural network and geometric classifier achieves high accuracy of corner detection. By learning the structure of corners, the detection is success-fully accomplished from warped images. The activate function generate two outputs, the one is success and the other one is failure of detection. In the process of Reconstruct, geometric transformation of grid patterns is conducted as shown in the Figure 3 (right). To detect grid patterns, grid density formula is used, written as where p is a number of corners in images, n is an area and S represents density. Detection algorithm select corners that have the higher density. In the next, experiments of the proposed approach are provided.

$$S = \frac{p^3}{nlog(n)}, \qquad (4)$$

4. EXPERIMENTS

This section provides the experimental results and the comparison results with Zhang's calibration method. Since a size of image affects to the computational complexity of an algorithm, in Table 1, size of images used in this work is provided.

Table 1. Image size and a number of squares in images

Status of images	Size	Number of corners
Before image transformation (before Padcrop)	1200×1200	8×7
After image transformation (after Padcrop)	1200×1200	8×7

Detection of chessboard that is used in the process of learning employs neural net-work algorithm. In total, 9,664 images are used as training set, and the learning is conducted in the cases of existence of corners and non-existence of corners. The formal case is defined as "Train_ok" and the latter one is defined as "Train_no". There are 4,732 are used as "Train_ok" and are 4,932 are used as "Train_no". The results of learning are shown in Figure 5.

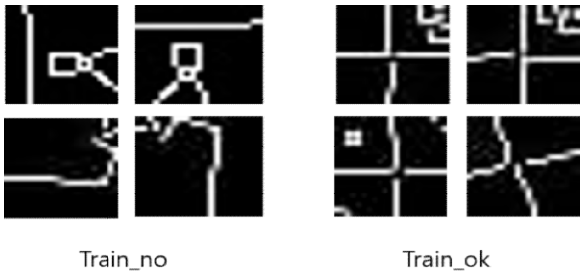


Figure 5. Training data for the cases of existence and non-existence of corners for calibration Table 2 provides the results of calibration using re-projection error when using a pair of images, i.e., stereo camera system. The proposed method and the existed method are compared. As shown in the Table 2, the proposed approach achieves the lower re-projection error.

Table 2. Heading and text fonts

Algorithm	Re-projection error
Zhang & Bouguet [2, 13]	1.9583
Proposed method	1.2735

In Table 3, the accuracy of calibration results by increasing iteration for optimization is provided. As shown in Table 3, accuracy is slightly increased.

Table 3. Accuracy (true-positive rate) with a number of iterations.

Number of iterations	Accuracy (True-positive)
50	0.9993
100	0.9996

5. CONCLUSION

In this paper, we proposed the approach to camera calibration by employing the concept of deep neural network. To increase performance of calibration, deep learn-ing is applied to the procedure of corner detection when checkerboard is used for the calibration work. The proposed approach not only detects the region of the board, but also extracts corners even if the images are warped. In particular, the proposed approach is promising because re-projection error has been decreased even though a number of captures has been significantly decreased. This work can contribute to simple and efficient calibration algorithm in diverse applications that require real-time computer vision techniques.

ACKNOWLEDGEMENTS

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding)

REFERENCES

[1] Liu, Y., Cheng, Y. and Wangm, W, (GIIC 2018) “A survey of the application of deep learning in comput-er vision.” In: Proceedings Volume 10835, Global Intelligence Industry Conference, SPIE, Beijing, China.

[2] Zhang, Z.: A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1330-1334 (2000).

[3] Kim, S-Y. and Han, J-H.: A Constrained Self - Calibration Technique. Journal of KISS Software and Applications 28(4), 358-368 (2001).

[4] Zhao, B. and Hu, Z.: Camera self-calibration from translation by referring to a known cam-era. Applied Optics 54(25), 7789-7798 (2015).

[5] Duan, S., Zang, H., Xu, M., Zhang, X., Gong, Q., Tian, Y., Liang, E. and Liu, X.: Camera self-calibration method based on two vanishing points. In: Proc. SPIE 9675, AOPC 2015: Image Processing and Analysis, LNCS, pp. 96752, Beijing, China (2015)

[6] Luong, Q.-T. and Faugeras, O.D.: Self-Calibration of a Moving Camera from Point Corre-spondences and Fundamental Matrices. International Journal of Computer Vision 22(3), 261-289 (1997).

[7] Heikkila, J.: Geometric camera calibration using circular control points. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(10), 1066-1077 (2000).

[8] Czyzewski, M.A.: An Extremely Efficient Chess-board Detection for Non-trivial Photos. arXiv:1708.03898v1 (2018).

[9] Kim, D-G., OpenCV Computer Vision Programming. 1st edn. Kame Publishing, Seoul (2014).

[10] Tsai, R.: A versatile camera calibration technique for high-accuracy 3D machine vision me-trology using off-the-shelf TV cameras and lenses. IEEE Journal on Robotics and Automa-tion 3(4) 323-344 (1987).

[11] Kim, T. and Park, T.: Calibration Method between Two 3D LIDARs for Autonomous Vehi-cle. In: 2012 KSAE Annual Conference Proceedings, LNCS, pp. 827-827, Gyeonggi Prov-ince, South Korea (2012).

[12] Provenzi, E., Carli, L., Rizzi, A. and Marini, D.: Mathematical definition and analysis of the Retinex algorithm. Journal of the Optical Society of America 22(12), 2613-2621 (2005).

[13] Leiner Barba, J. Lorena Vargas, Cesar Torres, Q.M. and Lorenzo Mattos, V.: Three-Dimensional Reconstruction Optical System Using Shadows Triangulation. AIP Conference Proceedings, LNCS, Vol 992, pp. 1073-1077, Campinas, Brazil (2008).

Author

Cheolhyeong Park is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on geometric image analysis, computer vision, computer graphics and machine learning. He is in the course of integrated B.S and M.S degree in computer engineering.

Jisu Kim is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on image processing, computer vision, signal processing and machine learning. He is currently pursuing his M.S degree in computer engineering.

Dr. Deokwoo Lee is an Assistant Professor in the department of computer engineering at Keimyung University. Dr. Lee has received B.S degree in electrical engineering from Kyungpook National University, Daegu, Republic of Korea, and M.S and Ph.D degree from North Carolina State University, Raleigh, NC, USA, respectively. He has been working on the areas of computer vision, image processing, signal processing and machine learning. In particular, he has been conducting camera calibration, bio-signal analysis and image denoising.

INTENTIONAL BLANK

EVALUATION OF DIFFERENT IMAGE SEGMENTATION METHODS WITH RESPECT TO COMPUTATIONAL SYSTEMS

[1]Ms. MehakSaini and [2]Prof.(Dr.)K. K. Saini

[1]Electronics & Communication Engineering,
Lovely Professional University, Jullunder, India
[2]Director, IIMT College of Engineering, Greater Noida, UP

ABSTRACT

Image segmentation is a fundamental step in the modern computational vision systems and its goal is to produce amore simple and meaningful representation of the image making it easier to analyze. Image segmentation is a subcategory of image processing of digital images and, basically, it divides a given image into two parts: the object(s) of interest and the background. Image segmentation is typically used to locate objects and boundaries in images and its applicability extends to other methods such as classification, feature extraction and pattern recognition. Most methods are based on histogram analysis, edge detection and region-growing. Currently, other approaches are presented such as segmentation by graph partition, using genetic algorithms and genetic programming. This paper presents a review of this area, starting with taxonomy of the methods followed by a discussion of the most relevant ones.

KEYWORDS:

Image segmentation , histogram analysis & Edge detectors.

1. INTRODUCTION

1.1 IMAGE SEGMENTATION

Segmentation is a pre-processing step where images are partitioned into several distinct regions and each is a set of pixels. Mathematically, image segmentation may be represented as P_n representing the regions of pixels R_n with n pixels the region can be described as union of all pixels connected, satisfying equation 1:

$$R_n = \cup_{i=1}^{n} P_i \quad (1)$$

Several image segmentation methods were proposed in the literature. However, a single method may not be efficient for a specific image class. Frequently, it is necessary to combine more than one method to solve interesting real-world problems. The main methods for image segmentation are based on histogram analysis, edge detection and segmentation by regions.

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 51–61, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90306

1.2 HISTOGRAM ANALYSIS

The image histogram analysis is a common method for image segmentation. A histogram is a graphical representation in which a data set is grouped into uniform classes such that in the horizontal axis the classes are represented and, in vertical axis, the frequencies in which the values of this class are present in the data set. Based on the central tendency or histogram variation it is possible to determine the cutoff point that will be used as threshold in these segmentation process. In this approach classes with high and low frequency are identified where a class with low frequency between two high frequency classes usually represents the best cutoff point to image threshold. An example of histogram analysis is presented in figure 1 (b) where classes with high and low frequencies can be seen.

An efficient approach for image segmentation based on histogram analysis is the Otsu method(Otsu, 1979).This method performs several iterations analyzing all possible thresholds to look for the best threshold T that presents the highest inter-class variance. This method assumes that the image to be segmented will be classified in two classes, object and background, and threshold point will be determined by the pixel intensity value that represents the minimum variance intra-class. In Otsu's method the threshold that minimizes the intra-class variance is exhaustively search and can be defined as a weighted sum of the variances of the two classes, as shown in equation 2:

$$\sigma^2_{intra} = W_a \sigma^2_a + W_b \sigma^2_b \quad (2)$$

Where, σ represents the variance of these classes and weights Wire presents the occurrence probability of each class being separated by a threshold T. Figure 1 (a) shows the original image in grayscale and figure 1 (c) shows image segmented by Otsu approach.

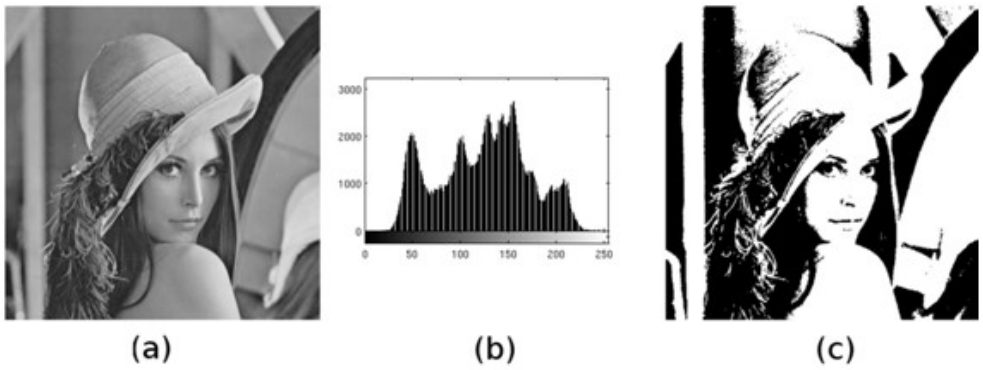


Figure 1. (a) Original gray-level image. (b) Original histogram image. (c). Segmented image using Otsu approach.

1.3 EDGE DETECTION

Edge detectors are common methods to find discontinuities in gray level images. An edge is a set of pixels of similar intensity level connected by adjacent points. We can find out edges estimating the intensity gradient. Edges in images can be divided in two distinct categories: edges of

intensity and edges of texture. In the first, the edges arise of abrupt changes in the image pattern and, in the second case, edges are detected by limits of textures in regions invariant to illumination changes(Tan, Gelfand, & Delp, 1989).

The Roberts edges operator(Fu & Mui, 1981) performs a simple 2-D spatial gradient analysis in digital images and emphasizes regions with high spatial gradient that can be edges. This method is a fast and simple convolution-based operator and usually the input of the method is a grayscale. Basically, Convolution masks can be applied to the input image to produce the absolute magnitude of gradient and the orientation. If applied separately it is possible to measure the gradient component in each orientation. The gradient magnitude is given by equation 3 and an example of image segmentation using Roberts edge operator can be seen in figure 2 (a):

$$|G| = |G_x| + |G_y| \quad (3)$$

In simple terms, the Sobel operator computes an approximation of the gradient of image intensity at each point and finds the contrast by a differentiation process (Duda& Hart, 1973). Thus, regions of high spatial frequency that correspond to edges are detected. The Sobel operator is a discrete differentiation operator and it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. Technically, the Sobel edge detector uses a simple pair of 3 x 3 convolution mask to create a series of gradient magnitudes like the Roberts edge operator. An example of image segmentation using Sobel edge operator can be seen in figure 2 (b).

In the edge detecting process, itis important consider three criteria: detection, localization and minimal response. Firstly, all edges occurring in image should be detected and there should be no-responses to non-edge. In other words, the signal-to-noise ratio should be minimized. Second, the error between the edge pixels detected and the real image edge should be minimized. A third criterion is to have a unique result to a simple edge, eliminating multiple detections to an edge (one edge should not response in more than one detected edge).

Based on these criteria the Canny edge detection algorithm is known as the optimal edge detector (Canny, 1986). To satisfy these requirements Canny calculated the variations that optimizes a given function and proposed your approach in five stages: noise reduction (smoothing), finding intensity gradients, non-maximum suppression, tracing edges through double threshold, and edge tracking by hysteresis.

Since the Canny method is sensitive to noise, a smoothing is necessary before trying to locate and detect edges. Thus, considering the good results and the facility to compute the Gaussian filter using a simple convolution mask, it is used in the Canny operator. A smoothed and slightly blurred image is produced after this step.

The next step is to find the intensity gradient of the image. So, the Canny operator uses algorithms to detect horizontal, vertical and diagonal edges in smoothed image. In this stage the Roberts or Sobel operators can be used to find the first derivative in the horizontal (Gx) and vertical (Gy) directions. From this edge gradient and direction, the edge direction angle can be determined, by using the arctangent function. The edge direction angle is approximated to one of

four angles representing vertical, horizontal and diagonals (0, 45, 90 and 135 degrees), as shown in equation 4:

$$\Theta = arctan\left(\frac{|G_x|}{|G_y|}\right) \tag{4}$$

After the directions are known, non-maximum suppression should be applied to trace along the edge in the edge direction and suppress any pixel value that is not considered to be an edge. Basically, this is done to preserve all local maxima in the gradient image ignoring anything else. This procedure will give the inline edge as result.

Tracing edges with double threshold is a way of eliminating streaking. Streaking is the breaking up of an edge contour remaining after non-maximum suppression. The edge pixels after non-maximum suppression step should be analyzed pixel by pixel. Probably many of them are edge pixels, but some may just be noise or color variations because of un even surfaces. To solve this problem it is possible to use double threshold with a high and a low value. Thus, edge pixels below the low threshold are considered non edges, edge pixels between low and high thresholds are considered weak edges, and edge pixels above the high threshold are considered strong edges. Finally, it edge tracing is done by hysteresis. In this process the strong edges are immediately included as edges in the final image and the weak edges are included if and only if they are connected to strong edges. Strong edges represent actual contours in the original image. However, noise and small variations have no influence enough to be characterized as edges. Thus, even in small numbers, the weak edges that are close to strong edges tend to compose the final result. The other weak edges distributed independently by image will be ignored. The final edge tracking process results is a binary image where each pixel is labeled as an edge pixel or non-edge pixel. Figure 2 (c) shows an example of image segmentation using Canny's operator.

Figure 2. (a) Segmented image using Roberts's method. (b) Segmented Image using Sobel's method. (c) Segmented image using Canny'soperator.

1.4 SEGMENTATION BY REGIONS

This section presents region-based methods for image segmentation using growth regions and splitting-merging approach.

In the splitting-merging approach, images are subdivided arbitrarily disconnecting regions and then these regions are again connected, or disconnected, to satisfy predefined constraints. The algorithm is iterative and first divides the image into four disconnected quadrants and then joins any adjacent region that satisfies the constraints. This process will be performed while one can split and join regions considering the constraints conditions(Trémeau & Borel, 1997).

In image segmentation by growing region, pixels are grouped based on predefined criteria that are started from a set of initial seeds. Starting from seed points, new regions are created by grouping neighboring pixels with similar properties. The selection of seed points in color images is a critical procedure. In this case, a set of descriptors based on intensity levels and spatial properties are required. The region growing images segmentation methods more cited in the literature are Watershed Transform and Mumford & Shah Functional(Raut, Raghuwanshi, Dharaskar, &Raut, 2009).

In Watershed Transform it is considered the gradient magnitude of image and surface topography. Pixels with gradient magnitude intensity (GMIs) represent the limits of the largest area and correspond to watershed threshold. This method has some variants and calculates the gradient magnitude intensity for all pixels in the image. With variance values of gradient it is possible to perform a topographic surface with valleys and mountains. The lower regions will correspond to low gradient while the highest regions correspond to the high gradient. The growth region procedure would be equivalent to flood performed at same speed in each local minimum, starting from the lowest one and then flooding the region until the maximum altitude reaches the global maximum.

In Watershed Transform the segmentation is perform with watershed regions that are formed flooding from local minimum. Flooding is controlled by dams that separate different local minima. The dams that emerge to the water surface are watershed lines. These dams are formed by closed contours around each regional minimum and correspond to ridges performing the segmentation by division line(Bleau& Leon, 2000). Figure 3 (b) shows an example of region growing image segmentation using watershed approach applied to figure shows in figure 3 (a).

Figure 3. (a) Original gray-level image. (b) Segmented image using watershed approach.

The Mumford and Shah functional algorithm assumes that each region is a group of pixels that behaves like an elastic material (like rubber). Regions can grow as long as possible to stretch this rubber. The basic principle is that the higher the variance between pixels in an area, the lower the elasticity of the material. Mathematically, the Mumford and Shah functional algorithm is very

simple and produces best results for general use when compared with other region growing algorithms. However, depending on the image, its time computational complexity can become a problem. In most cases it is necessary high investments to perform complex computational computation to obtain results(Jiang, Zhang, &Nie, 2009).

The image segmentation method is based on am a the metical model described by the energy equation of Mumford and Shah functional. This functional energy equation uses image grayscale variance considering that the higher gray level variance, the larger the difficulty to join these regions. This energy will determine if one can group with other regions, thereby delimiting the regions endings(Mumford &Shah, 1989). The simplified form of the Mumford and Shah functional expresses the segmentation problem as a minimization, as shown in equation 5:

$$E(u,K) = \int_{\Omega/k} \|u - g\|^2 \,\mathrm{dxdy} + \lambda l(K) \tag{5}$$

Where Ω is the domain of the image, K is a set of boundaries with total length l(K), g is a scalar or vector-valued function of the channels of the image on the domain Ω, u is a piecewise constant approximating scalar and λ is the regularization parameter for the boundaries. If λ is small, then a lot of boundaries are allowed and a good segmentation may result(Redding, Crisp, Tang, & Newsam, 1999). Figure 4 (c) shows an example of region growing image segmentation using Mumford and Sha mathematical model using the algorithm proposed by (Grady &Alvino, 2009) applied to figure shows in figure 4 (a). Figure 4 (b) shows the contours of the segmentation shown in figure 4 (c).

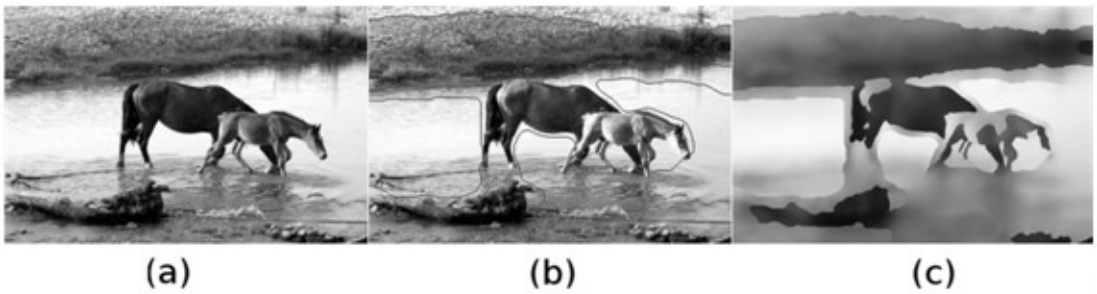


Figure 4. (a) Original gray-level image. (b) Original gray-level image with contours of the segmentation. (c) Segmented image using Mumford and Sha mathematical model and the algorithm proposed by (Grady & Alvino, 2009).

1.5 GRAPH PARTITIONING

In image segmentation by graph partitioning the initial image is partitioned as a weighted undirected graph. Each pixel is considered a node in the graph and edges are formed between each pair of pixels. Each edge weights is measured by similarity between pixels. The image is partitioned into disjointed sets aiming at removing edges that connect segments(Raut, Raghuwanshi, Dharaskar, &Raut, 2009). The optimal graph partitioning is performed by minimizing edge weights (energy function) that will be removed. Shi and Malik (Shi & Malik, 2000) proposed an algorithm to minimize normalized cut using a ratio that will can be standard for all the edges set.

In graph partitioning, G = (V, E)is an undirected graph with vertices Vi∈ V, where V is a segmented elements set and (Vi, Vj)∈ E are edges that corresponds to pair of neighboring vertices. For each (Vi, Vj) ∈ Eit is assigned a correspondent non-negative weight w(Vi, Vj) that indicates the dissimilarity measure between neighbor elements Viand Vj. In the case of image segmentation, the elements in V are pixels and the edge weights have the same dissimilarity measure between two pixels connected by same edge(Felzenszwalb & Huttenlocher, 2004).

Image segmentation by graph partitioning can be considered labeling problem. The object label (s-node) is set to1 and background (t-node) is set to 0. This process will be used in an energy minimization function by graph partitioning. For a reasonable segmentation, the graph partitioning should occur at the boundary between object and background. More specifically, object boundary energy should be minimized. The representation for this labeling can be L = {l1, l2, l3, … , li …, lp}, where p is the number of pixels of an image and li∈{0,1}. Thus, the L set is divided into two parts and pixels labeled to1 are part of theobject and theother are grouped as background(Yi & Moon, 2012). The energy function is defined by equation 6and can be minimized by approach presented in (Boykov & Funka-Lea, 2006):

$$E(\mathrm{L}) = \alpha\, \mathrm{R}\,(\mathrm{L}) + \mathrm{B}\,(\mathrm{L}) \qquad (6)$$

2. EVOLUTIONARY COMPUTING APPROACHES

A number of important segmentation image approaches based on evolutionary computing have been proposed in the literature, including genetic algorithms and genetic programming. Genetic algorithms and genetic programming are a stochastic search or optimization method based on the Darwin natural selection principles. This principle says that individuals better adapted to the environment have more probability of surviving and to generate descendants. Computing analyzing, individuals are candidate solutions to problems, adaptability is a quality of solutions and the environment is the problem instances where the solutions will be validated.

In genetic algorithms a possible solution is represented as individual composed by chromosomes, which are formed by genes. Several encoding schemes are used to represent chromosomes such as natural binary(the most usual), integer, gray code and real values. During the evolutionary process of genetic algorithms genetic operators, like crossover and mutation are applied. These operators change the individuals of a population to create new generations of solutions. The natural selection principle is performed by a selection procedure where individuals with good fitness evaluation have a higher probability to be selected for reproduction. The quality of a solution is evaluated by a fitness function that is calculated over an objective function. Both, fitness function and objective functions, evaluate how the solution is near the optimal value of the problem.

A. GENETIC ALGORITHMS

The use of genetic algorithms is motivated in the context of image segmentation by ability of to deal with a large and complex search space in situations where only a minimum knowledge is available about the objective function. An example of application is to adjust parameters in segmentation image algorithms as proposed in(Bhanu, Lee, & Ming, 1995). Usually, image segmentation algorithms have many parameters to be adjusted where the corresponding search

space is quite large and there are complex interactions among parameters. This approach allows determining the parameters set that optimize the output of segmentation. Other approach is proposed in(Bhandarkar, Zhang, & Potter, 1994)where genetic algorithms are used for edge detection that is cast as the problem of minimization of an objective cost function over the space of all possible edge configurations. A population of edge images was evolved using specialized operators. Finally, in another approach, the image to be segmented is considered as an artificial environment wherein regions with different characteristics, according to the segmentation criterion, are as many as ecological niches. In this approach genetic algorithms are used to evolve a population of chromosomes that are distributed all over this environment and each chromosome belongs to one out of a number of distinct species. The genetic algorithm-driven evolution leads to distinct species to spread over different niches. The distribution of the various species at the end of the evolution unravels the location of the homogeneous regions in the original image(Andrey, 1999).

B. GENETIC PROGRAMMING

Genetic programming is a method for automatic programs and it was derived from genetic algorithms. In Genetic programming the possible solutions for a problem are represented as programs in the form of trees. The functions are represented as internal nodes of trees and the inputs to the functions are represented as terminals (leaves) of the tree. The function set and terminals can be provided by the user and oriented to the specific problem to be solved. Recently, genetic programming has received interest as a methodology for solving computer vision problems because of its ability to select specific filters for detecting image features or to construct new features. It is possible to detect three approaches based on genetic programming: to detect low-level features, which have been predefined by human experts, such as corners, edges and vegetation indices using remote sensing; to construct new low-level features to specific problem domain and that is not necessary to be interpreted by human experts; and approaches to solve a high-level recognition problem such as object detection, image classification and texture segmentation(Olague& Trujillo, 2011).

An example of image segmentation based in genetic programming is presented in (Perlin& Lopes, 2013) that use genetic programming to solve a segmentation image problem. In this work the image segmentation problem is seen as classification problem where some regions of an image are labeled as foreground (object of interest) or background. A set of terminals and non-terminals composed by algebraic operations and convolution filters are provided for the genetic programming. The fitness function is defined as the difference between the desired segmented image and that obtained by the application of the mask evolved by genetic programming. Another approach describes a combined evaluation measure based on genetic programming. One of the greatest challenges while working on image segmentation algorithms is a comprehensive measure to evaluate their accuracy. Thus, the proposed approach allows nonlinear and linear combination of single evaluation measures and can search within many and different combination of basic operators of single measures to find a good set(Vojodi, Fakhari, & Moghadam, 2013). An example of image segmentation using genetic programming can be seen in figure 5 (a). This image was performed by (Perlin & Lopes, 2013) approach and was applied to image shows in figure 5 (b).

Figure 5. (a) Original color image. (b) Segmented image using Genetic Programming and algorithm proposed by (Perlin & Lopes, 2013).

3. CONCLUSION

In this paper a comprehensive analysis of concerning methods for image segmentation is discussed, that is a very important procedure in image processing and computer vision. We presented methods based on histogram analysis, edge detection, region-growing and graph partition. Finally, we pointed out some recent approaches based on genetic algorithms and genetic programming.

Image segmentation based on the analysis of the histogram is the most commonly used method, specially the method proposed by Otsu. Methods of edge detection and region growing have been widely used in image segmentation since they have pretty good results. The main methods of edge detection and region growing most frequent in the literature were Canny, Mumford & Sha and Watershed. We also present a more recent image segmentation using graph partitioning. This approach has attracted the interest of researchers who work in this area.

Finally, we presented an approach for image segmentation using evolutionary computation, particularly genetic algorithms and genetic programming. The use of evolutionary computation in the context of image segmentation has attracted the interest of the scientific community by the robustness that this methodology has the ability to handle large and complex search space.

REFERENCES

[1] Andrey, P. (1999). Selectionist relaxation: genetic algorithms applied to image segmentation. Image and Vision Computing, 17 (3–4), 175-187.

[2] Bhandarkar, S. M., Zhang, Y., & Potter, W. D. (1994). An edge detection technique using genetic algorithm-based optimization. Pattern Recognition, 27 (9), 1159-1180.

[3] Bhanu, B., Lee, S., & Ming, J. (1995). Adaptive image segmentation using a genetic algorithm. IEEE Transactions on Systems, Man and Cybernetics, 25 (12), 1543-1567.

[4] Bleau, A., & Leon, L. (2000). Watershed-Based Segmentation and Region Merging. Computer Vision and Image Understanding, 77 (3), 317-370.

[5] Boykov, Y., & Funka-Lea, G. (2006). Graph Cuts and Efficient N-D Image Segmentation. Int. J. Comput. Vision, 70 (2), 109-131.

[6] Canny, J. (1986). A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (6), 679-698.

[7] Duda, R., & Hart, P. (1973). Pattern Classification and Scene Analysis. New Jersei, USA: Wiley Publications.

[8] Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient Graph-Based Image Segmentation. International Journal of Computer Vision, 59 (2), 167-181.

[9] Fu, K., & Mui, J. (1981). A survey on image segmentation . Pattern Recognition, 13 (1), 3-16.

[10] Grady, L., & Alvino, C. (2009). The Piecewise Smooth Mumford & Shah Functional on an Arbitrary Graph. IEEE Transactions on Image Processing, 18 (11), 2547-2561.

[11] Jiang, X., Zhang, R., & Nie, S. (2009). Image Segmentation Based on PDEs Model: A Survey. 3rd International Conference on Bioinformatics and Biomedical Engineering, (pp. 1-4).

[12] Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics, 42 (5), 577-685.

[13] Olague, G., & Trujillo, L. (2011). Evolutionary-computer-assisted design of image operators that detect interest points using genetic programming . Image and Vision Computing , 29 (7), 484-498.

[14] Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics, 9 (1), 62-66.

[15] Perlin, H. A., & Lopes, H. S. (2013). A Genetic Programming Approach for Image Segmentation. In: A. Chatterjee, & P. Siarry (Eds.), Computational Intelligence in Image Processing (pp. 71-90). Springer Berlin Heidelberg.

[16] Raut, S. A., Raghuwanshi, M., Dharaskar, R. & Raut, A. (2009). Image Segmentation - A State-Of-Art Survey for Prediction. Proceedings of the 2009 International Conference on Advanced Computer Control (pp. 420-424). Washington, DC, USA: IEEE Computer Society.

[17] Redding, N. J., Crisp, D. J., Tang, D., & Newsam, G. N. (1999). An efficient algorithm for Mumford-Shah segmentation and its application to SAR imagery. Conference on Digital Image Computing: Techniques and Applications (DICTA-99) (pp. 35-41). Perth, Australia.

[18] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (8), 888-905.

[19] Tan, H., Gelfand, S., & Delp, E. (1989). A comparative cost function approach to edge detection. IEEE Transactions onSystems, Man and Cybernetics, 19 (6), 1337-1349.

[20] Trémeau, A., & Borel, N. (1997). A region growing and merging algorithm to color segmentation. Pattern Recognition, 30 (7), 1191-1203.

[21] Vojodi, H., Fakhari, A., & Moghadam, A. M. (2013). A New Evaluation Measure for Color Image Segmentation Based on Genetic Programming Approach. Image and Vision Computing, (In Press), 1-40.

[22] Yi, F., & Moon, I. (2012). Image segmentation: A survey of graph-cut methods. Proc. of International Conference on Systems and Informatics (ICSAI), 2012, (pp. 1936-1941).

AUTHOR

Prof.(Dr) K.K. Saini (DOB:31 July 1967) is B.E., M.Tech. & PhD in Electronics & Communication Engineering. Currently, He is at present Director, IIMT College of Engineering , GN, UP, India. His Research area and area of interest is Optical Communication, Chaos based comm., Satellite Communication and Reliability Engineering. He has published more than 500 research papers in various reputed international journals, international conferences and IEEE International Transactions. He has guided Dissertation of more than 150 M.Tech. students and 10 Ph.D. scholars. He has chaired various international conferences as a chief guest in India and in foreign countries also. He is very senior member of various research organizations throughout world. He is also awarded life time achievement award in 2015 in Bangkok, Thailand and delivered webinar in 2014. For further detail do refer is website [WWW.DRKKSAINI.IN]

INTENTIONAL BLANK

SMARTGRAPH: AN ARTIFICIALLY INTELLIGENT GRAPH DATABASE

Hal Cooper[1], Garud Iyengar[1], and Ching-Yung Lin[2]

[1]Department of Industrial Engineering and Operations Research
[2]Department of Electrical Engineering, Columbia University, New York, USA

ABSTRACT

Graph databases and distributed graph computing systems have traditionally abstracted the design and execution of algorithms by encouraging users to take the perspective of lone graph objects, like vertices and edges. In this paper, we introduce the SmartGraph, a graph database that instead relies upon thinking like a smarter device often found in real-life computer networks, the router. Unlike existing methodologies that work at the subgraph level, the SmartGraph is implemented as a network of artificially intelligent Communicating Sequential Processes. The primary goal of this design is to give each ``router'' a large degree of autonomy. We demonstrate how this design facilitates the formulation and solution of an optimization problem which we refer to as the "router representation problem", wherein each router selects a beneficial graph data structure according to its individual requirements (including its local data structure, and the operations requested of it). We demonstrate a solution to the router representation problem wherein the combinatorial global optimization problem with exponential complexity is reduced to a series of linear problems locally solvable by each AI router.

KEYWORDS

Intelligent Information, Database Systems, Graph Computing

1. INTRODUCTION

Data is often thought of in the context of input and output, to be used or analyzed by some external program or process. The structure of graph data, however, can indicate useful information about how to best execute graph-based algorithms on that structure. This is demonstrated by a key refrain for existing graph computing paradigms; to "think like a vertex" [1]. In this work, we demonstrate the benefits of further integrating graph data with the analytics run on that data by creating an artificially intelligent graph database. When graphs have knowledge of their own properties, the ability to send messages to other graphs, run calculations concurrently, and perform self-modification, a graph is no longer a static source of data. It instead begins to resemble a network of routers. In this work we replace the "think like a vertex" mantra with "think like a router", using a router inspired abstraction to create a "SmartGraph" database. The method distinguishes itself from existing graph databases through the artificially intelligent router abstraction; the routers that are defined by the subgraphs they encapsulate manage graph representation, concurrency and execution of operations themselves as opposed to being simple static data managed by an external process.

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 63–77, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90307

Asynchronous concurrent execution is difficult in traditional distributed graph computing systems like Pregel [2] in part because the "think like a vertex" mantra (that was chosen to make "reasoning about programs easier") is typically implemented using Bulk Synchronous Parallel [3] methodologies. This involves top-level maintenance of lists of "active" vertices (and/or edges) during each "superstep" that all perform the same vertex calculation. This can waste many computing cycles [4], since many graph algorithms do not converge at the same rate across different nodes. Proposed solutions to this problem often focus on maintaining sets of which nodes (or edges) need to be updated during each iteration, and therefore despite acknowledgment of asymmetric convergence rates, ultimately still use a synchronous iterative system.

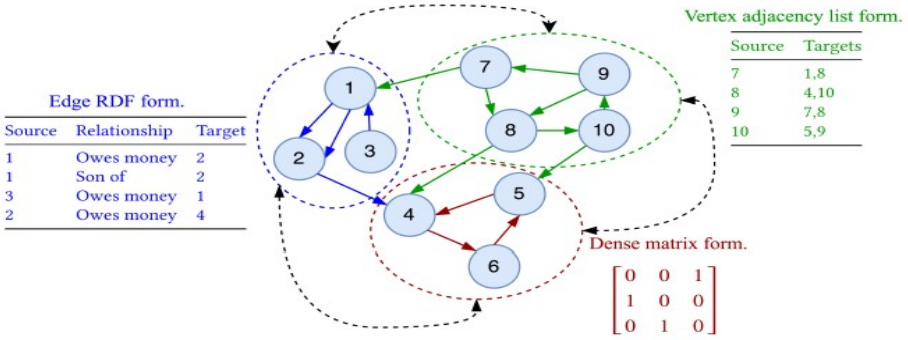


Figure 1. Toy graph (LHS) and a possible go routine router assignment

A key difference between the existing methods for graph computing and the system outlined in this paper is in having complex but manageable asynchronous concurrency of execution that eliminates the need for iterative supersteps and, more generally, the need for external macro-level management of the graph database and execution process. Though there exist previous works which have used subgraphs as a base unit for organizing computations, these works either limit their asynchronicity to "within subgraph" computations [5], or define subgraphs as connected components [6], [7]. In general, we wish to define subgraphs with less restrictions than existing methodologies allow in order to take advantage of more local structure, and to give these local structures individual autonomy.

There are many clear benefits to having highly concurrent asynchronous routers capable of "thinking for themselves". For example, routers can receive, execute, and transmit graph query operations without needing to know (and limit their speed to) the global supersteps, as well as automatically handle locally relevant operations concurrently with other routers; though we leave investigations of these properties to future work. In this paper, we focus on outlining the design of the Smart Graph itself and demonstrate how its asynchronous concurrency capabilities facilitate local artificial intelligence. In particular, we demonstrate the value of this approach through the "router representation problem", where "routers" use machine learning methods and solve local optimization problems to great effect (i.e. well approximating the globally optimal solution).

2. CONCURRENCY AND COMMUNICATING SEQUENTIAL PROCESSES

In a modern graph database, it is a functional requirement not only that the system be able to deal with large amounts of data, but that the system be able to deal with a large amount of different

requests with limited computational resources. The SmartGraph strives to facilitate massive concurrency involving graph operations on graph-based data by explicitly tying this graph structure into a concurrency management mechanism (through the router abstraction outlined in Section 3).

Concurrency is a property of a system that allows multiple processes (that may be related or entirely distinct) to have overlapping lifetimes. This does not necessarily mean that the multiple concurrent processes execute simultaneously at the hardware level (indeed, unlike parallelism, concurrency can be achieved on a single thread), it may simply refer to processes being able to be paused momentarily on a single thread, while other processes are given priority.

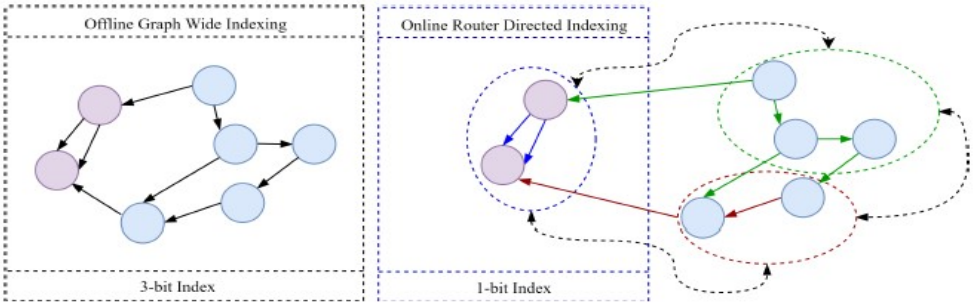


Figure 2. Toy comparison of graph-level and router-level vertex indexing

Communicating Sequential Processes are a method of implementing concurrency in programs based on message passing. This work takes some liberty with the theory of CSP as a whole, and instead focuses on CSP as it is implemented in the programming language Go (a.k.a. Golang) [8], developed by Google in 2009 with a focus on concurrency primitives as first-class citizens. Go is a language often described as a systems programming language, and is commonly used to design services, web-servers etc. It is with these capabilities in mind that we have created the Smart Graph in the Go language.

In Go, CSP is implemented through two main concurrency primitives: Go routines and channels. A go routine is essentially a very lightweight function that can be multiplexed onto different threads to be run concurrently with one another (and the Main, which is itself a goroutine). Goroutines communicate not by sharing memory, but “share memory by communicating” through the use of the channels. Of note is that although goroutines act like threads, they are not threads. This allows us to design a more lightweight concurrency mechanism, with many more go routines than available threads. This is ultimately of critical importance, as it allows us to simplify the management of graphs with a number of partitions at scales typically unheard of in the literature [9].

3. THE ROUTER

The SmartGraph concept is implemented in Go by explicitly tying goroutines to the graph-structured data that is to be explored or used in algorithms to be executed. Individual goroutines define (not merely control) the subgraphs of interest, with the relevant graph structure, vertex and edge properties defined in the variable size stack corresponding to the particular goroutine. The goroutines are our method of implementing the routers, and the highly concurrent functionality of

the goroutine is the mechanism through which we implement the communication and execution of analytics.

This approach contrasts strongly with the very strict definition of a subgraph (as connected components) used in prior works [6], [7] as it allows us to arbitrarily assign vertices and edges to routers. Though this may seem similar to partitioning done at the network or thread level in distributed graph computing [9], recall that goroutines are far more numerous than threads, and balancing occurs by multiplexing goroutines onto threads, in contrast to finding a partition that provides a balanced cut of the graph to all available threads. This allows us to partition the graph into many more pieces than might be indicated by the number of threads. Having separate graph structures in the routers facilitates taking advantage of local graph structure or optimizing with respect to locally requested operations. An example of this phenomenon, which we refer to as the "router representation problem" is explored in Section 4.

Through viewing this as the graph managing its own concurrent execution, one can do away with the top-level maintenance of sets of vertices, or algorithm iteration supersteps required by BSP methods. Concurrent execution is instead automatically handled by the router routines, as they will (through their status as a goroutine) immediately signal to the scheduler when they are ready (having received required data or messages from parent goroutines) to execute desired operations. This allows the system to be highly asynchronous, yet manageable thanks to the router abstraction. A basic example encapsulating a SmartGraph in goroutine routers is given in Figure 1. Observe on the LHS, a toy directed graph with no particularly noteworthy node or edge features. The RHS shows an example set of three goroutine routers that encapsulate the SmartGraph. Each router, defined by a dotted ellipse of a primary color, contains a subset of nodes and edges from the overall graph. The dotted black connections on the RHS represent channels between the source and destination routers.

The network router is useful not only as an abstraction model for efficiently implementing and executing graph algorithms, it also inspires useful functionality for the SmartGraph. Network routers have the ability to maintain useful information in memory, such as routing tables. They also possess the ability to perform custom routing logic independently of other routers. There are both obvious and subtle ways in which a graph database and analytics platform aping this functionality provides benefits. For example, a straight-forward method of employing the benefits of local representation is in saving memory address space as highlighted by Figure 2. The “local” aspect of the router AI can also be taken further than simple indices. With routers able to act independently, we introduce the notion of “local subgraph representations”, where the routers encapsulating a subgraph do so via different storage mechanisms and graph formats. For instance, one router AI may organize the subgraph it encapsulates using vertex adjacency lists, with another router AI learning to use an RDF framework. An example of this sort of router dependent representation is given in Figure 1. This approach has both computational benefits (since some operations have faster implementations in certain graph representations) and storage benefits (e.g. when using a sparse representation like CSR, or when compressing a subgraph consisting of nodes with high similarity).

4. THE ROUTER REPRESENTATION PROBLEM

The representation problem, as defined in this paper, is closely related to the more generic problem of selecting data structures either at compile-time or run-time, in order to minimize

memory usage, execution time, or some combination of the two. In contrast to this work, the literature focuses almost exclusively on optimization with respect to standard container objects (such as lists, sets, arrays, hash Maps etc) rather than graph specific data structures. A small number of papers due address the problem in a more graph specific context; for example, [10] directly considers the impact of basic graph operations and representations on execution time. However, it only uses the vertex adjacency list form and investigates how the single graph representation can be constructed by different container types, and so does not directly consider multiple graph specific data structures (i.e. alternatives to the adjacency list). Furthermore, [10] considers representations that "change" over time, but at any given time applies a single data structure to the entire graph. This means that the method cannot benefit from local graph structure. In contrast, [11] explicitly investigates different graph structure representations (both the adjacency list and the adjacency matrix), but again only considers their choice in application to the entire graph.

The data structure selection problem has been approached from three primary directions. The first is that of optimization, where benchmarking is used to create a function that approximates the execution time of a series of operations [10]. The second is using machine learning to generate rule sets that can be used to determine the choices of representations (e.g. "if BFS is called on a graph with density greater than 25 percent, use an adjacency matrix, else use an adjacency list") [12]. The final commonly used method is to simply provide a framework for implementing swap rules, and allow the user to specify precisely what those rules are manually [13].

Unique to this work, we explore methods for choosing local graph representations (that is, structural representations that apply locally rather than to the whole graph) by solving a set of optimization problems over learned models that consider both the local graph data, as well as the graph operations requested in relation to that data. Furthermore, we do so by deconstructing the problem such that a problem that initially appears exponential in complexity becomes linear in the number of routers by the number of router representations. In Section 4.1 we introduce the basics of the representation problem. In Section 4.2, we introduce our method for solving the representation problem under idealized circumstances, and in Section 4.3, outline how we learn the functions required to solve the problem in practice. This ultimately involves using average router behaviour as an approximation to input, which we argue in this section and demonstrate in Section 5, is an approximation that ultimately works very well.

4.1. REPRESENTATION METHODOLOGY

Let $G = (V,E)$ be a known property graph, i.e. a graph with properties attached to vertices and edges, where we allow multi-edges (e.g. representing different types of relationships between the same two objects). Let $P = (P_1,P_2,P_3,...,P_k)$ denote a partition of the edge set E. Let $V_i = \{u: (u,v)$ or $(v,u) \in P_i\}$. Note that we allow for a vertex v to be an element of more than one V_i. In a SmartGraph database, a separate router $\mathcal{R}_i$ encapsulates the subgraph (V_i,P_i).

We emphasize that the partitioning introduced here should not necessarily be considered in the same vein as traditional graph partitioning [9], as we are operating on a single machine unconstrained by threads, and as such, we do not overly care about partition balance.

Let S denote the set of all graph representations allowed by the SmartGraph. Let $R_i \in S$ denote the representation employed by router $\mathcal{R}_i$ for (V_i, P_i), $i = 1, ..., k$. In this work, the allowed set S of representations includes the Matrix Adjacency List (MAL; though we typically use the term adjacency list in full), wherein vertices are stored in a hash-table like structure, with lists of outbound vertices), and the Resource Description Format (RDF), which stores separate tables of edges and vertices. We deliberately use the term tables here, as the use of RDF can be thought of as functionally similar to implementing a graph database in a traditional relational database system. These two representations are by far the most common method of representing graphs in practice. For example, the RDF format is used in Spark, and it's GraphX [14] and GraphFrames [15] packages, whereas adjacency lists are used by Neo4j [16], a leading commercial graph database system.

There are many more graph representations, including adjacency matrices (often impractical in practice due to large storage requirements), sparse matrix representations (where-in a user is more focused on efficiently performing mathematical operations than graph operations; though the two are often closely connected), and representations with particular properties (such as allowing directed graphs only, disallowing multi-edges, cyles) etc. There are also custom representations designed to work efficiently with GPU operations [17], [18]. Thus it should be noted that the combination of representations used in this paper (and solved by the SmartGraph router abstraction herein) is not intended to argue that it will always result in the most efficient representative structure. Instead, we are arguing for a methodology for combining arbitrary different representation possibilities, and using the RDF and adjacency list as our representations for purposes of demonstration.

4.2. PREDICTING EXECUTION TIMES AND CHOOSING REPRESENTATIONS

We represent a database job $\mathbf{x} \equiv \{x_1, x_2, ..., x_M\}$ as a sequence of M known operations, that we classify as either "complex", or "basic". Complex operations are those that internally execute other operations (either complex or basic) to complete their execution, those that require knowledge of multiple routers (such as inspecting the number of vertices in a given router that are duplicated in other routers), and those that requiring complex input. What remains we describe as basic operations. We allow for directed acyclic dependence between operations of a job, i.e. all parents of an operation x_a, $\pi(x_a)$ must be executed before x_a. Let $\mathcal{B}$ denote the set of l basic operations. We assume that the average run time for a basic operation $b \in \mathcal{B}$ on a graph with n vertices and m edges, and representation $r \in S$ is given by a function $h(b, m, n, r)$ (which we assume is given, Section 4.3 will outline how we learn this and other functions).

Furthermore, we define a function $\mathbf{h}(\mathbf{b}, m, n, r)$, where $\mathbf{b} \in \mathbb{N}^l$ is a vector of basic operation counts, and $\mathbf{h}$ is the time to execute $\mathbf{b}$ (i.e. the linear combinations of functions h with appropriate parameters; where here we are not considering timing related to concurrent execution). Note that an operation initiated on router $\mathcal{R}_i$ may have to initiate calls to other routers in order to complete the operation. For example, since vertices are possibly duplicated, a "traverse neighbors" operation may require calls to other routers to check if those vertices exist there too, and get neighbors of such duplicates also. Let $c_{ij}(x_a, |x_a|) \in \mathbb{N}^l$ denote the number of

calls of basic operations $\mathcal{B}$ initiated on router $\mathcal{R}_j$ when the operation is initiated on router $\mathcal{R}_i$. We separately include the size of the input (e.g. number of vertices), $|x_a|$ involved in the operation for reasons that will become clear shortly. The approximate execution time $T(x_a,|x_a|,i,R)$ for an operation x_a initiated on router $\mathcal{R}_i$ is then given by

$$T(x_a,|x_a|,i,R) = \mathbf{h}(c_{ii}(x_a,|x_a|),|V_i|,|P_i|,R_i) + \sum_{j\neq i}\mathbf{h}(c_{ij}(x_a,|x_a|),V_j,|P_j|,R_j). \quad (1)$$

However, in practice jobs are sequences of operations, where the number of relevant graph objects for a job can depend upon its parents. For example, consider a job $\mathbf{p}:=\{x_1,x_2\}$ with $\pi(x_2)=x_1$ on a directed weighted graph G with partition P, where x_a is an operation "traverse edges from current vertex set if $w\geq\overline{w}$". For x_1, we assume that we have some initial vertex $z\in V_i$ as the current vertex set. However, we do not know in advance the size of the vertex set to which x_2 will apply. This means that in addition to knowing how many basic operations (and their type) $\mathbf{b}$ are required to execute a given operation, we also need a count of how many graph objects (e.g. vertices or edges) are filtered through by parent operations. We define $f_j(x_a)\in\mathbb{N}$ as a function that takes an operation and returns a count of the graph objects within partition part j that survived the filtering of the parent operations $\pi(x_a)$. Then the execution time of job $\mathbf{p}$ is given by

$$T(\mathbf{p},i,R) = T(x_1,|x_1|,i,R) + \sum_j T(x_2,f_j(x_2),j,R), \quad (2)$$

where the first term corresponds to execution x_1 on the initial router, and the sum represents running the subsequent operation on the "filtered" nodes. We have overloaded the notation T to correspond to runtimes for either jobs or operations, with the meaning clear from context. Note that $c_{ij}(x_a,1)\cdot|x_a|\neq c_{ij}(x_a,|x_a|)$ because there may be duplicates (e.g. multiple edges pointing to the same vertex) that could otherwise cause a vast overestimation (e.g. consider a fully connected graph) in the number of objects for the next operation, and therefore a large error in time estimation. It is easy to see how this approach extends to a much more general job $\mathbf{x}$ as

$$T(\mathbf{x},i,R) = \sum_{x_a\in\mathbf{x}}\sum_j T(x_a,f_j(x_a),j,R), \quad (3)$$

where we assume $f_i(x_1)$ is some known set, and $f_j(x_1)=\phi\ \forall i\neq j$. With this execution time function, we therefore wish to solve the following optimization problem:

$$\min_{\{R=(R_1,\ldots,R_k):R_i\in S\}} T(\mathbf{x},i,R) = \min_{\{R=(R_1,\ldots,R_k):R_i\in S\}}\sum_{x_a\in\mathbf{x}}\sum_j T(x_a,f_j(x_a),j,R). \quad (4)$$

The Problem (4) is a very complex combinatorial problem (as there are exponentially many, $k^{|S|}$, possible choices of the router representation vector R), where the representation of any given router i can influence the performance of operations sent to different routers j. Thus the choice of router representations must be made globally, with no opportunity for parallelization. However, we observe in Problem (4) that we have an important opportunity, by virtue of having established total counts of basic operations and their router occurrence locations.

Consider that the number of routers, and the length of the job $\mathbf{x}$ is finite. We can therefore interchange the order of summation on the RHS. Also, each complex operation ultimately results in a set of basic operations distributed across the routers, and that *we know this distribution of basic operations before solving the optimization problem* because it is not dependent upon representation because of how complex and basic operations are defined. This means, that supplied with appropriate c,f functions, we can rewrite the job $\mathbf{x}$ as a job $\mathbf{y}$, such that $\mathbf{y}$ consists only of basic operations (i.e. executing $\mathbf{y}$ executes $\mathbf{x}$). Furthermore, we can group the basic operations by router, such that for $\mathbf{y} = \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_k\}$, $\mathbf{y}_i$ is the set of all basic operations for router $\mathcal{R}_i$. That is, we have the following very significant result:

$$\underset{R=(R_1,...,R_k):R_i \in S\}}{\operatorname{argmin}} \sum_{x_a \in \mathbf{x}} \sum_{j} T(x_a, f_j(x_a), j, R) = \bigcup_{i=1,...,k} \underset{R_i \in S}{\operatorname{argmin}} \boldsymbol{h}(\boldsymbol{y}_i, |V_i|, |P_i|, R_i), \tag{5}$$

where we are abusing notation in assuming $\mathbf{y}_i$ also has an appropriate vector form of counts of basic operations on $\mathcal{R}_i$. There are two key points to Problem (5). The first is that we have successfully decomposed a very complex interdependent problem with exponentially many possible solutions, to a linear number (in the number of routers) of problems with a linear number of solutions (in the number of representations), a vast reduction in problem complexity. Furthermore, these optimization problems are local to each partition part. This means that each router can locally solve the simple problem associated with itself (potentially concurrently), and these locally optimal solutions together combine to form the *globally* optimal solution of Problem (4).

4.3 LEARNING FILTERING, COUNTING, AND TIMING

In practice, we do not know the c_{ij}, f_j, and h functions unless we run the operations (which is contrary to our intent of choosing representations before execution). Instead, we seek to learn approximation functions $\overline{c}_{ij}$, $\overline{f}_j$, and $\overline{h}$. To learn these functions, we send sample operations to *each router* (note the contrast to our learning of basic operation runtimes, which is done as a global precomputation) and learn based upon the average responses of each router. In job $\mathbf{p}$, different routers may have different vertex degree (affecting basic operation counts and $\overline{c}_{ij}$ because edges must be checked, even if they do not satisfy the condition to be passed to the next operation), and different edge weight distributions (affecting $\overline{f}_j$ and the number of objects passed to the next operation).

Unlike in distributed graph computing systems, where a job might represent a call to compute PageRank etc. over the entire graph, we explicitly identify the database functionality of the SmartGraph as the reason to define *read-only* jobs as queries. We wish to have a range of operations that is sufficiently flexible such that they can be composed into more complex graph operations. This means that their combination should be sufficiently expressive. In the extensive literature on graph query languages [19]–[21], the expressiveness of query formulations has been a central area of research, with expressiveness often traded off against complexity and efficiency. In practice, industrial graph query languages like Cypher and Gremlin have not been theoretically analyzed to any significant degree due to the complexity facilitated by the wide range of possible query operations permitted [22].

Query methodologies typically fall into two categories; path queries [19], and pattern matching [23]. For purposes of demonstration within this paper, we choose to use a path query approach rather than a pattern matching approach. We choose this approach because it is clearly suitable for the methodology outlined in Section 4.2. Consider that a path query typically has a form $I \overset{cond.}{\rightarrow} J$ where we are seeking paths from graph object set I to (potentially unknown) graph object set J that must satisfy conditions $cond.$ along the path. There is an extensive literature on graph query algebras, and we refer the reader to [20] for a recent survey on the topic. For our purposes, we define path queries by chaining single hops with conditions on edges and/or vertices. Observe that this maps very neatly into the functions c_{ij}, f_j, and h. We have some path query that is decomposed into a series of stages, and conditions upon those stages. Each stage requires observing a certain number of graph objects (related to c_{ij}), only some of which survive to the next stage (related to f_j), and these are internally composed of basic graph operations (related to h, see Table 1) with timing dependent upon representation.

In contrast, we define *write-only* jobs as operations that modify the structure or properties of the graph. Indeed, the existence of properties within the graph database strongly differentiates this problem, even from the existing attempts at data structure optimization, as they focus entirely on graph structure [10], [11], [24]. The existence of properties is critical to our methodology, as it is not possible to ignore the schema of the graph when profiling simple operations as in existing works (e.g. profiling a GetVertex() operation is not sufficient, since the schema of the vertex added will have a large impact on performance). We leave support for write-operations to future work, in this version of the representation problem we focus on read operations as graph-querying, a fundamental read operation, is so elemental to the everyday usage of graph databases.

In-line with our view of routers as artificially intelligent structures, we learn approximations to $\overline{c}_{ij}$ and $\overline{f}_j$ independently at each router. That is, each router generates and executes sample queries according to the routers schema, and stores the resulting models for solving a problem akin to Problem (5), where we make the substitutions $c_{ij} \leftarrow \overline{c}_{ij}, f_j \leftarrow \overline{f}_j,$ and $h \leftarrow \overline{h}$. Thus we are using average router performance as a model for these functions, something that is facilitated at finer grains as we increase the number of routers (recalling that it is the router abstraction and the implementation of routers without go-routines that allows us to fathom router numbers in the tens of thousands, which is greatly distinct from the relatively few shards used in typical thread-focused graph computing).

Table 1. Examples of basic and complex operations

Basic Operation	Complex Operation
AddInternalEdge	AddExternalEdge
AddVertex	GetSharedVertices
RemoveEdge	ExploreNeighborsOutsideRouter
RemoveVertex	FilterEdges
GetVertex	FilterVertices
GetInternalEdge	RoutersConnected
VertexExists	
InternalEdgeExists	
UpdateInternalEdgeProperties	
UpdateVertexProperties	
ExploreNeighborsInRouter	

5. EXPERIMENTAL RESULTS

We use random forests [25] to learn each of the $\overline{c}_{ij}, \overline{f}_j, \overline{h}$ approximations owing to their ease of use combined with their ability to capture non-linear interactions. To demonstrate the efficacy of the method, we construct an artificial graph system consisting of four vertex types; Person, Product, Location, and Mail (with 300,400,500, and 600 vertices of each type respectively). We split the graph by vertex type, such that a given router primarily contains a single vertex type. As we allow edges both within the routers and between them, each router also contains a number of duplicated vertex types copied from other routers due to external edges. We constructed the graph using the Erdos-Renyi model [26] such that average degree is given as in Table 2. We note here that we do not expect our method to require a graph substantially similar to the graph presented here-in, or even to partition by vertex type as we have done. Instead, this example was generated for ease of understanding and reproducibility. We have similar designed the average between router degree to be much smaller than the within-router degree. This is because methods of graph partitioning typically involving minimize communication between partitions [9], and so we believe this a reasonable choice (note that this work does not deal with the question of graph partitioning, since it would need to be solved jointly with the assignment problem; such investigation is left to future work).

In terms of edge properties, we add two properties to each edge; a bivariate Gaussian with correlation ρ chosen $\sim \mathcal{U}[0,1]$ for each edge type. We wish to demonstrate multiple properties in this early work in order to demonstrate that queries that interact with multiple properties are not an inconvenience for the described method.

5.1 BASIC OPERATIONS

For the purposes of learning basic operations, we simply run many instances of the basic operation on each edge and vertex type described in the schema, for varying sizes of the graph (that is; the only input available is the total number of vertices and edges in a graph). In this instance, we are not training specifically on the graph described in Table 2, merely graphs of

Table 2. Average out-degree for each vertex combination

	Person	Product	Location	Mail
Person	32	0.08	0	0.06
Product	0.1	26	0.12	0.04
Location	0.01	0.02	255	0.01
Mail	5	3	8	100

different sizes with the same vertex and edge types. For this reason, the learning of $\overline{h}$ functions, which we do independently for each representation type (RDF and MAL in our example), can be performed independently of any particular graph, only depending upon the schema. This indicates that if we expect multiple graphs with the same schema, we need only learn our approximations $\overline{h}$ once. In contrast, the methods described below require training on each individual graph.

5.2 COMPLEX TO BASIC COUNT MAPPING

In order to learn $\overline{c}_{ij}$, we generate a number (up to 10,000) of single-step test queries. These queries are of various types (e.g. filter a list of edges, filter in-bound from a list of vertices, filter out-bound from a list of vertices etc). In order to ensure that our methodology can be applied more broadly, we generate these test queries by randomly sampling vertices and edges from each router (the number of which is chosen uniformly from zero to the number of edges/vertices in the router). We then generate test queries randomly designating a property on the sampled graph object (in this paper, we only consider properties on edges), such that we are searching for edges that are greater than or less than the property at the sampled edge. Clearly this technique will generate queries that span the range of allowable property values, without any need to know in advance the distribution of any given property on the graph. We tag each query with a unique identifier, and track that identifier as results in the execution of basic operations (that is, for any given complex operation, we know exactly where basic operations occur, which operations they are, and how many of them occur).

We then teach a random forest to learn this mapping, such that test input includes the number of vertices or edges we wish to start the query from. In addition, we include the number of vertices shared between the source router (i.e. the router where the query is first initiated) and the other routers. We found that the inclusion of the latter was key to obtaining random forest models with very high accuracy. We observe that the random forest models quickly learn to become very accurate. We note that this is not a case of overfitting; all experiments used half of the samples for training, and half for evaluation. Indeed, the nature of our sampling methodology means that the sample space of possible queries is extraordinarily large. With $R^2 \to 1$ for many mapping functions, we observe the power of local learning, where the models quickly pick up on router characteristics.

5.3 LEARNING FILTERING

The method for learning filtering (that is, the number of objects that will survive a query) is very similar to that described above. However, here we explicitly use the parameter values of the queries as inputs to our random forest models. In addition to the relationships between the complex and basic operation tags, the system also keeps track of how many graph objects survive each filtering operation. As in the learning of $\overline{c}_{ij}$, we quickly obtain $R^2 \to 1$. Of course, we expect that the difficulty of learning the filtering mapping will increase as the number of variables that we perform filtering on increases; future work will investigate this potential issue. Nonetheless, as demonstrated by Figure 3, in this example the random forests are capable of learning the filtering mapping approximation to a very high degree of accuracy, given sufficient samples.

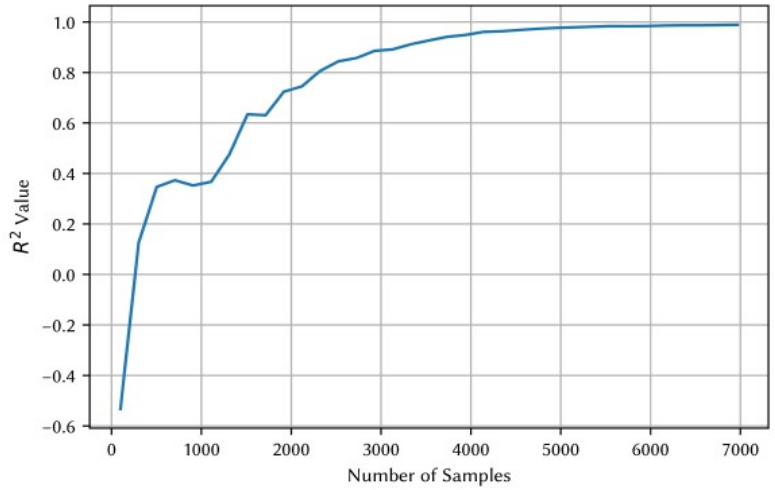


Figure 3. R^2 for learning vertex output and router locations given list of vertices and filtering conditions

5.4 THE REPRESENTATION PROBLEM

Finally we use all of the above methods to solve the representation problem. This problem involved sending 100 random two-step queries (that is, each query had an output that had to be fed into the next stage) to each of the four routers (as described by Table 2). We ran the queries in each of the sixteen possible sets of representation choices (in order to demonstrate the fully optimal solution), and then fed the query specifications to our learned models of $\overline{c}_{ij}, \overline{f}_j, \overline{h}$. We observed very encouraging results (as seen in Figure 3); using our machine-learning models gave us representations that on average were only ~14% slower than optimal, with the average router representation being more than 13 times slower than the optimal. As optimal representations run the gamut from all-MAL and all-RDF to everything in between, the value of an approach that gets a near optimal solution is particularly important given that the consequences of incorrect assignment can be so drastic. Note that finding the true optimal solution requires running the set of jobs with every possible combination of router assignments. As mentioned previously, this requires an exponential number of router representation tests. As the size of the jobs increases, it

becomes infeasible to look for an optimal representation through brute force (and doing so is redundant anyway, as completing the job is the purpose of the execution on the database). In contrast, the SmartGraph method will scale with the machine learning models used to learn $\overline{c}_{ij}, \overline{f}_j, \overline{h}$, and in many instances will not need to be retrained at all (e.g. if we keep the same graph and move from two-step to three-step traversal queries).

6. DISCUSSION AND FUTURE WORK

In this work we demonstrated a new methodology for constructing graph databases, and used the representation problem to demonstrate how giving artificial intelligence to "router-like" subgraphs can be used to solve highly complex problems. We showed that the representation problem is an exponentially large combinatorial optimization problem, but that it can be solved to near-optimality by having each router learn machine-learning model representations of

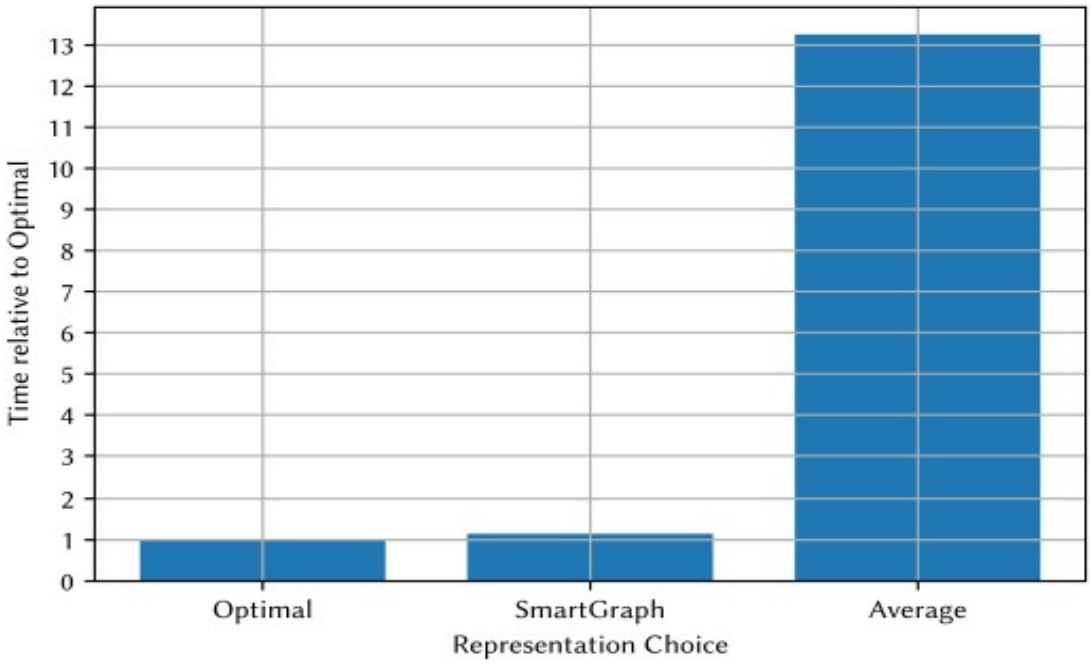


Figure 4. Representation Problem on our Test Cases

themselves and their neighboring routers. There is clearly more to explore in the representation problem alone. For example, solving the problem on real-world graphs (and investigating if scale-free properties are problematic for the approach), solving the problem with the inclusion of write methods, investigating how changing the number and size of routers influences the learning capability of each router, and exploring joint-learning with the graph partitioning problem. In addition, the SmartGraph itself is rife with potential for further research. For example, in investigating how the routers can facilitate completely asynchronous querying capabilities, the real-world performance metrics of SmartGraphs with an enormous number of routers, and more. By adding artificial intelligence at the "edge" of graph database technology, we increase the computing capabilities of the graph database, and continue to muddle the space between graph computing and graph databases.

REFERENCES

[1] R. R. McCune, T. Weninger, and G. Madey, “Thinking Like a Vertex,” ACM Comput. Surv., 2015.

[2] G. Malewicz et al., “Pregel,” Proc. 28th ACM Symp. Princ. Distrib. Comput. - Pod. ’09, p. 6, 2009.

[3] L. G. Valiant, “A Bridging Model for Parallel Computation,” Commun. ACM, vol. 33, no. 8, pp. 103–111, Aug. 1990.

[4] V. Kalavri, S. Ewen, K. Tzoumas, V. Vlassov, V. Markl, and S. Haridi, “Asymmetry in Large-Scale Graph Analysis, Explained,” in Workshop on Graph Data Management Experiences and Systems, 2014.

[5] Y. Tian, A. Balmin, S. A. Corsten, S. Tatikonda, and J. McPherson, “From ‘think like a vertex’ to ‘think like a graph,’” Proc. VLDB Endow., 2013.

[6] Y. Simmhan et al., “GoFFish: A sub-graph centric framework for large-scale graph analytics,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, vol. 8632 LNCS, pp. 451–462.

[7] D. Yan, J. Cheng, Y. Lu, and W. Ng, “Blogel: A Block-centric Framework for Distributed Computation on Real-world Graphs,” in Proc. VLDB Endow., 2014.

[8] A. Donovan and B. W. Kernighan, The Go programming language. Addison-Wesley Professional, 2015.

[9] A. Buluç, H. Meyerhenke, I. Safro, P. Sanders, and C. Schulz, “Recent Advances in Graph Partitioning,” pp. 1–37.

[10] B. Schiller, C. Deusser, J. Castrillon, and T. Strufe, “Compile- and run-time approaches for the selection of efficient data structures for dynamic graph analysis,” Appl. Netw. Sci., pp. 1–22, 2016.

[11] A. Kusum, I. Neamtiu, and R. Gupta, “Safe and flexible adaptation via alternate data structure representations,” in Proceedings of the 25th International Conference on Compiler Construction - CC 2016, 2016.

[12] C. Jung, S. Rus, B. Railing, N. Clark, and S. Pande, “Brainy: effective selection of data structures,” in PLDI ’11:Proceedings of the 32nd ACM SIGPLAN Conference on Programming Language Design and Implementation, 2011.

[13] O. Kennedy and L. Ziarek, “Just-In-Time Data Structures,” in Onward!, 2015.

[14] R. S. Xin, J. E. Gonzalez, M. J. Franklin, I. Stoica, and E. AMPLab, “GraphX: A Resilient Distributed Graph System on Spark,” First Int. Work. Graph Data Manag. Exp. Syst., p. 2, 2013.

[15] A. Dave, A. Jindal, L. Li, R. Xin, J. Gonzalez, and M. Zaharia, “GraphFrames: An Integrated API for Mixing Graph and Relational Queries,” in Proceedings of the Fourth International Workshop on Graph Data Management Experiences and Systems SE - GRADES ’16, 2016.

[16] J. Webber, “A Programmatic Introduction to Neo4J,” Proc. 3rd Annu. Conf. Syst. Program. Appl. Softw. Humanit., pp. 217–218, 2012.

[17] O. Green and D. A. Bader, "cuSTINGER: Supporting dynamic graph algorithms for GPUs," 2016 IEEE High Perform. Extrem. Comput. Conf. HPEC 2016, 2016.

[18] S. Che, B. M. Beckmann, and S. K. Reinhardt, "BelRed: Constructing GPGPU graph applications with software building blocks," 2014 IEEE High Perform. Extrem. Comput. Conf. HPEC 2014, 2014.

[19] P. Barceló, L. Libkin, A. W. Lin, and P. T. Wood, "Expressive Languages for Path Queries over Graph-Structured Data," ACM Trans. Database Syst., vol. 37, no. 4, p. 31:1--31:46, 2012.

[20] R. Angles, M. Arenas, P. Barceló, A. Hogan, J. Reutter, and D. Vrgoč, "Foundations of modern query languages for graph databases," ACM Comput. Surv., vol. 50, no. 5, p. 68, 2017.

[21] L. Libkin, W. Martens, and D. Vrgoč, "Querying Graphs with Data," J. ACM, 2016.

[22] F. Holzschuher and R. Peinl, "Performance of graph query languages: comparison of Cypher, Gremlin and native access in Neo4j," 16th Int. Conf. Extending Database Technol. EDBT' 13, no. March 2013, pp. 195–204, 2013.

[23] B. Gallagher, "Matching Structure and Semantics : A Survey on Graph-Based Pattern Matching," AAAI FS, 2006.

[24] B. Schiller, J. Castrillon, and T. Strufe, "Efficient data structures for dynamic graph analysis," in 11th International Conference on Signal-Image Technology & Internet-Based Systems, 2015.

[25] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[26] P. Erdos and A. Rényi, "On the evolution of random graphs," Publ. Math. Inst. Hung. Acad. Sci, vol. 5, no. 1, pp. 17–60, 1960.

AUTHORS

Hal Cooper (left) is a doctoral candidate in the Department of Industrial Engineering and Operations Research at Columbia University, where he is advised by Professor Garud Iyengar (right),who also serves as the Department Chair. Ching-Yung Lin (bottom) is an Adjunct Professor of the Department of Electrical Engineering at Columbia University. Ching-Yung is also the CEO of Graphen, Inc., and a former Chief Scientist for Graph Computing at IBM.

INTENTIONAL BLANK

NETWORK SECURITY ARCHITECTURE AND APPLICATIONS BASED ON CONTEXT-AWARE SECURITY

Hoon Ko[1], Chang Choi[2], Pankoo Kim[3] and Junho Choi[4]

[1]IT Research Institute, Chosun University, Gwangju, South Korea
[2]IT Research Institute, Chosun University, Gwangju, South Korea
[3]Department of Computer Engineering, Chosun University, Gwangju, South Korea
[4]Division of Undeclared Majors, Chosun University, Gwangju, South Korea

ABSTRACT

The number of services and smart devices which require context is increasing, and there is a clear need for new security policies which provide security that is convenient and flexible for the user. In particular, there is an urgent need for new security policies regarding IT vulnerability layers for children, the elderly, and the disabled who experience many difficulties using current security technology. For a convenient and flexible security policy, it is necessary to collect and analyze data such as user service use patterns, locations, etc., which can be used to distinguish attack contexts and define a security service provision technology which is suitable to the user. This study has designed a user context-aware network security architecture which reflects the aforementioned requirements, collected user context-aware data, studied a user context analysis platform, and studied and analyzed context-aware security applications.

KEYWORDS

Context-aware Security, Network Security Policy, Malicious Code Detection

1. INTRODUCTION

As the internet of things develops and becomes commercialized, security threats targeting a variety of mobile smart devices are increasing. Among the threats currently being introduced are attacks which can install malicious code on washing machines and refrigerators, attacks which allow hackers to illegally control automobiles remotely, and security vulnerabilities in medical devices such as pacemakers. However, the security solutions for responding to these threats have not taken user convenience into account, and users are inconvenienced by them as a result. It is believed that in the future all devices will connect to networks at high speeds, and the problem of network security will become more important. Children, the elderly, and the disabled in particular must be considered, as they experience difficulties in understanding and using current security solutions. Current security studies are being conducted toward the end of increasing user convenience while maintaining system security. The goal is to resolve the underlying problem by accurately understanding the principles related to security threats and providing the security factors which current solutions omit or fail to satisfy. The user context-aware network security

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 79–90, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90308

solution proposed in this study is a security framework which can simultaneously provide convenience and security to the normal user. It collects information which can determine the user's context such as user location information, service use patterns, etc., and it automatically provides the most appropriate security service to the user. This paper describes the user context-aware technology that is needed to collect data from the user device's sensors and nearby communications devices and becomes aware of the user context through machine learning and real-time big data analysis technology in order to provide optimized security services.

2. USER CONTEXT-AWARE NETWORK SECURITY TECHNOLOGY TRENDS

2. 1 CONTEXT DEFINITION

Context describes the situation that an entity is experiencing, and context can be collected to analyze a scenario. There are various kinds of context such as people, places, and things. Context is categorized as physical or logical according to its nature, as shown in Table 1.

Table 1. Context Categories.

Category	Physical Context Data	Logical Context Data
Acquisition Method	Acquired through sensors; sensors observe status through measurements	Acquired through record data and activity situation
Processing	For example, magnetic sensors measure a magnetic field's strength and direction, and inertia sensors measure angular displacement and changes in angular displacement	For example, the data indicates whether a person is performing an activity or resting, or whether the person is talking on the phone or sending a message, etc.

2.2 DATA COLLECTION

The types of data which can be collected by a device are very diverse, but the problems with collected data such as integrity, availability, user privacy, etc. are also very diverse. Table 2 shows recent studies related to data collection.

Table 2. Data Collection.

Category	People-centric Opportunistic Sensing [1]	The algorithmic foundations of differential privacy [2]	Continuous user authentication on mobile devices [3]	Soft authentication with low-cost signatures [4].	Joint segmentation and activity discovery using semantic and temporal priors [5]
Feature	When the results of devices performing sensing are accessed by Wi-Fi, they	Device's own security and user privacy	Use of wrist motion sensor's accelerometer (hallmark usage)	Comparison of behavior patterns of registered smartphone owner and illegal user	Uses non-parametric integrated model, divides time into specific segments and

	are sent to the database server and data is collected				uses time points where activities are changing
Advantages	Data collection is possible even though separate infrastructure is not allocated	Strengthens security of Android market	No need for passwords, fingerprints or RFID tags	Higher accuracy than existing methods when noise of sensor results is below a fixed level	Flexible segment size possible by specifying one super sample
Disadvantages	Problems related to participants' privacy, data integrity, and usability	User information leaks when requesting permission to access specific information on the devices during app installation	Accuracy increases only if the time when the user uses the object and the time when the wrist motion sensor collects the data are the same	There is still battery consumption, and usage range (coverage) is limited	Higher accuracy than existing methods when noise of sensor results is below a fixed level

A. SECURITY INFORMATION / EVENT MANAGEMENT

Systematic measures are necessary to perform early detection and respond to the various security threats that use advanced/large-scale network infrastructure. It is also necessary to have a security control system which collects large amounts of event data such as logs, packets, etc. generated by various security devices, network infrastructure, server/storage devices, and service applications, and which uses big data solutions to perform security analysis. Security Information and Event Management (SIEM) is a commercially-oriented solution which collects virtual/actual networks, service applications, system logs and event data. It then categorizes these and analyzes them to create quick reports, and it gives warnings if an additional intervention or altered response is needed [6][7][8][9]. The security tools provided by SIEM perform the core role of a security operations center (SOC) which performs central duties related to security in an organization or business' IT structure. It is also used to record security logs and generate compliance reports [9].

B. TRENDS IN SECURITY COLLECTION/ANALYSIS TECHNOLOGY USING THE CLOUD

In Security as a Service (SESaaS), the security service provider takes responsibility, and services are provided for authentication, anti-virus, anti-malware/spy-ware, intrusion detection, and security event management [10]. In cloud computing, this is defined as Cloud Security as a Service or SECaaS in which the cloud provider (CP) provides security in the form of SESaaS. The services are categorized as shown in Figure 1. The provided services include identity and access management, data loss prevention, web security, email security, security assessments, intrusion management, security information and event management, encryption, business continuity and disaster recovery, and network security.

Figure 1. SECaaS WG defined 10 categories

3. USER CONTEXT-AWARE NETWORK SECURITY STRUCTURE DESIGN

3.1. CONSIDERATIONS DURING ARCHITECTURE DESIGN

Recently, there has been an increase in the device networks which need authentication in the IoT environment as well as the range of performance and usage environments for these devices. Because of this, it must be possible to use existing authentication methods in the proposed user context-aware network structure. The network security structure design considerations are shown in Table 3.

Table 3. User context-aware network structure design considerations.

Considerations	Features
User usability Issue	Because smart devices are always carried by the user as they move around, usability must be considered more for cases in which authentication occurs often [11] In practice, user accessibility is decreased if errors such as false positives and false negatives occur often or if it takes a long time to perform authentication and use the device
Capacity issue of Mobile devices	If there is no interface which provides knowledge-based or biometric-based authentication methods due to the device characteristics: It is necessary to entrust security related features to an external server such as a hub, gateway, or cloud There can be a problem with having to completely trust the entity which performs this role, and higher security such as additional authentication is required

Privacy issue	There is a possibility that GPS information which user data gives can expose the user's location[12] User call logs describe the user's personal relationships. It is necessary to consider the privacy of users which have information

3.2. DATA COLLECTION ARCHITECTURE

The device's sensors are used, and logs such as the user's contact information and text information are collected, and a behavior fingerprint is defined for each. For example, if the user is a student, the student's movement pattern may take a regular form such as moving between home and school. If the user accepts a phone call from an unregistered number, the probability that the number will be accepted in the future increases. If the user does not accept the call, the probability that it will not be accepted in the future increases. These user behavior patterns are used in the creation of a behavior fingerprint (BFP). BFPs which are created this way are called attributes. The accurate creation of attributes is more important than anything in increasing the accuracy of user authentication. This is because external factors such as the smartphone's location, user's behavior, current weather, temperature/humidity, etc. create noise, and this reduces accuracy when distinguishing the user. Therefore, in order to distinguish users by using attributes, confidence must be increased continually by combining nearby attributes and renewing them.

```
// Data Collection Algorithm. The attribute is a value created by the BFP. The
confidence analysis is the accuracy analysis stage. The confidence set is the value
created after the confidence analysis. lpis describes the least privacy invasion set.

c ← Collect(attribute1+attribute2+…+attributen);
a ← Analysis(c);
s ← Confidence.set(a);
p ← privacy.analysis(s);
r ← result(accuracy(p);
attributei ← lpis (r);

SEND attributei TO Server;

If attributei == DB(attributei)
    ACCEPT (authentication);
OR
    RECOMMEND(OtherAuthentication);
    If matched {
      DECREASE(Confidence.RATE(attributei));
      UPDATE(Confidence) TILL matched;
      p ← privacy.analysis(s) IN External Server;
      FEEDBACK(s, r);
    }
```

In the proposed architecture's processing order, the data is first collected according to the attribute, and then the confidence analysis is performed. The confidence analysis is a module which calculates whether a certain combination of attributes is necessary to guarantee confidence over a certain threshold value. Based on these calculations, the confidence set is created and analyzed. The privacy analysis analyzes the degree of privacy invasion for each attribute. The least privacy invasion set is selected, and the set's attributes are transmitted to an external server.

The external server provides feedback on the confidence analysis and the privacy analysis based on the analyzed results. If the comparison with the fingerprint that the user has already saved does not produce a match, a different authentication method is proposed. If the different authentication method succeeds, the attribute's confidence is lowered, and the confidence is continually updated (FEEDBACK Step). In FEEDBACK, if privacy continually brings in the same result, similar attributes are continually sent out. However, because it is easy to violate privacy, the external server's feedback results are reflected in the privacy analysis to prevent transmission of the same attribute.

3.3. CONTEXT-AWARE MULTI-FACTOR AUTHENTICATION

Multi-factor authentication requires different types of information on the authentication target in order to control access to a resource, such as information that the user knows and a possession that the user has [13]. This authentication method is already widely used. For example, a transaction with a bank's ATM card requires both knowledge (password -- information the user knows) and the ATM card (possession -- thing the user has) to confirm the user. Because 2 pieces of information are required, this corresponds to a type of multi-factor authentication known as two-factor authentication. If the user's context is used as one piece of authentication information in multi-factor authentication, the level of security can be maintained while also ensuring convenience. Information acquired as the result of soft sensing is used to determine the user context information. This method can confirm whether or not the device or environment which is to be authenticated matches the context experienced by the user, and the complexity of the authentication stage is only alleviated if it does match. If this method is used, an additional two-factor or three-factor authentication stage is required if the user context information does not match. By providing such a stage, a higher level of authentication security can be achieved. If the user context information does match, the authentication stages can be simplified to increase user convenience.

4. USER CONTEXT-AWARE NETWORK SECURITY STRUCTURE DESIGN

4.1. Definition of Data

A. COLLECTION DEVICES

In the IoT era, the number of sensor types and the places where they are used have both increased. Devices that users carry and connected sensors can extract features from the user's activities, behavior, etc. There are convenient and effective authentication methods which calculate a risk score and make requests for a suitable authentication approach according to the risk. Types of collection devices include smart phones, PCs, tablet PCs, wearable devices, smart sensors / hubs, smart door locks, smart TVs with IoT hubs, and other connected devices.

B. COLLECTABLE DATA

Table 4 shows the collectable data. For example, a smartphone that a user carries or a wearable device can detect the user's location, activities, and characteristics through physical sensors (hard sensing) such as accelerometers, magnetometers, and gyroscopes. Effective authentication methods which are suitable for the user context can be requested through logical sensors (soft sensing) which detect screen state of devices, battery consumption, phone records, data usage, etc.

Table 4. Collectable data.

Device	Sensors (Physical/ Logical)	Features
Smartphone, PC and tablets/ Wearables	- Phone Info/Calls - Location, Accelerometer - Magnetometer	User's current location and activities
Smart Sensors/ Hub	- Motion, Light	User's environments and activities
Smart Door Lock	- Location	User at home or not
Smart TV with IoT Hub	-Usage time -User channel properties, etc.	Duration, User's personal interest

4.2. SECURITY/PRIVACY ISSUES

Context- aware opportunistic user authentication systems collect biometric contexts through the user's smart device and are able to understand the user's current context. By doing so, they have the compelling benefit of providing appropriate security solutions. However, by collecting this data, problems can occur which are related to user privacy and collected data integrity and usability. Also, as the use of smart devices increases, malicious apps may be offered on app markets. Even when an app is not malicious, mistakes by a developer who does not take security into account may cause problems in which users accidentally encounter other malicious programs through advertisements, etc. and expose their personal information [10]. If context- aware opportunistic user authentication is being used, there is a risk that the device may become a privacy vulnerability point (privacy hole) and even personal information stored on other connected devices could be leaked [14].

4.3. COLLECTION TECHNOLOGY

Data collection for user context- awareness is categorized as physical sensor data which can be directly obtained from the devices that the user is operating, logical sensor data which changes according to the user's habits and device usage patterns, and soft sensing data in which new data is derived by using the data of several sensors [15][16]. Each type of sensor data can be collected continuously like GPS data or collected intermittently like phone conversation data. In short, there are many types of data for context awareness, and they have characteristics that make it difficult to integrate collection methods. As such, context-aware data requires analysis and processing before inferences which utilize the data can be made. Depending who is in charge of processing, this analysis includes (a) methods which directly process sensed data on the user device and (b) methods in which the user device sends data to a server and it is processed. The features of each method are shown in Table 5.

Table 5. Collection technology types.

Type	Advantages	Disadvantages
(a) Data is processed directly on the user device	After processing, the size of the data is reduced so that storing it is easy and sending it to the server is easy.	The user device cannot guarantee sufficient performance and it is difficult to update the algorithms needed for data processing
(b) Data is sent to a server and processed	It is possible to guarantee sufficient performance to process the collected data. It is easy to change the data characteristics or analysis standards and update the algorithm.	The collected data must all be sent to the server, which may put a burden on the user device or server depending on the situation.

4.4. COLLECTION TECHNOLOGY

Figure 2 shows the proposed data collection architecture. Confidence analysis analyzes the user fingerprint accuracy for each attribute. It can analyze 1 attribute or analyze the accuracy of a set of several attributes simultaneously. The accuracy analysis is first done on the basis of reference data, and later machine learning is used to continue learning about accuracy analysis. At first, an accuracy above a fixed level is seen. However, because the accuracy for each attribute can vary due to flaws in the sensors themselves or changes in user patterns, the accuracy is updated by setting a score for the accuracy based on the authentication results.

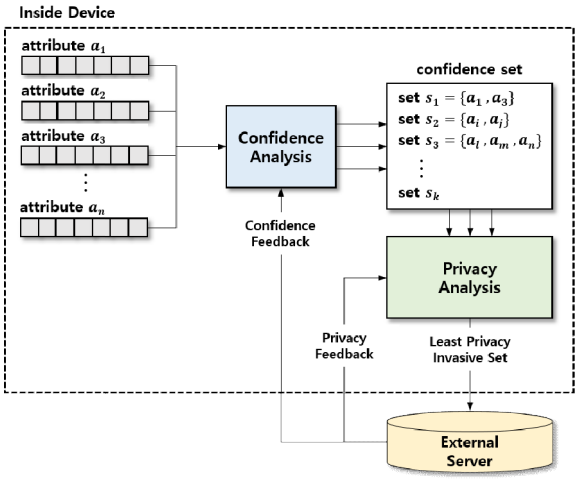


Figure 2. Data collection architecture

When confidence analysis is finished, a set of attributes which has an accuracy above a fixed threshold value is produced. The threshold value can be selected experimentally, and it can be updated later based on the authentication results. The set of attributes which is above the threshold value is called the confidence set. Privacy analysis is performed for a total of k confidence sets.

Privacy analysis shows the privacy given by the attributes in the confidence set. The user's sensitivity to the privacy of each attribute varies. Because the analysis of sensitivity to privacy varies by each user, different levels of sensitivity are used for each user. Providing high accuracy means distinguishing the user more easily, and this ultimately means showing the user's personal information more accurately. When the privacy analysis is finished, one set with the lowest sensitivity, i.e. the set with the least user privacy leaking, is selected. The selected set is defined as the least privacy invasive set, and it is transmitted to the external server for the next stage.

5. APPLICATIONS

This section defines 5 context-aware applications based on the previously defined context-aware network security architecture and context-aware data collection method. These applications are context-aware authentication technology, context-aware access control technology, a context-aware personalized warning system, context-aware security settings, and a malicious code detection system. Table 6 describes each application.

Table 6. Applications.

Category	Goal	Processing	Features
Context-aware authentication technology	Provides a suitable user context authentication process in which the service user is pre-enrolled to provide context-aware service to the user	- Collects data from acceleration sensor, microphone, GPS location service, and touch screen for authentication - The acceleration sensor measures the user's movement state and speed. The microphone receives the user's voice. - The GPS can accurately know the user's location. The touch screen collects the user's touch behavior. The collected information is used to perform authentication	- Authentication using the user's movement state - Authentication using the user's voice - Authentication using the user's GPS - Authentication using the user's touch records
Context-aware access control technology	Determines the access control which is suitable for the context based on context data acquired by using a variety of sensors, GPS, etc.	- The Profiler module extracts features from collected raw data and defines objects related to context models and profiles - The Classifier module uses the extracted features to train the context classification model - Transmits the "classification" and the "confidence value" which shows the degree of risk in the context model to the Access Control Layer - The Access Control Layer module determines access control based on the received data	- Knows outdoor places that the user visits often - Knows indoor places that the user visits often - Integrates GPS data with interactions between the user and the device UI and performs user location analysis and current user status analysis
Context-aware personalized warning system	Provides custom warnings according to the user context and effectively conveys the warning content to the user	- Creates different warnings to transmit to the user depending on the context analysis results - Creates customized warnings which can effectively convey warning content to the user based on each context such as the level of specialized knowledge possessed by the user, the user's current status (working, sleeping, resting, etc.), level of risk for abnormal contexts, etc.	- Warnings according to the user's system specialty - Warnings according to the user's status - Warnings according to the warning context's risk level - Bot detection warning: If the user is not a person, distinguishes the user as a bot and creates a captcha
Context-aware security settings	Uses the user's context-aware data to resolve the hassle of having to go through a complex security settings process and provides a dynamic and simple security settings method	- Collects the user's context-aware data and uses it to dynamically set the security settings level of the devices used by the user according to the situation - The security settings levels are defined as normal, high, and user-specified - The context that is used to change the security settings utilizes information such as the degree to which the user understands security-related information, the user's current status, the user's activity data, and the user's location data	- Security settings using user status data - Security settings using user location data - Security settings using user activity data
Malicious code detection system	Provides malicious code detection system which is customized according to the user status to establish measures for systematically responding to and preventing malicious code attacks	- Goes through static analysis, dynamic analysis, and trace analysis and categorizes the features which show the execution file as classes - Determines groups which are estimated by the relationship with the target execution file via class categorization stage - Categorizes the developer's classes based on the data extracted from the execution file and infers the developed and circulated execution file from the specific structure	- Malicious code detection through static analysis - Malicious code detection through dynamic analysis - Malicious code detection through usage tracking

6. CONCLUSIONS

To implement user context-aware network security technology, it must be possible to determine contexts by utilizing user usage patterns and movement patterns which are based on user position data. To utilize safe user context awareness, data must be collected from the user device's sensors and nearby network devices, and the user's nearby context must be recognized by using machine learning and real-time big data analysis technology to provide optimized security service. In order to implement a safe user context-aware service, this study examined trends in context-aware network security technology and designed a user context-aware network security architecture. The trends in user context-aware network security technology were organized into data collection in which the collected data's integrity, usability, and user privacy are processed, context analysis which is an analysis stage for performing various analyses on the collected data, and context-based applications which are for safely using the analyzed contexts. In the user context-aware network security architecture design, the context information analysis and the user's nearby context, which have been collected by the cloud server, were judged together to measure risk, and based on these results, security services which can deliver both user satisfaction and security were provided. In the results of the study, five kinds of applications were introduced, and the safety of the user context-aware applications were analyzed. Based on the results of this study, a variety of safe user context-aware applications are expected in the future.

ACKNOWLEDGEMENTS

This research was supported by Korea Electric Power Corporation (Grant number: R18XA06-12) and the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP)(No.NRF-2016R1A2B4012638).

REFERENCES

[1] Berker Agir, Jean-Paul Calbimonte and Karl Aberer, "Semantic and Sensitivity Aware Location Privacy Protection for the Internet of Things," PrivOn'14 Proceedings of the 2nd International Conference on Society, Privacy and the Semantic Web - Policy and Technology, Vol. 1316, pp. 58-63.

[2] Dwork, Cynthia, and Aaron Roth. "The algorithmic foundations of differential privacy," Foundations and Trends in Theoretical Computer Science 9.3-4 (2014): 211-407.

[3] Vishal M. Patel, Rama Chellappa, Deepak Chandra, and Brandon Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," IEEE Signal Processing Magazine 33.4 (2016): 49-61.

[4] Buthpitiya, Senaka, Anind K. Dey, and Martin Griss. "Soft authentication with low-cost signatures," Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on. IEEE, 2014.

[5] Seiter, Julia, et al. "Joint segmentation and activity discovery using semantic and temporal priors," Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on. IEEE, 2015.

[6] "SIEM: A Market Snapshot,"Dr. Dobb's Journal, Feb. 2007.

[7] J. Hayes, "Cybersecurity and the Big Yellow Elephant," Cloudera Vision Blog, May 2015.

[8] K. M. Kavanagh, O. Rochford, and T. Bussa, “Magic Quadrant for Security Information and Event Management,” Gartner, Aug. 2016.

[9] Bhatt, P. K. Manadhata, and L. Zomlot, “The operational role of security information and event management systems,” IEEE Security & Privacy, vol. 12, no. 5, 2014.

[10] Mosaic Security Research, “Log Management & Security Information and Event Management (SIEM) Software Guide,” Mosaic Security Research, (accessed May 2014).

[11] G. Abowd and A. Dey, "Towards a better understanding of context and context-Awareness," In International Symposium on Handheld and Ubiquitous Computing, pp. 304-307, 1999.

[12] H. Witte, C. Rathgeb and C. Busch, "Context-Aware Mobile Biometric Authentication based on Support Vector Machines," 2013 Fourth International Conference on Emerging Security Technologies, Cambridge, 2013, pp. 29-32.

[13] William E. Burr, Donna F. Dodson, Elaine M. Newton, Ray A. Perlner, W. Timothy Polk, Sarbari Gupta, and Emad A. Nabbus, “Electronic Authentication Guide,” Special Publication 800-63-2, NIST, 2013.

[14] Kapadia, Apu, David Kotz, and Nikos Triandopoulos. "Opportunistic sensing: Security challenges for the new paradigm." 2009 First International Communication Systems and Networks and Workshops. IEEE, 2009.

[15] H. Witte, C. Rathgeb and C. Busch, "Context-Aware Mobile Biometric Authentication based on Support Vector Machines," 2013 Fourth International Conference on Emerging Security Technologies, Cambridge, 2013, pp. 29-32.

[16] T. Gisby, “”Soft” Sensors Are Breaking Into Four Major Industries,” Aug 2015.

[17] Harbach, Marian, et al. "Sorry, I don’t get it: An analysis of warning message texts." International Conference on Financial Cryptography and Data Security. Springer Berlin Heidelberg, 2013.

[18] Anderson, Bonnie Brinton, et al. "How polymorphic warnings reduce habituation in the brain: Insights from an fMRI study." Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, 2015.

AUTHORS

Hoon Ko received B.S. at Howon University in 1998, and M.S. in 2000, Ph.D. in 2004 at Soongsil University, S. Korea. He had worked at GECAD/ISEP/IPP in Portugal as a Post-Doc from 2008 to 2013, and at J. E. Purkinje University, Czech Republic as a research professor from June 2013 to July 2015. Now he been working at IT Research Institute, Chosun University, S. Korea as a research professor since Oct 2017. His research area is CPS Security, Cyber-Security, context-aware, multicast, IoT, Bio-information security.

Chang Choi received his Ph.D. degree in Computer Engineering from Chosun University in 2012, and is now working as a research professor at the Chosun University, South Korea, and is a Senior member of the IEEE. His research interests include Intelligent Information Processing, Semantic Web, Smart IoT System and Intelligent System Security.

Pankoo Kim received his B.E. degree from the Chosun University in 1988 and M.S. and Ph.D. degrees in Ccomputer Engineering from Seoul National University in 1990 and 1994. Currently, He is now working as a full professor at Chosun University. He is an EIC of IT CoNvergence PRActice Journal. His specific interests include semantic web techniques, semantic information processing and retrieval, multimedia processing, semantic web and system security.

JunHo Choi received the PhD degree from Chosun University, Gwangju, Korea in 2004. He is currently an Assistant Professor with the Department of Undeclared Majors, Chosun University, Gwangju, Korea. His research interests include Security, Cloud Computing, Semantic Analysis, Ontology Modeling.

SIMPLIFICATION OF COMPILER DESIGN COURSE TEACHING USING CONCEPT MAPS

Venkatesan Subramanian[1] and Kalaivany Natarajan[2]

[1]Department of Information Technology, Indian Institute of Information Technology, Allahabad, India.
[2]Australia

ABSTRACT

The students of undergraduate expect simple, interactive and understandable method of teaching. But, in many universities, instructors simply follow text books and solve the examples given in it. In such case, students cannot learn anything other than the text book contents defined in the syllabus. This method of teaching will not motivate students to understand the subject in depth unless the industrial and practical applications of the subject explained. Once students do not show interest, instructor also loses the interest which leads to poor performance of students. Hence, it is mandatory to re-design the teaching method especially for the undergraduate students. This paper use concept maps to simplify teaching and learning of a compiler design subject with assignments and problems related to research and industry.

KEYWORDS

Compiler Design, Concept Map, Computer Science, Education

1. INTRODUCTION

The IT stalwarts are citing that most of the engineering graduates are not fit for hiring. The major reason for such issue is, while studying the undergraduate engineering, students does not understand subjects in depth. Students simply read the important topics of the subjects' from text books and pass the exam. Some do solve questions based on the same text book. We but can categorize the students in to four based on their expectation from the subject study as shown in figure 1.

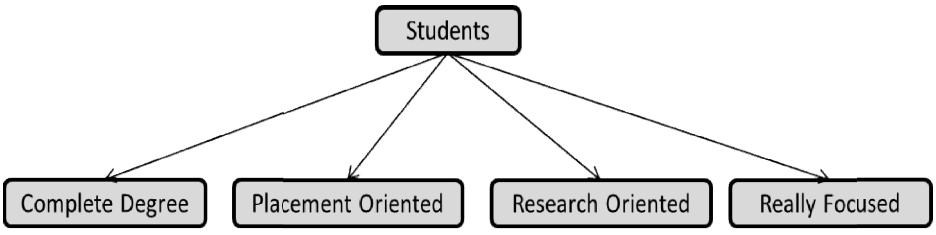


Figure 1. Categorization of Students

The first category of students studies the subject just to pass and get the degree. Second category of students concentrates only on the subjects those helps in industry placement and they consider other subjects are the one just to complete the degree. Third category of students interested to do research will concentrate on subjects, which they feel it is required for research. Fourth category of students really focuses on all subjects and study thoroughly. Usually, students of first and last

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 91–102, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90309

category will be very less and other two categories will have more students especially the second category that is placement oriented. All first three categories of students will realize the issue of not concentrating on all subjects while they do job or research because subjects are linked with each other and required while solving research or industry problems. However, it is very late to realize. Hence, students should concentrate on all subjects while they study. It can be achieved only when students are motivated towards the subject and the instructor change the traditional text book based teaching method. To understand the students' objectives towards each subject of the course, we conducted the survey with the following three questions.

a) What are the subjects you are interested in?
b) What are the subjects you feel required for placement?
c) What are the subjects you feel required for research?

There are 94 third year undergraduate students participated in the survey. The survey results are shown in table 1. The Interested column shows the percentage of students having interest towards the subject. The Placement and Research column shows the percentage of students feel that the subjects are required respectively to get industry placement and do research.

Table 1. Survey Results of students opinion on subjects (in percentage)

	Interested	Placement	Research
Introduction to Programming	52	45	2
Data Structure	95	81	17
Algorithms	95	72	36
Database Management System	86	65	10
Automata	11	29	33
Compiler Design	14	14	19
Software Engineering	21	14	1
Operating System	82	44	32
Computer Architecture	26	19	26
Programming Principles	16	12	11
Networks	52	37	9
Artificial Intelligence	20	48	83

We can understand from the survey results that students are not having enough interest on subjects such as Automata, Software Engineering, Computer Architecture and Compiler Design. But an experienced technologists and teacher knows the importance of these subjects. Hence, it is necessary to encourage and motivate students for these subjects by changing the teaching method. In this paper, we concentrate on Compiler Design course, which is considered by all categories of students as one of the complex and very less productive. Students keen to study Programming are not showing interest towards Compiler Design. They feel that Introduction to Programming helps in understanding programming language and writing programs but Compiler Design is not having any productive outcome. We need to beat this notion by highlighting important of compiler design course. Hence, it is essential to re-design the teaching style of Compiler Design course to motivate the students. It is possible when we have the proper lesson and concept plan in the beginning.

In this paper, we propose concept maps for Compiler Design course by the innovative ideas [2 and 3]. The concept map is discussed in overall and phase wise in the paper. We have listed core concepts of compiler in each phase and provided the problems to be discussed in the class room and assignments for students as homework.

The remaining paper is organized as follows: Section 2 discusses the existing works, Section 3 briefly discusses the overall concept map of Compiler Design course, Section 4 discusses about the phases of compiler with concept map. Section 5, has the discussion of teaching method. Finally, Section 6 concludes the paper with the directions of future work.

2. LITERATURE REVIEW

In the past, different approaches were proposed to motivate the students to concentrate on the compiler design course. Akim Demaille et al [12] introduced set of compiler construction tools for educational projects. This makes learning and evaluation easy. Li Xu and Fred G. Martin [13] proposed a Chirp system that provides a realistic and engaging environment to teach compiler course. However, concepts and its relationship may not reach the students. Tyson R. Henry [14] proposed a Game Programming Language (GPL) based teaching to motivate students towards the compiler project. Elizabeth White et al. [15] proposed an approach that enables students of compiler course to examine and experiment with a real compiler without becoming overwhelmed by complexity. S. R. Vegdahl [16] proposed visualization tool to teach compiler design to bring the interest of students on the subject. Marjan Mernik and Viljem Zumer [17] proposed a software tool called LISA for learning and conceptual understanding of compiler construction in an efficient, direct, and long lasting way. Henry D. Shapiro and M. Dennis Mickunas [18] replaced the term project on compiler design with several smaller, independent, programming assignments for better understanding and to motivate students. Martin Ruckert [19] argues that teaching compiler with unusual programming language is a good choice. Divya Kundra and Ashish Sureka [20] [21] discussed about case based teaching and learning of compiler design and proposed case studies for different concepts to make learning easier and interesting. Somya Sangal et al [22] proposed a Parsing Algorithms Visualization Tool (PAVT) tool to teach the process of parsing algorithms. Learners can visualise the intermediate steps of the parsing algorithm through the tool.

Even though game based, case based and tool based teaching makes learning easy and interesting, concepts and its relationship with some assignments related to current research and industry requirement should be discussed for motivation and long lasting remembrance. Also students should prepare the assignments for each concept based on their understanding. The concept map [2 and 3] based teaching is well tested with different subjects and proved that students can easily understand the concept. Hence, we propose concept map for teaching compiler design course. A concept map will simplify the teaching and make students understand the subject. In addition, we propose that for each concept; problems based on latest research and industry requirements to be discussed to motivate the students.

3. COMPILER CONCEPT MAP

Compiler is software that reads a program written in one language and translates it into another (target) or assembly, later to machine language without changing the semantics [1]. Interpreter is an alternate for compiler however it translates and executes directly instead of converting the complete code to the target language. Compiler does various operations to translate an input to output language. These operations are categorized into different phases. Figure 2 shows the overall concept map of the compiler with concepts and its relations.

It is important for an undergraduate students studying Compiler Design course to understand how the compiler is designed and program is getting translated. The concept map in figure 2 is designed with the motive to encourage the students to deeply focus on this subject. The concept map includes the input from the pre-processor and pre-requisite for this subject such as

instruction set, Context free and context sensitive grammar, regular expression, finite state machine and push down automata. The proposed concept map gives the complete overview of compiler design course and relationship among the concepts.

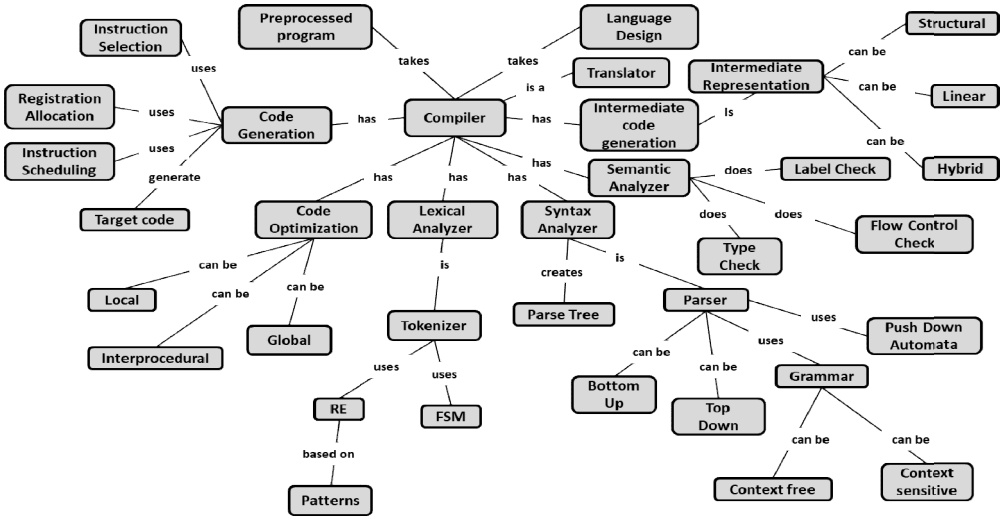


Figure 2. Basic concept map of Compiler

4. PHASES OF COMPILER

Compiler has six phases however there are references with seven phases dividing optimization into two parts that is machine dependent and independent optimization. In this paper, we have considered machine independent optimization in the code optimization phase and machine dependent optimization in code generation phase itself. In the beginning, it is required to explain to the students that what is the need for six phases instead of 1 or 10 phases? The major reason is that if we have all operations in one phase then it will increase the computational time and having more phases such as 10 or 15 will have the redundant process. Compiler cannot be efficient if we have 1 or 10 phases. Each phase of the compiler with its concept map is discussed in the following.

4.1. Lexical Analysis

Lexical analysis is the first phase of the compiler that takes the pre-processed program as input and produces tokens as output. The concept map for the lexical analysis phase is shown in figure 3. The important concepts to be discussed in lexical analysis phase are Tokenization, Buffering, Finite State Machine, Regular Expression and Symbol Table and obviously how these concepts are interlinked to generate the tokens. To tokenize the input programs, lexical analyser utilizes the pre-defined regular expression of the programming language and recognizes using the finite state machine. Lexical analyser reads the program characters from the buffer one by one to recognize the tokens. Instructor should also discuss the difference of single and pairs of buffer.

Even though, students studied regular expression and finite state machine in the pre-requisite subject Automata theory, the concepts need to be revised with an example of programming language patterns. In the classroom, instructor may solve the following problems for the better understanding of each concept.

- Create the regular expression for C Language variables.
- Construct the finite state machine for C variables using the regular expression generated in the previous problem.
- Take the hello world C program as input and recognize the tokens using single and double buffering. Assume buffer size as 10 bytes and show the advantage of double buffering over single or greater than two buffering.

In additional students should be given assignments and problems on each concept to solve.

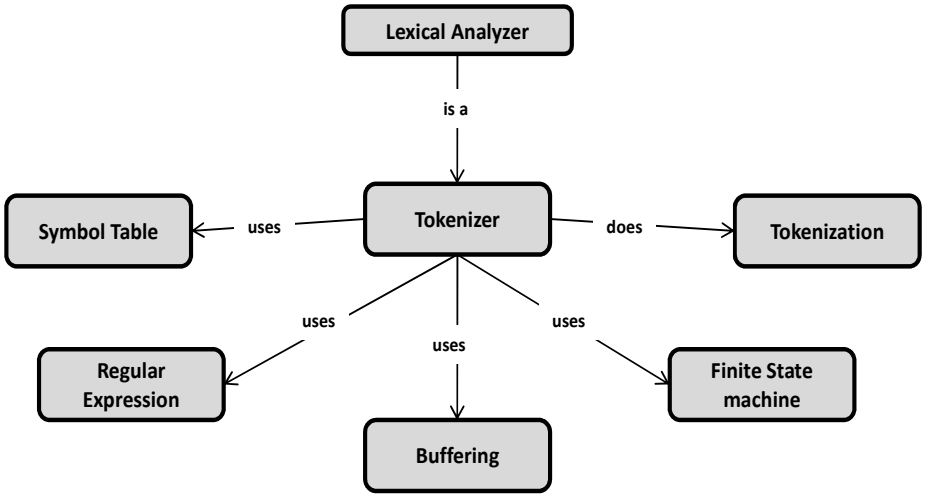


Figure 3. Concept map of Lexical Analyser

4.2. Syntax Analysis

Syntax Analysis takes the input as tokens from the lexical phase and produces the syntax tree, which will be used by the semantic phase. The concept map of syntax analyzer is shown in figure 4. Task of the syntax analyzer is to check whether the syntax present in the program is part of the programming language or not. To do this, Syntax Analyzer uses parser with Finite State Machine, Push down Automata and Context Free and Sensitive Grammar. Parser can be of universal, top down or bottom up approach and it will be chosen based on the developer requirement. The LL (Left-to-right, Leftmost derivation) and LR (Left-to-right, Rightmost derivation) parsers work only on the unambiguous grammar to parse in linear time. For LL parser, the unambiguous, deterministic and non-left recursive grammar will be taken as input and computes the first and follow. Using the first and follow, the parsing table will be constructed. In case of the LR parser, the finite state machine based item set for the unambiguous grammar will be generated and parsing table will be constructed from it.

The LL and LR parser takes the token from the lexical phase, parsing table and utilizes the push down automata to check the syntax and in parallel it produces the syntax. Instructor need to briefly discuss about the universal parsers like Earley's and CYK, which can take any type of grammar and parse it however the complexity of syntax validation will be more. Instructor need to solve one problem for each concept to make the student understand it. As a simple case, following problems can be solved in the class using the arithmetic operation grammar in figure 5.

- Eliminate Left recursion and compute the first & follow for the grammar.
- Compute the LR Automaton for the grammar.
- Compute LL and LR table using the first & follow and LR Automaton output respectively.
- Parse the input "a * b + c" through LL, LR and CYK parser using the parsing table computed in the previous assignment.

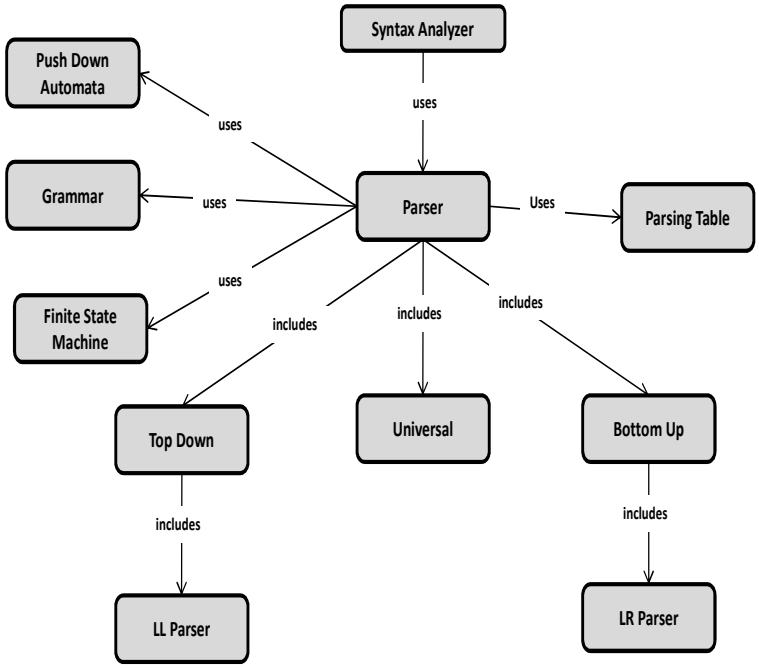


Figure 4. Concept map of Syntax Analyser

E- > E+T | E-T | T
T - > T*F | T/F | F
F - > Int

Figure 5. Sample Grammar for Arithmetic Operation

4.3 Semantic Analysis

Figure 6 shows the concept map of the semantic analyser. It takes the parse tree as input from the syntax analyzer and checks the semantics of the program such as type matching, parameter matching, label check, etc. This layer uses the attribute grammar or direct method to verify the semantics. The concepts in the semantic analyser are the Symbol Table, Attribute Grammars, Semantic check, Syntax Directed Translation and Definition. Attribute grammars can be represented using the Syntax Directed Translation (SDT) or Syntax Directed Definition (SDD) and use Dependency Graph at the time of evaluation to maintain the order. Semantic analyzer utilizes the symbol table to check the semantics of the program. Many high level programming languages do not have the attribute grammars for semantic check and it checks the semantics using the symbol table.

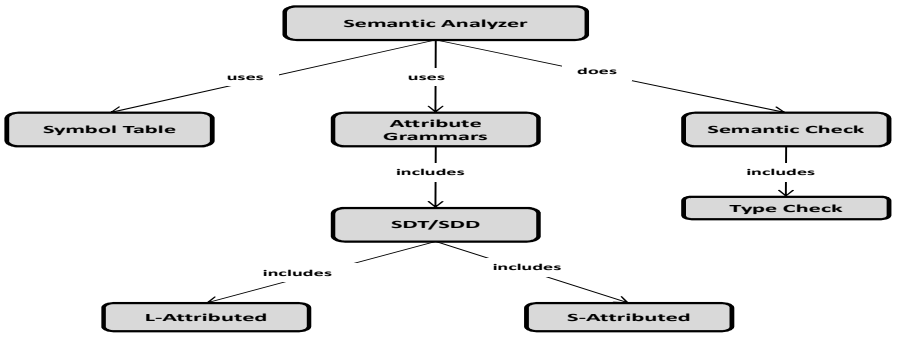


Figure 6. Concept map of Semantic Analyser

Instructor can verify the semantics of the snippet given in figure 7 using the C language constraints and show every step of the process in the class for better understanding

```
func(int a, float b):
int d, e
d = 0
e = a / round(b)
if (e > 5)  {
        printf("%d", d)
}
```

Figure 7. Sample snippet for semantic check

4.4. Intermediate Code Generation (ICG)

Figure 8 shows the concept map of the intermediate code generation phase of the compiler. ICG takes the input as syntax tree from the semantic analyzer and provides the Intermediate Representation. Intermediate Representation can be of structural, linear or hybrid. In most of the programming languages, linear representation is used and in that three address code is mostly preferred with any storage representations such as quadruple, triple and indirect triple. This phase may use the syntax directed translation to convert into three address code.

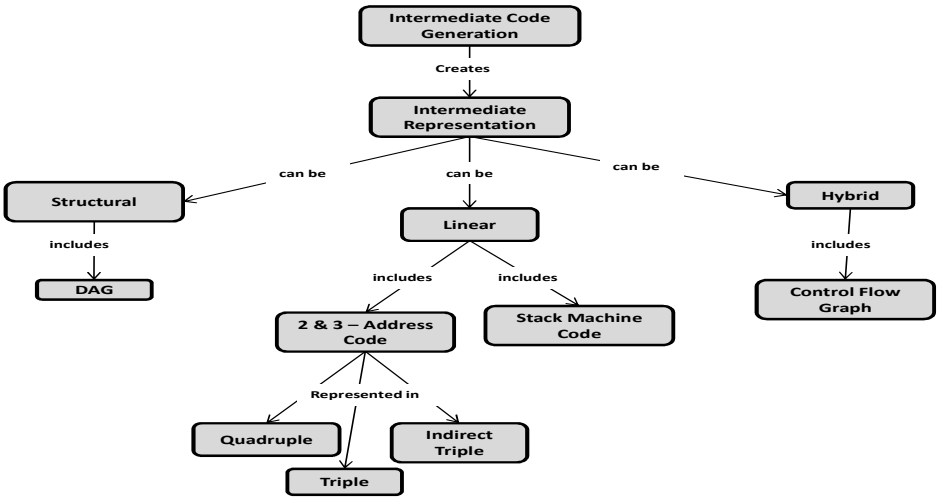


Figure 8. Concept map of Intermediate Code Generation

As an example, instructor can take the bubble sort program and generate the Direct Acyclic Graph (DAG) and three address code for it.

4.5. Code Optimization

Code Optimization is to make the program run efficiently after compilation. It helps in optimizing the time and space complexity of the program without changing the semantics. Code optimization can be done locally, globally or inter-procedural and it can be machine dependent or independent. Figure 9 shows the concept map of the code optimization phase. The different optimization concepts are Copy Propagation, Dead code elimination, Code Motion, Global Common Sub expressions, constant folding, loop optimization, unreachable code elimination, etc. Intermediate Representations from the previous phase will be grouped as blocks and then optimization will be applied locally and globally. In the classroom, instructor has to show one example for each optimization concept to make the concepts clear for students.

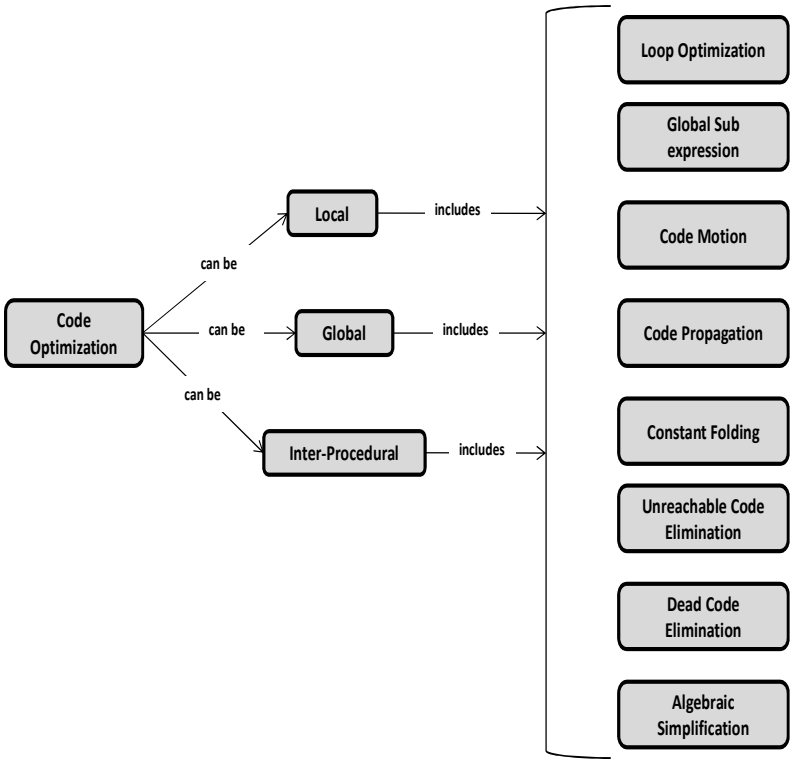


Figure 9. Concept map of Code Optimization

4.6. Code Generation

Code Generation is the final phase of the compiler that takes the optimized code and generates the target code. Target code may be another High level language code or Assembly level language code. To generate the assembly level, it is important to consider the Instruction Set, Register Allocation and Instruction Scheduling. The Next use information and Basic block concepts supports in register allocation for the target code. Figure 10 shows the concept map of the code generation.

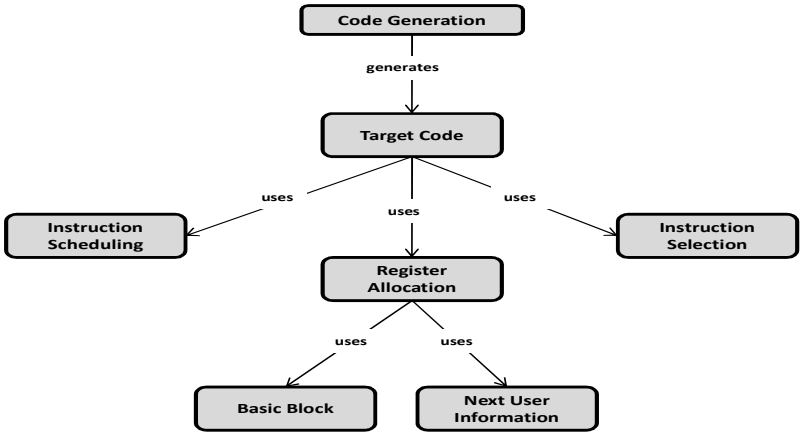


Figure 10. Concept map of Code Generation

Instruction set will be based on Reduced Instruction Set Computer (RISC) or Complex Instruction Set Computer (CISC) or Micro-op (mix of CISC and RISC). Since CISC architecture is used in popular Intel processors, CISC instruction set can be taken for solving the example in classroom. However, students should be asked to generate the target code for all architecture.

5. DISCUSSION

In the beginning of the Compiler Design course, instructor need to describe the course through the overall concept map, which includes the pre-requisites and new concepts students are going to learn. A concept map with the relationship among concepts can motivate the students to focus on the course. Every concept in the phase wise concept map should be discussed in the class with example problems that may be from the text book. In addition, the current research works and the industry products related to each concept should be briefly discussed to increase the interest of students. However, there should be assignments to the students on each concept from the recent research work and industry developments. Table 2 shows the sample assignments and problems related to each phase. Since students solve the research and industry oriented problems or assignments, they will be more interested on the course. Some assignments in table 2 especially research oriented assignments can be solved referring the literature cited.

In addition to the regular assignments, students should be asked to develop a compiler considering Classroom Object Oriented Language (COOL) [4] and Decaf [5]. This will help students to clearly understand the working process of each phase of the compiler. In case Compiler Design course contains the laboratory component then it should be used as Compiler design laboratory instead of programming laboratory.

Table 2. Assignments and Problems for each phase of the compiler

Concept	**General Problem**	**Industry oriented assignment**	**Research oriented assignment**
Lexical Analyser	How tokenization can be done for strings that are more than the length of two buffers.	Lexical Analysis supports in log analysis. Take a sample log file of an industry and apply lexical analysis to analyse the log file.	Apply Lexical analysis to prevent or detect SQL Injection Attack [8].

Syntax Analysis	How do we find the Shift/Shift or Shift/Reduce conflict in the grammar?	Syntax Analysis supports in log analysis. Take a sample log file of an industry and apply syntax analyser to analyse the log file.	Create the EMail syntax verification parser [7].
Semantic Analyser	Show SDT based semantic check is better than simple symbol table based check or otherwise.	Take a sample feedback file of an industry and verify the semantics.	Analyse the type checking [6].
Intermediate Code Generation	Convert any simple arithmetic expression parse tree to three address code and stack machine code. Identify the best with respect to time complexity.	Analyse the usage of Intermediate Representation in NetASM [7]	Analyse the graph-based higher-order intermediate representation [11]
Code Optimization	Remove the common sub-expression in the C code segment and find the time complexity. X+ = (4*j + 5*i) Y+ = (7 + 4*j)	Take the software or program from the Github and apply the code optimization. Analyse the time difference in execution.	Analyse the Optgen: A generator for local optimization [9]
Code Generation	Take any arithmetic expression (c = a + b) three address code and find how many move instructions are required to store the result? Assume there is only two registers.	Analyse how the Industry could generate code for the embedded systems.	Generate the Snowflake architecture instructions for calculator C program [10].

6. CONCLUSION AND FUTURE WORK

The Compiler Design course is chosen in this paper because from the survey results it is evident that many students are pre-decided that this subject is of less scope in the research and development. This paper introduced a concept map for the Compiler Design course with necessary class room problems and assignments based on recent research and industry development. With the concept map and assignments, students can understand the importance and relationship of the Compiler Design course with other courses. Instructors teaching Compiler Design course need to update the students' assignments based on the latest research and industry progress. Hence, students will get motivated and focus on this subject. The future work of this paper is to make the students to design the problems and assignments for each concept related to industry and research and measure the accuracy of the students' knowledge on the concept and show the relevance of each concepts and phases in recent Research and Development.

ACKNOWLEDGEMENTS

We would like to thank the authors of different books, web articles and research papers from which the assignment problems and class room examples are taken and discussed in this paper. Also would like to thank the students participated in the survey.

REFERENCES

[1] Alfred V. Aho, Monica S. Lam, Ravi Sethi and Jeffrey D. Ullman, *Compilers: Principles, Techniques and Tools*, Second Edition Book, 2006.

[2] Joseph D. Novak and Alberto J. Ca´nas, The Theory Underlying Concept Maps and How to Construct Them,http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptMaps.bck-11-01-06.htm [accessed on 10 January 2019]

[3] Joseph D. Novak, A theory of education. NY: Cornell University, 1977.

[4] CoolAid: The Cool 2016 Reference Manual, URL: http://pabst.cs.uwm.edu/classes/cs654/handouts/cool-manual.pdf [accessed on 11 January 2019].

[5] Julie Zelenski, Jerry Cain and Keith Schwarz, Decaf Specification, URL:http://web.stanford.edu/class/archive/cs/cs143/cs143.1128/handouts/030%20Decaf%20Specification.pdf [accessed on 11 January 2019].

[6] Francisco Ortin, Daniel Zapico, and Juan Manuel Cueva, Design Patterns for Teaching Type Checking in a Compiler Construction Course, *IEEE Transactions on Education*, Vol. 50, No. 3, pp.273-283, 2007.

[7] Muhammad Shahbaz , Nick Feamster, The case for an intermediate representation for programmable data planes, *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, 2015.

[8] L. Ntagwabira, S. L. Kang, Use of query tokenization to detect and prevent sql injection attacks, *3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT) 2010*, Vol. 2, pp. 438-440, 2010.

[9] S. Buchwald. Optgen: A generator for local optimizations, *Proceedings of the 24th International Conference on Compiler Construction (CC),* pp. 171-189, Apr. 2015.

[10] Andre Xian Ming Chang, Aliasger Zaidy and Eugenio Culurciello, Efficient compiler code generation for Deep Learning Snowflake coprocessor, *1st Workshop on Energy Efficient Machine Learning and Cognitive Computing for Embedded Applications*, pp. 24-28, 2018.

[11] R. Leißa, M. Köster, and S. Hack, A graph-based higher-order intermediate representation, *Proceedings of the 13th Annual IEEE/ACM International Symposium on Code Generation and Optimization*, pp.202-212, 2015.

[12] Akim Demaille, Roland Levillain and Benoît Perrot, A set of tools to teach compiler construction, in *Proceedings of the 13th Annual Conference on Innovation and Technology in Computer Science Education*, (ITiCSE '08), pp. 68-72, 2008.

[13] Li Xu and Fred G. Martin, Chirp on crickets: Teaching compilers using an embedded robot controller, *Proceedings of the 37th SIGCSE technical symposium on Computer science education*, Vol. 38, no. 1, pp. 82-86, 2006.

[14] Tyson R. Henry, Teaching compiler construction using a domain specific language, *Proceedings of the 36th SIGCSE technical symposium on Computer science education*, Vol. 37, no. 1, pp. 7-11, 2005.

[15] Elizabeth White, Ranjan Sen, and Nina Stewart, “Hide and show: Using real compiler code for teaching, *Proceedings of the 36th SIGCSE technical symposium on Computer science education*, Vol. 37, no. 1, pp. 12-16, 2005.

[16] S. R. Vegdahl, “Using visualization tools to teach compiler design,” in Proceedings of the Fourteenth Annual Consortium on Small Colleges South eastern Conference (CCSC ’00), pp. 72-83, 2000.

[17] Marjan Mernik and Viljem Zumer, An educational tool for teaching compiler construction, *IEEE Transactions on Education*, Vol. 46, no. 1, pp. 61-68, 2003.

[18] Henry D. Shapiro and M. Dennis Mickunas, "A new approach to teaching a first course in compiler construction," *Proceedings of the ACM SIGCSE-SIGCUE technical symposium on Computer science and education*, pp.158-166, 1976.

[19] Martin Ruckert, "Teaching compiler construction and language design: Making the case for unusual compiler projects with postscript as the target language, *Proceedings of the 38th SIGCSE technical symposium on Computer science education*, Vol. 39, no. 1, pp. 435-439, 2007.

[20] Divya Kundra and Ashish Sureka, Application of Case-Based Teaching and Learning in Compiler Design Course, arXiv:1611.00271v1 [cs.PL] 1 Nov 2016.

[21] D. Kundra and A. Sureka, "An experience report on teaching compiler design concepts using case-based and project-based learning ap-proaches," in *International Conference on Technology for Education (T4E 2016)*, 2016.

[22] Somya Sangal, Shreya Kataria, Twishi Tyagi, Nidhi Gupta, Yukti Kirtani, Shivli Agrawal and Pinaki Chakraborty, PAVT: a tool to visualize and teach parsing algorithms, *Education and Information Technologies, An official Journal of the IFIP Technical Committee on Education*, Vol. 23, No.6, pp.2737-2764 pp.2018.

SOLVING THE CHINESE PHYSICAL PROBLEM BASED ON DEEP LEARNING AND KNOWLEDGE GRAPH

Mingchen Li[1], Zili Zhou[2], Yanna Wang[1]

[1]College of Physics and Engineering, Qufu Normal University, Qufu 273165, PR China
[2]School of Software Engineering, Qufu Normal University, Qufu Shandong 273165, PR China

ABSTRACT

In recent years, problem solving, automatic proof and human-like test-tasking have become a hot spot of research. This paper focus on the study of solving physical problem in Chinese. Based on the analysis of physical corpus, it is found that the physical problem are made up of n-tuples which contain concepts and relations between concepts, and the n-tuples can be expressed in the form of UP-graph (The graph of understanding problem), which is the semantic expression of physical problem. UP-graph is the base of problem solving which is generated by using physical knowledge graph (PKG). However, current knowledge graph is hard to be used in problem solving, because it cannot store methods for solving problem. So this paper presents a model of PKG which contains concepts and relations, in the model, concepts and relations are split into terms and unique IDs, and methods can be easily stored in the PKG as concepts. Based on the PKG, DKP-solving is proposed which is a novel approach for solving physical problem. The approach combines rules, statistical methods and knowledge reasoning effectively by integrating the deep learning and knowledge graph. The experimental results over the data set of real physical text indicate that DKP-solving is effective in physical problem solving.

KEYWORDS

Knowledge Graph, Deep Learning, Problem Solving, & Physical Problem

1. INTRODUCTION

Developing machine solver [1] has become an active research topic since the 1960s and recently it has developed rapidly under the influence of fast advance of artificial intelligent (AI), Human-Computer Interaction [2], machine learning [3, 4], and pattern recognition. It can be applied to various courses, for all stages of the students to provide intelligent answer services. In this paper, we consider how to make machine solve physical problem in Chinese.

For physical problem in Chinese, its sentence structure and semantics [5] are complex and diversity, this improves the difficulty of problem solving, on the other hand, the machine does not have the congenital ability of understanding and analysing, it cannot correctly identify the semantic information of physical problem. However, it is simple for human. Human can apply unique brain thinking to get the semantic information for problem, and to solve them. So, it is a challenge which makes the machine completely imitate human learning and problem solving.

To make machine think like human in dealing with physical problem, one of the important prerequisites is to let machine understand the semantic information of physical problem and

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 103–115, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90310

transform them into a knowledge representation that machine can understand. The study found that knowledge graph [6] can allow machine to understand the semantic information in the physical problem accurately. However the existing knowledge mapping model cannot achieve good results in problem solving [7] because of the lack of expression ability. So this paper presents a model of PKG for physical problem solving, it contains concepts and relations which are split into terms and IDs. From the perspective of knowledge expression, every physical problem is made up of n-tuples which contain concepts and relations between concepts. These n-tuples can be expressed in the form of UP-graph based on PKG. By using UP-graph and the methods stored in the PKG, The physical problem can be solved finally.

The paper presents a novel approach called DKP-solving (solving the physical problem combine deep learning [8, 9] and knowledge graph). It consists of two stages: Firstly, the physical concepts and the relations are extracted by using deep learning, natural language processing (NLP) [10] and term set in PKG and grouped them in the form of n-tuples. By using PKG, UP-graph is generated which is the semantic expression of physical problem. Secondly, the plan of solving problem is given based on the UP-graph and methods stored in PKG. This novel approach of solving the problem is not just for the physical filed, it's also for other discipline, such as mathematic, biology and geography.

The reminder of this paper is organized as follows: Section 2 summarizes the related work about some researches on the automation of problem solving. The proposed novel approach is introduced in Section 3. Section 4 introduces the experiment and the process of physical problem solving. The paper closes with the conclusions and future work in the last section.

2. RELATED WORK

Around the task of problem solving, many companies and laboratories from domestic and foreign have made related studies. Here are some related works of problem solving from some authorities and companies.

In 1964, Bobrow et al. [11] developed the earliest intelligent teaching system-STUDENT which could solve the problem, this system can understand algebraic problem expressed in English, it transforms the text with the form of natural language into a relational model. By manipulating these relational models to achieve the problem of automatic solution. However, STUDENT store sentence structure is very limited.

After the 1980s, with the development of cognitive psychology, especially the classification of problem from semantic layer, promoting the further development of semantic understanding. Riley et al. [12] made a further solving of addition and subtraction problem in 1983, dividing it into three categories: combining problem, transferring problem and comparing problem. And they segmented 14 sub-categories from these three categories. In 1985, Kintsch et al. [13] made a deep cognitive research on one step addition and subtraction application problem. They put forward a characterization model of one step application to solve problem. In 1986, Dellarosa et al.'s [7] research based on the above two theories, the ARITHPRO system was developed which simulates the human cognitive process, to achieve the understanding of one step addition and subtraction application problem.

After the 1990s, the method of statistical natural language processing based on a large corpus were widely used in the field of natural language understanding and achieved good results. Andes Tutorial System is an intelligent tutoring system based on physical mechanics, through understanding the meaning of the problem, completing the interaction with the students [14]and providing problem-solving feedback.

In 2014, Stanford University published Modelling Biological Processes for Reading Comprehension [15] in the EMNLP. Firstly, the biological information process in the text through the forecast to build a rich representation structure, and then use this structure to answer the question, but only for biological knowledge .The Google Deep Learning Group used end-to-end neural networks to solve the problem of reading comprehension, with the application of image-based Attention mechanism in a recurrent neural network (RNN) [16] and the emergence of Attention-based machine translation model [17] and Attention-based image [18] generation. It is usually said that RNN is unable to remember the text of longer distance text in the network related to timing information [19], so Google deep mind mainly uses Attention's bidirectional Deep Long Short-Term Memory (LSTM) [20], which is an improved version of the bidirectional RNN for solving the reading comprehension. Facebook and other companies and laboratories are also actively studying this issue. Especially Facebook's babi projects [19] in 2015, Facebook proposed a new Memory Network for understanding problem' meaning and solving problem according to MeNN [21]. Huawei's Noah's Ark Laboratories experimented with Babi's data set and achieved better results than Facebook's original test results. They used reasoning model [22] based on GRU [23] neural network. In addition, in 2016, Facebook built another data set for test and training neural network with solving the comprehension reading—children book test, referred to as CBT. In 2016, IBM Watson's researchers built a new set of depth models-- Attention Sum Reader [24], referred to as ASR which is a model that uses Attention to predict answers directly. This model draws on a structure called Point Network [25]. By summing the Attention values of the same word appearing at different text positions, and the value as the final answer probability distribution value. This model uses the deep learning library theano.

Domestic and foreign related companies and research institutions have made a great contribution for improving the machine ability of problem solving, some researches are in the study of reasoning logical structure, and some using deep learning to solve the problem, but the key point of the present research is to allow machine to understand human intend, and to imitate human behaviour. Chinese sentence structure has unique syntax, complex semantic environment, which deepen the difficulty of problem solving, furthermore some of the knowledge in the application problem is hidden knowledge, more difficult to perceive. Therefore, DKP-solving is proposed to solve the above problems, and the experiment proved that it did improve the machine's intelligence.

3. DKP-SOLVING ARCHITECTURE

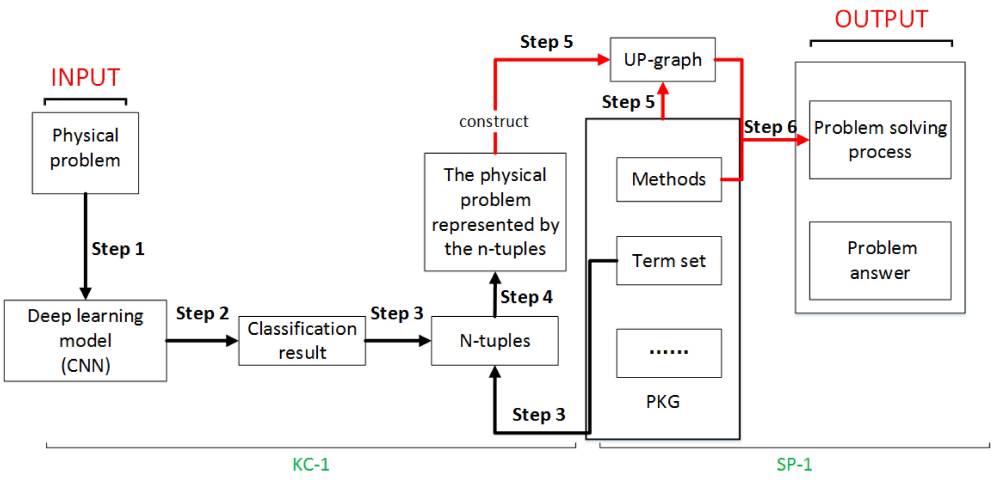


Figure 1. The sketch of DKP-solving

As shown in Figure 1, DKP-solving has two stages. The first stage is called KC-1(First-generation model for extraction of n-tuples) which steps are marked by black arrows, its purpose is to extract n-tuples, which combines the knowledge graph and the Convolutional Neural Network (CNN) [26]. Physical problem (INPUT) is entered the deep learning model (CNN) which built in step1. The classification results for every sentence in physical problem are got by step2. Combining those results with term set in PKG the n-tuple for every sentence are acquired by step3. Finally, physical problem can be represented by the n-tuples in step 4. At the second stage where the steps are marked by red arrows is called SP-1(First-generation for solving problem using knowledge graph), problem's n-tuples will form an UP-graph which is semantic understanding of physical problem in step 5. Under the premise of understanding the meaning of the problem, using the methods stored in the PKG to solve the problem in step 6, and finally get the problem solving process and answer (OUTPUT).

3.1. KC-1's Building

Extracting n-tuples from physical problem helps to express physical semantic information accurately. By matching n-tuples and concepts in PKG, the purpose of building the UP-graph can be achieved which can help solve physical problem.

3.1.1. Definition 1 (Sentence Structure)

By analysing, it is found that the structure of the sentence in physical problem is limited. So the sentences can be divided into 5 types. These five categories are shown in Table 1.

Table 1. Sentence classification

Category	A	B	C	D	E
Sentence Structure	$[h_i\_h_i r_i t_i\_]$	$[h_i r_i t_i u_i]$	$[\_h_i h_i\_h_i r_i t_i]$	$[h_i\_h_i r_i t_i u_i]$	$[\_h_i\_h_i r_i h_i\_t_i\_]$

In Table 1, h_i is a conceptual term in physics, r_i is the relational nouns, characterizing the relations between h_i and t_i , t_i is usually the value of the physical quantity. u_i is the unit of physical concept. "_" is usually a word that does not represent the actual meaning in the sentence and does not affect the overall semantic meaning of the sentence, such as: "此" (this), "的" (of), and "已知" (a known). In order to reasonably explain the meaning of five categories, we give an example, as follows: "小物块的质量为 0.5" (The quality of small block is 0.5), "小物块" (small block) and "质量"(quality) are h_i, "为" (is) is r_i, "0.5" is t_i. It belongs to category A.

3.1.2. Definition 2(Classification-Model)

In order to classify the physical sentence, we use CNN. In physical text analysis, CNN can make it possible to deal with this kind of problem because of the limited length and the compact structure of physical sentence. In this model, the first layer embeds the word into the low-dimensional vector, and the text sequence is expanded into the word vector sequence; The second layer is convoluted with some length different convolution kernels, convoluted followed by the activation function; The third layer is the pooling layer and add dropout; The fourth layer is the fully connected layer and is classified by SoftMax [27].

3.1.3. The Progress of KC-1

Firstly, each sentence in the problem can get the structural classification by the CNN. Secondly, every sentence in physical problem is segmented. Finally, the n-tuples for physical problem are extracted according to the term set of PKG and its struction.

3.2. SP-1's Building

The problem solving process which called SP-1, in this progress, we made a novel knowledge graph model (PKG) which can store semantic information and methods about physical problem.

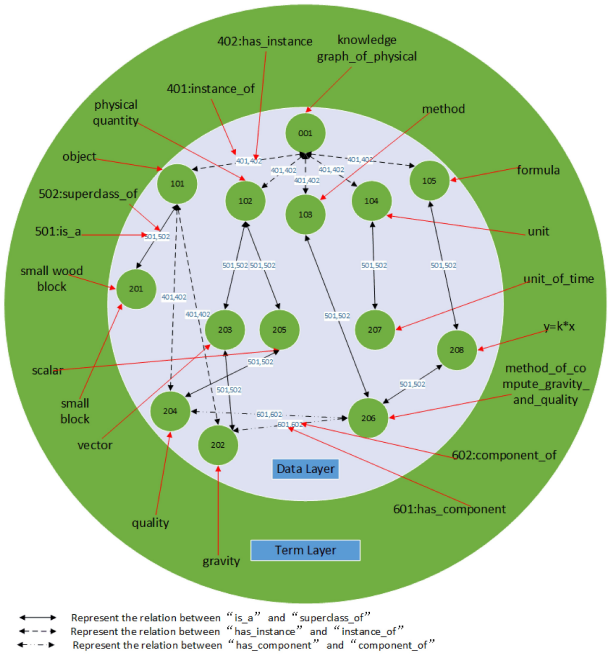


Figure 2. PKG's model structure

Figure.2 is a visual display of PKG's model. The definition of semantic elements based on the PKG model is composed of natural language symbols (terms in the figure) and unique identifier of concepts (represented by ID). Different terms of concepts or relations which have the same meaning use the same ID, such as: "small wood block" and "small block" use the same ID number (201), that is, use the same semantic unit. In the model, Methods are seen as concepts, and have its semantic information and its ID, for example, the method which computes the gravity is split into term "method_of_compute_gravity_and_quality" and ID (206). Furthermore, the method has relations with gravity (202) and quality (204).

3.2.1. The Construction of PKG

Based on the model of PKG, we have built our own PKG which store physical concepts, relations between physical concepts, methods (calculation formulas) in concepts etc. In PKG, we split the physical concept into term and unique ID, term is semantic information of physical concept, and unique ID can be identified by machine, so the representation of physical concepts has two parts, one is term set, and the other is ID set. By this way, machine will be easy to solve the problem about synonymy and polysemy. PKG contains five physical upper concepts, such as "object", "physical quantity", "method", "unit" and "formula", and their unique ID number are "101",

"102", "103", "104", and "105" respectively. There are a number of child nodes below each parent node, such as the parent concept "object" has the child concept "quality" and the child concept "gravity", in data layer, they are "101", "204" and "202".

In PKG, concepts are connected by a certain relation pair which contains two relations, for example, there are two relations ("is_a" and "superclass_of") between "object" and "small block", their unique ID number are "501" and "502". There are two relations ("has_instance" and "instance_of") between "object" and "quality", their unique ID number are "402" and "401", and there are two relations ("has_component" and "component_of") between "quality" and"methos_of_compute_gravity_and_quality", their unique ID number are "601" and "602". The concepts and relations formed UP-graph finally. A child concept inherits the nature of the parent concept, for example, child concept "small block" inherits the "quality" and "gravity" of the parent node "object", unique ID number of four concepts are "201", "204", "202" and "101".

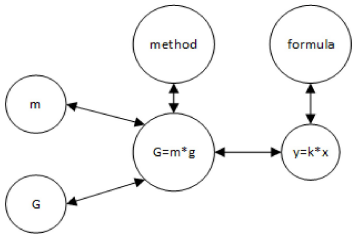


Figure 3. The method in PKG

Method is also a concept and it is the key to solving physical problem, Each method node has its own inputs and outputs, as shown in Figure 3, which provides a method of (G=m*g). The method is applied to calculate gravity (G), quality (m), and gravitational acceleration (g). Gravitational acceleration is constant, it is known that any one of gravity (G), quality (m), according the method (G=m*g) can be obtained unknown one. Every method has its own calculation form which can be expressed by formula in math, for (G=m*g), its form is (y=k*x). Table 2 shows some of the methods and the corresponding formula.

Table 2. Methods and formulas

methods	describe	formula
G=m*g	Calculate gravity, quality, gravitational acceleration.	y=k*x
$m = \rho * v$	Calculate quality, density ,volume	
S=v*t	Calculate distance, speed ,time	y=x1*x2
F=P*S	Calculate force, pressure, area	
U=I*R	Calculate voltage, electric current, resistance	
I=I1+I2	Calculate the sum of electric current (series connection)	y=x1+x2
P=ρgh	Calculate pressure, density, height, gravitational acceleration	y=k*x2*x3

Summary: This graph mainly including the four blocks. There are physical quantities associated with objects, the physical quantity needs to be measured by the unit, solving the problem requires

methods, they are connected between each other. PKG can make machine understand the semantic of physical problem and then solve them. Figure.4 is our PKG (part).

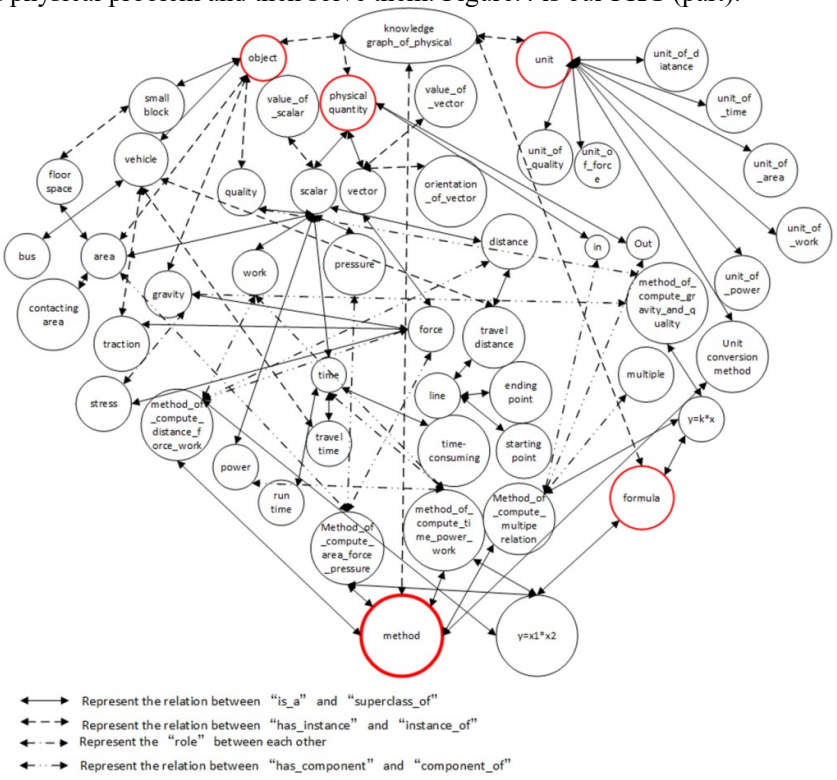


Figure 4. PKG (part of physical knowledge graph)

3.2.2. The Process of SP-1

By KC-1, each sentence of physic problem will form an n-tuple, which contains some words. By searching the term set of PKG, it will be found that the corresponding concept for each word in every tuple, every concept will found their associated concepts in PKG, these concepts will form a tree for every word. Traversing the trees formed by each concept, trees and trees will be directly linked by related concepts, if no link between each other, with the help of their upper node in PKG to establish contact. So the UP-graph is formed. And then, according to the UP-graph, the known value, unknown value can be found. Simultaneously, the methods in the PKG also can be found, finally, this physic problem will be solved through the above steps.

4. EXPERIMENTS AND ANALYSIS

4.1. Corpus

Different from the open source corpus of foreign language, physical corpus in Chinese are lacking. Therefore, a corpus was established for the experiment, in which data (25000) were crawled from the network about the physical problem. Per category has 5000 dates.

4.2. Data-Pre-Processing Rules and Parameter Settings

In the data pre-processing, we use Jieba and PKG for word segmentation. In order to make the classifier better, we set some parameters for CNN model, three kinds of kernels which window's

size are 3*128, 4*128, and 5*128 are chosen. For the dropout rate, batch size, learning rate, we set to 0.5, 128, and 0.0001.

4.3. The Process of Solving Physical Problem

As far as we know, DKP is the first approach about solving the physical problems based on knowledge graph and deep learning. This also increases the difficulty of approach comparison, in order to make readers understand DKP better, we used one example to describe it. The result of the calculation is also given at the same time.

Physical problem (example M) is shown below:

一辆大巴车从甲地到乙地运送游客，已知该大巴车在行驶过程中受到的牵引力是重力的 0.15 倍， 质量为 5t，时间为 11hour，线路的长度为 70km，大巴车牵引力做功的功率是多少？(A bus transport tourist from A to B, in the process of this line, traction is 0.15 times that of the gravity, the quality is 5t, time is 11hour, the length of the line is 70km, how much is the power traction work of the bus?)

4.3.1. Extract N-Tuples

The classification is shown in Table 3

Sentence	Category
一辆大巴车从甲地到乙地运送游客，(A bus transport tourists from A to B,)	D
已知该大巴车在行驶过程中受到的牵引力是重力 0.15 倍，(in the process of this line, traction is 0.15 times that of the gravity,)	E
质量为 5t(the quality is 5t,)	B
时间为 11hour,(time is 11hour,)	B
线路的长度为 70km，(the length of the line is 70km,)	D
大巴车牵引力做功的功率是多少？(how much is the power traction work of the bus?)	C

Table 3. Sentence classification of M

The n-tuples extraction result for M is:

[['大巴车','甲地','乙地',], ['大巴车', '牵引力', '重力', '0.15', '倍'],['大巴车','质量', '5', 't'], ['大巴车','时间', '11', 'hour'], ['线路', '长度', '70', 'km'], ['大巴车', '牵引力', '功率', '？ ']]([['bus', 'A', 'B',], ['bus', 'traction', 'gravity', '0.15', 'times'], ['bus', 'quality','5', t'], ['bus', 'time', '11', 'hour'], ['line', 'length', '70','km'], ['bus', 'traction', 'power', '？ ']])

4.3.2. SP-1's of M

M's UP-graph is formed according to the relations and concepts in Figure 5, "bus" has the attribute of "travel distance ","travel time ","traction", "gravity","quality ". Each attribute has a value and a unit. Such as: "quality" has a value of "5", the unit is "t"(ton). There is a relation between the value of "traction" and "gravity". X is the gravity value. Y is the traction value, Z is the work value, K is the power value. They represent unknowns. This graph lets the machine understand the meaning of each known or unknown physical quantity, and the relations between

them. This part is called UP-graph, the other part stored in the figure are the methods. These methods can be used to solve the M. such as :"method_of_compute_time_power_work". M's solving progress are as follows: The value of the "gravity", "traction" and the "work" is the prerequisite for obtaining the value of "power".

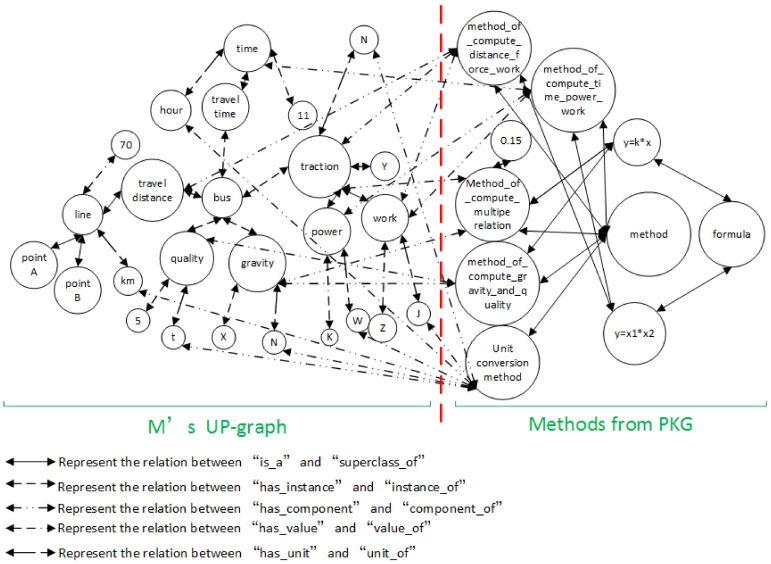


Figure 5. M's knowledge graph

The machine will use the graph to think like human and use logical reasoning to solve this problem, and it will use the graph to find the question of M is the "power", the known physical quantities' value are: "quality", "line ", "time "the unknown physical quantities' value are: "gravity", "traction", "work", "power" .

M's solving progresses and result are as follows:

Step 1: Units are normalized according to "Unit conversion method".
The step (1) for **M**'s solving progress is shown in Figure 6.

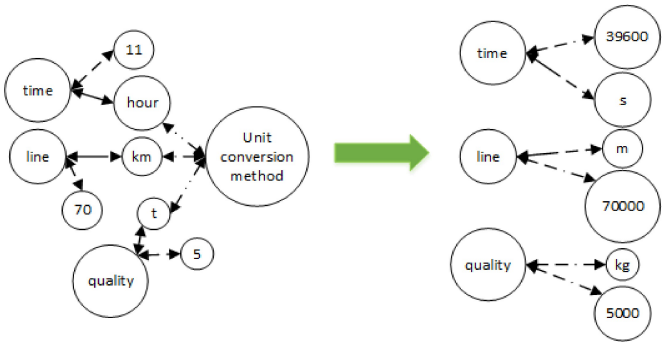


Figure 6. The step (1) for M's solving progress

By the Figure 6, the value and unit of "time" become "396000" and "s"(second), the value and unit of "line" become "70000" and "m"(meter), the value and unit of "quality" become "5000" and "kg"(kilogram).

Step 2: According to the problem to be solved, we need to find the method of getting the value of "power". So, the method which named "method_of_compute_time_power_work" is suitable.

Step 3: Find formula in graph which matching the method, and the formula (y=x1*x2) is got.

Step 4: Find other physical quantities other than "power" in the method. They are "work" and "time".

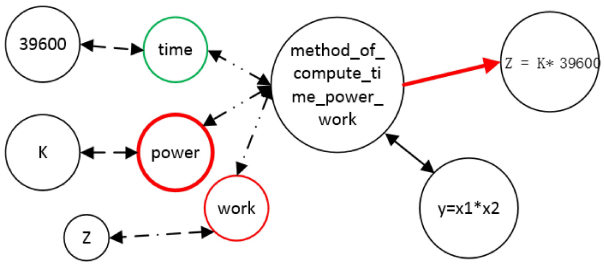


Figure 7. The step (2-4) for M's solving progress

The step (2-4) for **M**'s solving progress is shown in Figure 7. Physical quantity in red circle is uncertain, and physical quantity (time) circled by green circle has value (39600). So,

$$K = \frac{Z}{39600} \tag{1}$$

Step 5: Find the method which can get the value of "work", and the method "method_of_compute_distance_force_work" is suitable.

Step 6: Find the formula which matching this method, the formula (y=x1*x2) is got.

Step 7: By the method, "traction" and "line" are found.

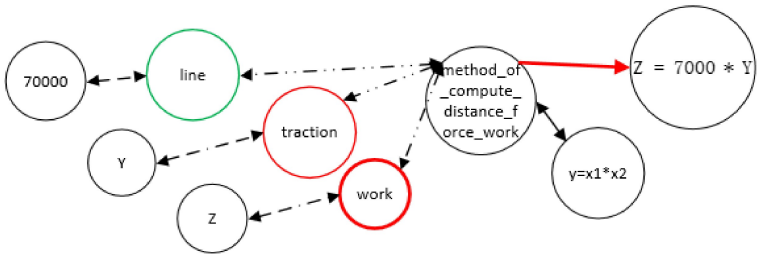


Figure 8. The step (5-7) for M's solving progress

The step (5-7) for **M**'s solving progress is shown in Figure 8. Physical quantity in red circle is uncertain, and physical quantity (line) circled by green circle has value (70000).That is,

$$Z = 70000 * Y \tag{2}$$

Step 8: Find the method which can solve the value of "traction", by the graph, we can find the "gravity" and "traction" has a multiple relation, and the method "Method_of_compute_multipe_relation" can get the value of "traction" .

Step 9: Find the formula to match the method, and the formula is (y=k*x).

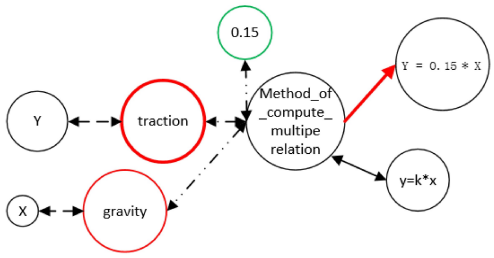


Figure 9. The step (8-9) for M's solving progress

The step (8-9) for **M**'s solving progress is shown in Figure 9. Physical quantity in red circle is uncertain, and physical quantity (times) circled by green circle has value (0.15).We can get:

$$Y = 0.15 * X \tag{3}$$

Step 10: Find the quantity which has no value, we get "gravity".

Step 11: Find the method which can get the value of "gravity", the method which named "method_of_compute_gravity_and_quality" is got.

Step 12: Find the formula which matching the method, this formula (y=k*x) is hit.

Step13: Find physical quantities other than "gravity" in the method. It is "quality", and it has value.

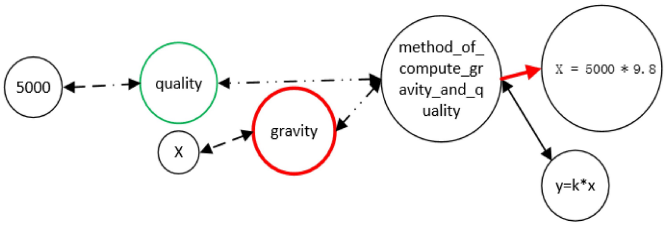


Figure 10. The step (10-13) for M's solving progress

The step (10-13) for **M**'s solving progress is shown in Figure 10. Physical quantity in red circle is uncertain, and physical quantity (quality) circled by green circle has value (5000). Gravitational acceleration has default value (9.8). So,

$$X = 5000 * 9.8 \tag{4}$$

Flow chart of M's solving problem is shown in Figure 11. Finally, the value of "power" and M's solving progress can be output.

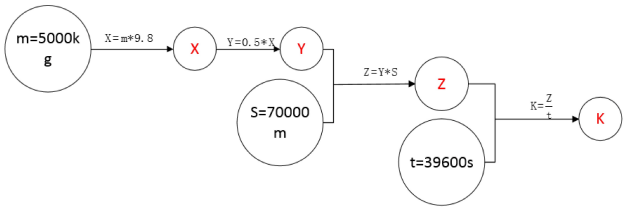


Figure 11. Flow chart of M's solving problem

$$X = 5000 * 9.8 = 49000\ N \tag{5}$$

$$Y = 0.15 * X = 0.15 * 49000 = 7350\ N \tag{6}$$

$$Z = Y * 70000 = 7350 * 70000 = 514500000\ J \tag{7}$$

$$K = \frac{Z}{39600} = \frac{514500000}{39600} = 12992.42\ w \tag{8}$$

So, the power of bus traction power is $12992.42\ w$.

4. CONCLUSIONS

The paper presents a novel approach of solving the physical problem based on Deep Learning and Knowledge Graph: DKP-solving. It combines the CNN and knowledge graph. CNN is used to obtain the sentence structure in physical problem. And then according to the different syntax structure to extract n-tuples. In the process of extracting n-tuples, we are using the classifier and term set of PKG. So every sentence in the problem will be represented by the n-tuples which are used to form the UP-graph. PKG is play an important role in DKP-solving, different from other knowledge graphs, a new model of knowledge graph is constructed which contains concepts, relations, the model can effectively express methods which are key contents in solving problems. By UP-graph and methods, the machine can solve the physical problem effectively and intelligently. In this paper we only use a very traditional way to extract the n-tuples, In the future, we will try more effective n-tuples extraction method, machine solving method, in order to solve the physical problem more effectively. This method also can contribute to other fields, such as Chinese reading Comprehension, mathematical problem solving, and biological computing analysis and so on.

ACKNOWLEDGEMENTS

This research is funded by the Science and Technology Commission of Shanghai Municipality (No. 16511102702) and by Xiaoi Research, by Shanghai Municipal Commission of Economy and Information Under Grant Project No. 201602024 and by the Science and Technology Commission of Shanghai Municipality (No.15PJ1401700).

REFERENCES

[1] V. Tsiriga and M. Virvou, A Framework for the Initialization of Student Models in Web-based Intelligent Tutoring Systems: Kluwer Academic Publishers, 2004.

[2] D. Bell, T. Koulouri, S. Lauria, R. D. Macredie, and J. Sutton, "Microblogging as a mechanism for human–robot interaction," Knowledge-Based Systems, vol. 69, pp. 64-77, 2014.

[3] J. R. Quinlan, C4.5: programs for machine learning: Morgan Kaufmann Publishers Inc., 1993.

[4] P. Langley, Elements of machine learning: Morgan Kaufmann Publishers Inc., 1995.

[5] C. Musto, P. Lops, M. D. Gemmis, and G. Semeraro, "Semantics-aware Recommender Systems exploiting Linked Open Data and graph-based features," Knowledge-Based Systems, 2017.

[6] J. Pujara, H. Miao, L. Getoor, and W. Cohen, "Knowledge Graph Identification," in International Semantic Web Conference, 2013, pp. 542-557.

[7] D. Dellarosa, "A computer simulation of children's arithmetic word-problem solving," Behavior Research Methods, Instruments, & Computers, vol. 18, pp. 147-154, 1986.

[8] L. Deng, J. Li, J. Huang, K. Yao, D. Yu, F. Seide, M. Seltzer, G. Zweig, X. He, and J. Williams, "Recent advances in deep learning for speech research at Microsoft," in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 2013, pp. 8604-8608.

[9] T. M. Mitchell, "Does machine learning really work?" AI magazine, vol. 18, p. 11, 1997.

[10] N. Kalchbrenner, E. Grefenstette and P. Blunsom, "A convolutional neural network for modelling sentences," arXiv preprint arXiv:1404.2188, pp. 655-665, 2014.

[11] D. G. Bobrow, "Natural Language Input for a Computer Problem Solving System," in Semantic Information Processing, 1964, pp. 281-8.

[12] M. S. Riley, "Development of children's problem-solving ability in arithmetic.," 1984.

[13] W. Kintsch and J. G. Greeno, "Understanding and solving word arithmetic problems.," Psychological review, vol. 92, p. 109, 1985.

[14] R. C. Murray and K. VanLehn, "A comparison of decision-theoretic, fixed-policy and random tutorial action selection," in International Conference on Intelligent Tutoring Systems, 2006, pp. 114-123.

[15] J. Berant, V. Srikumar, P. Chen, A. Vander Linden, B. Harding, B. Huang, P. Clark, and C. D. Manning, "Modeling Biological Processes for Reading Comprehension.," in EMNLP, 2014.

[16] V. Mnih, N. Heess and A. Graves, "Recurrent models of visual attention," in Advances in neural information processing systems, 2014, pp. 2204-2212.

[17] D. Bahdanau, K. Cho and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.

[18] K. Gregor, I. Danihelka, A. Graves, D. J. Rezende, and D. Wierstra, "DRAW: A recurrent neural network for image generation," arXiv preprint arXiv:1502.04623, 2015.

[19] J. Weston, A. Bordes, S. Chopra, A. M. Rush, B. van Merriënboer, A. Joulin, and T. Mikolov, "Towards ai-complete question answering: A set of prerequisite toy tasks," arXiv preprint arXiv:1502.05698, 2015.

[20] K. M. Hermann, T. Kocisky, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, and P. Blunsom, "Teaching machines to read and comprehend," in Advances in Neural Information Processing Systems, 2015, pp. 1693-1701.

[21] J. Weston, S. Chopra and A. Bordes, "Memory networks," arXiv preprint arXiv:1410.3916, 2014.

[22] B. Peng, Z. Lu, H. Li, and K. Wong, "Towards neural network-based reasoning," arXiv preprint rXiv:1508.05508, 2015.

[23] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, pp. 1724-1734, 2014.

[24] R. Kadlec, M. Schmid, O. Bajgar, and J. Kleindienst, "Text understanding with the attention sum reader network," arXiv preprint arXiv:1603.01547, pp. 908-918, 2016.

[25] O. Vinyals, M. Fortunato and N. Jaitly, "Pointer networks," in Advances in Neural Information Processing Systems, 2015, pp. 2692-2700.

[26] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.

[27] H. B. Demuth, M. H. Beale, O. D. Jess, and M. T. Hagan, Neural Network Design: China Machine Press, 2002.

INTENTIONAL BLANK

A BASIS OF CULTURAL EDUCATION: SAFEGUARDING INTANGIBLE HERITAGE THROUGH A WEB-BASED DIGITAL PHOTOGRAPHIC COLLECTION

Chen Kim Lim[1], Kian Lam Tan[2] and Nguarije Hambira[3]

[1,2,3]Department of Computing, Faculty of Arts, Computing and Creative Industry, Universiti Pendidikan Sultan Idris (UPSI), 35900 Tanjung Malim, Perak, Malaysia

ABSTRACT

The 21st century came with its own challenges as much as it brought various benefits through the advancement in technology. Cultural heritage is one such "casualty" of the developments in the 21st century in that there has been a decline in appreciation and awareness of the importance of cultural heritage. Thus, the present study was necessitated with the primary aim of (i) preservation of the intangible cultural heritage of the people of Georgetown through the development of the E-George Town Digital Heritage (E-GDH) system (ii) develop an effective GUI for the E-DGH system in order to stimulate and captivate the attention of users with the aim of raising awareness as well to educate the masses on the importance of cultural heritage and (iii) to evaluate the effectiveness of the developed system in relation to its objective through the administration of questionnaires to target respondents. To this effect the study employed the use of the waterfall model to develop this E-GDH website. The study found that respondents (prior to using the E-GDH system) had no previous experience in terms of oral story telling from their parents. Overall, it was found that the GUI was pleasant and attractive for use by respondents and that they were able to learn easily as a result. Based on the fact that respondents where able to learn with ease due to an effective GUI, the study also revealed that the content they were learning from this website was actually easy for them to understand and that this website was indeed helpful in helping them to understand and appreciate cultural heritage. The meaning and conduct of the education sector in this era of advanced technology has shifted a lot over the years changing from teachers as the primary source of information to what is termed as "learner –centred" where they are given the leeway to learn, explore and make sense of the world around them and the findings from this study falls no short from this notion. The E-GDH website could be used by schools in subjects such as history where the teacher could use this website as reference point for a certain lesson outcome that deals with digital cultural heritage or intangible cultural heritage. Thus the study contributes immensely to the understanding of cultural heritage by raising awareness as well as stimulating the inters of the young generation to appreciate and learn more about their cultural heritage. The prominence of digitalising the intangible cultural heritage cannot be emphasised enough as recent study has shown decline in interests in these area so the development of the E-GDH is one such positive call to action in response to UNESCO's 2003 call for preservation of intangible cultural heritage and by extension, educating and raising awareness on the importance of cultural heritage.

KEYWORDS

Cultural Education, Digital Library, Digital Cultural Heritage, Digital Asset Management System, Digital Preservation

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 117–130, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90311

1. INTRODUCTION

Culture education has to a lot with a type of learning that combines the beliefs, history, principles as well as the viewpoint of various people that comes from a variety of backgrounds that might be different not only culturally but also in terms of their way of life. Cultural heritage symbolizes the identity of the country while the awareness of this heritage among citizen illustrates the future of the country. The knowledge of the people about the history of a country's cultural heritage is very important to foster the love for their nation, especially in a multiracial nation such as Malaysia. According to Blake (2008) cultural heritage can be in the form of things you can physically touch or quantify (tangible cultural heritage) such as monuments or it can be in the form of things we cannot literally touch (intangible cultural heritage) such as oral storytelling, the traditional knowledge of a certain practice for instance surviving the scorching heat of the desert by using locally available materials such as plants and animal products. In the year 2003, the United Nations Educational, Scientific and Cultural Organization (UNESCO) acknowledged the importance of distinctly preserving the intangible cultural heritage the same way they do with tangible cultural heritage because intangible cultural heritage requires a different method of preservation. The preservation of intangible cultural heritage can be achieved in various ways and the one propagated by this study is that of digital preservation through the development of a web system. The issues that necessitated the present study were twofold: (i) there was a lack of interest among the younger generation to learn about their cultural heritage, (ii) there was also a lack of adequate online reading materials where people could read up on cultural heritage. Education in the 21st century has changed from what was once the accepted norm of "teaching and learning". In this new era of technology, education has been "decentralised" in the sense that it has now shift to become learner centred compared to previously when it was teacher centred. To this effect, development of this web system as propagated for by the present study aims to preserve and educate the importance of local cultural heritage of the people of Georgetown by stimulating interest in this subject as well to educate the masses through the content provided forth on this website where various aspects of the intangible cultural heritage of the people of Georgetown would be taught and in so doing preserve these artefacts while raising awareness on these issues. Developing detailed systems that digitally manage the assets of a given museum at first glance might appear to be a difficult task but it does not necessarily have to. The main function of this system was to provide a solution to expose people to the wealth of cultural heritage in Malaysia. This research aims to capitalize on the fact that recent research has proven that there has been a drastic increase in the use of online content as a way to access information among internet users. The system provides a service whereby cultural heritage that is photographic in nature could be accessed digitally. Besides, the user would be able to use the search function to find the information they need.

Usually, users get their knowledge from oral stories taught by their grandparents as well as previous generation. Tebeau (2013) and Matusiak, Tyler, Newton and Polepeddi (2017) further expounded on this notion by highlighting that oral history plays a pivotal role of ensuring that the public is made aware and able to understand the origin as well as the way local communities socialise among each other. Thus this study was carried out with these primary objectives; (i) to preserve the cultural heritage of people at Georgetown through the development of the E-GDH system, (ii) ensure the effectiveness the E-GDH system through the development of an effective Graphic User Interface (GUI) that aims to attract and keep the young generation interests and (iii) evaluate the effectiveness of the E-DGH system through the deployment of questionnaire to target respondents. With the creation of the E-GDH, cultural heritage of the people of Georgetown, Penang can be preserved more effectively due to organization of content about cultural heritage and the user can also upload content to add more depth to the repository. Besides, the online discussions on E-GDH will keep users engaged substantially. The authors are of the strong believe that this study benefits the young generation by educating them on the importance of

cultural heritage through the interactive content on the E-DGH website. Consequently, having these cultural artefacts digitalised helps by contributing to the UNESCO goals of intangible cultural heritage preservation.

2. LITERATURE REVIEW

According to Olson (2013) the discipline of archaeology is one that is disruptive in nature and the methods that are employed by most archaeologists with their task of recording and preserving such findings represent visuals through printing of two dimensional of objects that appears in 360 degrees' viewpoint. However, there is hope for this phenomena thanks to advancement in the technology of 3D that stands to bring about an overhaul to the discipline of archaeology. Advancement in 3D technology has resulted in development of software packages that makes it possible for the generation of more precise 3D models. One such software to be used in the discipline of archaeology was that of PhotoScan Pro that was used to record and digitalise heritage artefacts with greater accuracy, allowing for a new avenue through which data in this discipline could be disseminated.

Al-Barakati et al., (2014) indicated that storage of media information through the process of compressing and encryption is normally associated with digital asset management systems (DAMS). For the purposes of detecting temperament of these media files or for the management of copyrights related issues, it then becomes necessary for these compressed files to be watermarked. However, watermarking these media files is not as easy as it seems seeing that the process of compressing media files reduces the bits' value, also encrypting these files in addition to compression, leads to randomization of the bit stream. As a result, any attempt to watermarking these files leads to loss of quality. To this effect, it then becomes of paramount importance to ensure that the appropriate encryption is chosen that will provide counter measures to issues concerning loss of quality. Al-Barakati et al., (2014) suggested the use of these schemes to watermark the media files: Rational Dither Modulation (RDM) and Spread Spectrum (SS). The summary of other related articles to digital asset management is as follows (Table 1).

Table 1. Summary of Articles Related to Digital Asset Management (DAM).

Name Of Paper (Authors, Years)	Brief Description	Strength(s)	Weakness(es)
The Application of Workflow Management to Digital Heritage Resources (Al-Barakati et. al, 2014)	An experimental Workflow management System (WfMS) whose aim were verifying the validity of the integration of management workflow in relation to Digital Heritage Resources (DHRs).	The difficulties brought about by implementing DHR was easily overcome through the deployment of the WFMS.	Implementing DHR "DISPLAYS" framework was only done as a testing for the WFMS.
Robustwatermarking of Compressed and Encrypted JPEG2000 Images (Subramanyam et al., 2012)	DAMS typically deals with compression and encryption of media files.	For the purposes of tempering detection as well as copyright issues, the watermarking of encrypted compressed media files becomes even more crucial.	Watermarking these encrypted files results in the loss of the quality of the media file.
Digital Curation Beyond the ''Wild	As an alternative to the commonly used theory	This alternative approach is said to view	The current digitalization

Frontier': A Pragmatic Approach (Dallas, 2016)	of digitalization, this study proposed a more logical approach as compared to the former.	digitalization as a practice where a range of actors such as artists, local communities and so on can use the system beyond the normal or rather expected life cycle.	strategy is limited in that the funding of applications, marketing, persuasiveness in attracting users as well as setting up of intellectual properties.
Streaming The Archives: Repurposing Systems to Advance a Small Media Digitization and Dissemination Program (Anderson, 2015)	The present study gives account of the journey taken by an archivist in their first year of using the system.	It was reported that the system was able to provide for good storage during the time of use.	Regardless of its inability to provide for responsive metadata, YouTube is seen to continue being the source of information distribution across various platforms.
The Tel Akko Total Archaeology Project (Akko, Israel): Assessing the Suitability of Multi-scale 3D Field Recording in Archaeology (Olson et al., 2013)	Due to successful testing, the use of PhotoScan Pro has been adopted for use in the Akko Project given its commercial success since launching.	The use of PhotoScan Pro for archaeological purposes has been proven to be a success due to its ability to increase the accuracy in terms of the way it records and digitalises heritage artefacts. In so doing, it provides for an alternative way that allows for the distribution of data in the area of archelogy.	The use of this software sometimes leads to loss of quality in the media files.
Camera-Based Whiteboard Reading For Understanding Mind Maps (Vajda et al., 2015)	The study was aimed at producing digitally illustrated mind maps which would allow for digitally managing assets through storage and retrieval of such records.	This system was founded on the idea of imaging through the use of cameras thereafter followed by extracting the whiteboard segments for any written texts as portrayed in the mind map.	The developed system was evaluated using only evaluation that was experimental and unconfined.
Descriptive Metadata: An Analysis of British Path Newsreel Collections from World War Two (Addica, 2017)	A digitalized system that allowed for the archiving and search functions.	The use of simple search phrases as well as a clearer scheme could have made the system user friendly by providing ease of access and use.	The conversion of the negative films into digital form lacked information that was complete and

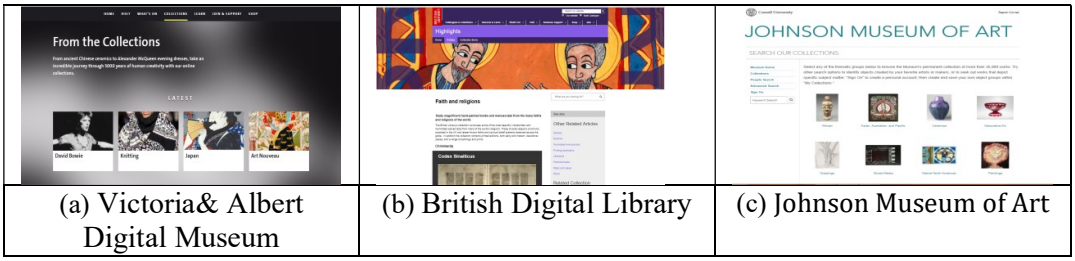
			detailed.
Systems Information Modelling: Enabling Digital Asset Management (Love et al., 2016)	The use of Computer-Aided-Design (CAD) provides managers with "on the go" documentation.	The use of CAD is not suitable for use across the board and this is evident with electrical engineering systems that involves a great deal of complexity due to its various components that incorporates the use of geometrical concepts that are not supported by CAD at the moment.	The optimization of performances when it comes to asset systems has been compromised due to the inability of having the required information at the required time.
Advocating for Sustainability: Scaling-Down Library Digital Infrastructure (Montoya, 2016)	The present study illustrated why those in charge of the library administration were ought to campaign systematically producing platforms that are scaled down in nature.	Environments keep changing technically and therefore it creates a need for libraries to adopt platform cultures that are more sustainable.	The nature of Content Management Systems (CMSs) is one that is complicated and forever evolving and this creates a scenario where infrastructures are not able to achieve the goals or rather the needs of small projects.
Digital Asset Management: Where to Start (McGovern, 2013)	The study proposed a framework that institutions could use to digitally manage their assets through planning which includes standards setting tasking as well as decision making.	The creation of a more complete document planner stands to provide institutions a cornerstone of all their digital programs as well as providing for a RFP in the future for a possible management system for their assets.	There is a need for DAM strategies that aims to tackle the distribution of digitalised assets that could be used in creating and maintenance of websites that are engaging in nature.

3. DIGITAL LIBRARIES

Ross (2012) highlighted the importance of ensuring that preservation of bit streams and the way it is represented by stating that within the context of digitalized collections maintenance of such collections should always be sort after. However, there are some challenges towards implementation of their goals of digitalizing their artefacts faced by many of the small-to-medium sized institutions that are built with the primary aim of cultural heritage. These challenges include issues such as reduced man power, reduced monetary support and infrastructures that technically are not compatible with latest technology. All these challenges reduce the institutions' capability of preserving their content digitally (Schumacher et al., 2014). There are a few case studies that are used in this research such as Victoria & Albert (V&A) Digital Museum, British Digital Library and Johnson Museum of Art (Figure 1). Boasting an amazing collection of well over two million artefacts that spreads over a period up to 5 000 years, the V and A is considered the front runner in terms of museums concerned with artistry and designs. This museum is home to a wealth of study materials for disciplines ranging from architecture, textile, book art to furniture and jewellery.

Another such library is that of the United Kingdom, it is the national library of the country and among the world's greatest of all times. With a collection up to 150 million artefacts that covers a wide range of well over 400 languages on top of up to 3 million pieces being added all year round. Their users can easily retrieve these collections, thanks to a library staff policy that allows for up to 30 minutes fully dedicated to tenting to users' enquiries as far as their collection goes (Schumacher et al., 2014).

Another museum that is equally committed to providing service to an audience that is diverse, is none other than the Herbert F. Johnson Museum of Arts. It opened its doors in the 1973 and it continued offering their services free of charge up until this very day. This museum does a great job when it comes to preserving its cultural artefacts for the present and future generations. The museum accomplishes this tasks through by connecting their audience with past arts and guides them in exploring latest developments in current arts.



(a) Victoria& Albert Digital Museum	(b) British Digital Library	(c) Johnson Museum of Art

Figure 1. Various Digital Museum (a) Victoria& Albert Digital Museum, (b) British Digital Library, (c) Johnson Museum of Art

4. METHODOLOGY

The present study adopted the use of the four phase waterfall model as the methodology for this research (Figure 2). The first phase was the requirement analysis and it entailed the collection of information through a survey questionnaire as well as interview sessions. Heritage photos characterising the people of Georgetown, Penang on 80-an were collected. These were than categorized into themes like race, type of attire, daily work and the facial appearance of people of Georgetown. The importance of the requirement phase could not be emphasised enough as the

overall functions and system limitations begins with this phase in terms of what the user expects from the system and what the completed system offers to the users.

In the design phase, the activities carried out were website design sketches, database designs, input and output designs, and website interface sketches that were used throughout the implementation of this website. The primary software used in developing this website was the Brackets, a free and open source code editor designed and developed by Adobe Systems. This platform was chosen due to its "live" editing abilities, that is to say with Brackets all changes made to the source code is directly reflected in the browser widow, in so doing it reduces the need for a developer to constantly keep refreshing the page to see the latest changes which saves a developers' time a great deal. Website design is said to be of utmost importance when it comes to translating business requirements into system designs that goes on to meet the needs of its intended users. To this effect the researchers saw to it that the design process was carefully executed in order to ensure that the finished website would be easily usable in providing complete and accurate information about the cultural heritage of the people of Georgetown. The design phase was divided into two sections; the first was more concerned with the database design whereas the second phase was primarily concerned with the graphic user interface (GUI) design. As for the development phase, the database was developed through the use of PhP MyAdmin and made use of the SQL language. This platform was chosen due to its widely available resources and tutorials, also it was chosen due to the fact that is freely available and thanks to a large community of developers and contributors it is considered relatively more secure. The E-GDH system required a database so as to keep, retrieve and manipulate cultural heritage of the people of Georgetown in real time. To connect the front-end (website GUI) to the back-end (database) the use of the PhP server based scripting language was adopted for use in this study.

The fourth and final phase included the evaluation of various components of the system such as the GUI evaluation by both the administrator as well as normal user respondents. Both of these evaluations were executed in terms of a questionnaire survey distributed to target respondents. The user evaluation was carried out on 20 randomly selected respondents from the Faculty of Human Sciences, Department of History who successfully completed and submitted the questionnaire survey.

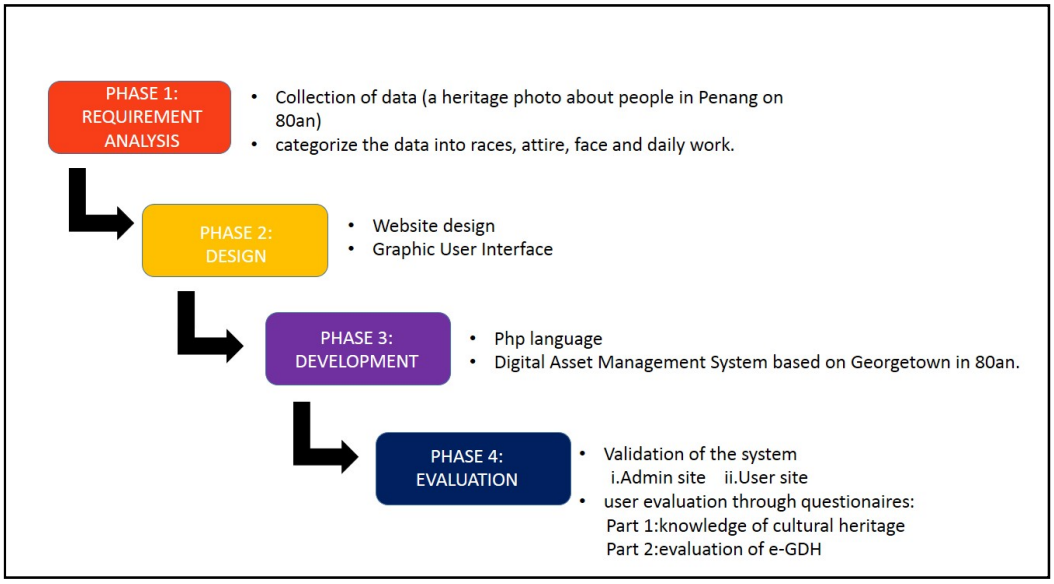


Figure 2. Research Methodology

5. DEVELOPMENT

A web-based digital asset management system based on Georgetown in 80-an was developed by using the PHP language (Figure 3). The index page contained the form element that allowed the administrator or normal user login options into the system (Figure 3, a). The system made use of the 'H' symbol to represent heritage. The researchers of the present study adopted the name electronic Georgetown digital heritage (E-GDH) as the official name for this system. There two main users for this system can either press the 'login as admin' if they are admin users which will redirect them to the admin page login session while 'login as user' button will direct to user login page. For the user logged in as admin, the main page will display the various functions they could perform (Figure 3b). When the admin clicks the button 'upload picture', the system will display an interface as shown in Figure 3(c). The page will request the admin to insert picture name, the description of the picture, select the race and type of attire, daily work as well as face. Next, the admin will have to click the button 'choose file' to select the image from their computers. The system will accept all type of image file format like JPEG, GIF, BMP, TIFF and PNG the admin could then press the submit button to finish this process. If the image was successfully inserted, the system message will display a message "Congratulations, the image was successfully uploaded" meanwhile if the process is not success the system will display a fail message. Cultural heritage photos can be viewed according to race such as British, Chinese, Indian and Malay (Figure 3 d-g). The system will display the title of the picture as well as its description. On the admin panel, each photo provides a button for update and delete the record. The searching function can help the user to get information that they want in a quick time (Figure 2i). The system will display a selection of the race and type of attire, daily work and face.

(a) Index Page of E-GDH | (b) Main Page of Admin of E-GDH | (c) Upload Image Page of E-GDH

(d) View Page of British Photos | (e) View Page of Chinese Photos | (f) View Page of Indian Photos

(g) View Page of Malay Photos | (h) Admin Update Photos Page | (i) User Searching Photos Page

Figure 3. A Walkthrough of a Web-based Digital Asset Management System (E-GDH)

6. EVALUATION

The purpose of this evaluation was to investigate if the objectives as outline in the first section of this paper has been achieved or not through the distribution of questionnaires to the target respondents. This study made use of students from Sultan Idris Education University as their primary target respondents in evaluating the E-GDH system. Twenty (20) respondents from the history education (AT32) course were able to complete and submit the evaluation form. Range of respondent's age ranged from 20 years old and below, 21-23 years old, 24-26 years old and 27 years old and above. The respondents were first exposed to the E-GDH system and after spending sometime interacting with the website they were asked to complete the questionnaire by providing evaluation on the usability of the system.

The questionnaire was made up of three sections, whereby the first section asked for demographic information of the respondents. The second section, was drawn up to test for the target respondents' basic knowledge about cultural heritage whereas the third section looked at evaluation of the E-GDH system. Upon completion of the data collection process, the data analysis was done through the use of IBM SPSS statistical software version 21. Statistical tests executed in SPSS were the mean and standard deviation tests. These test were tested on the Likert scale questionnaire items as indicated in the questionnaire survey. However, using descriptive analysis the tendency level was interpreted into several stages which were low (mean score: 1.00-2.40), average (mean score: 2.41-3.80) and high (mean score: 3.81-5.00).

Based on the section of basic cultural heritage knowledge, majority of the students (20) answered 'yes' that indeed they had a basic understanding of the cultural heritage concept (Figure 4). Majority (15) of the target respondents indicated that they did not have experience with oral story telling by their parents or elders compared to a minortiy (5) of the respondents who had experienced oral story telling from their elders. A slight majority (11) of the respondents had no knowledge of the importance of cultural heritage for the younger generation and the same holds true when it comes to the question item that asked if they knew the causes for low interest in cultural heritage by the younger generation, most respondents (15) indicated that they did not know what the factors were.

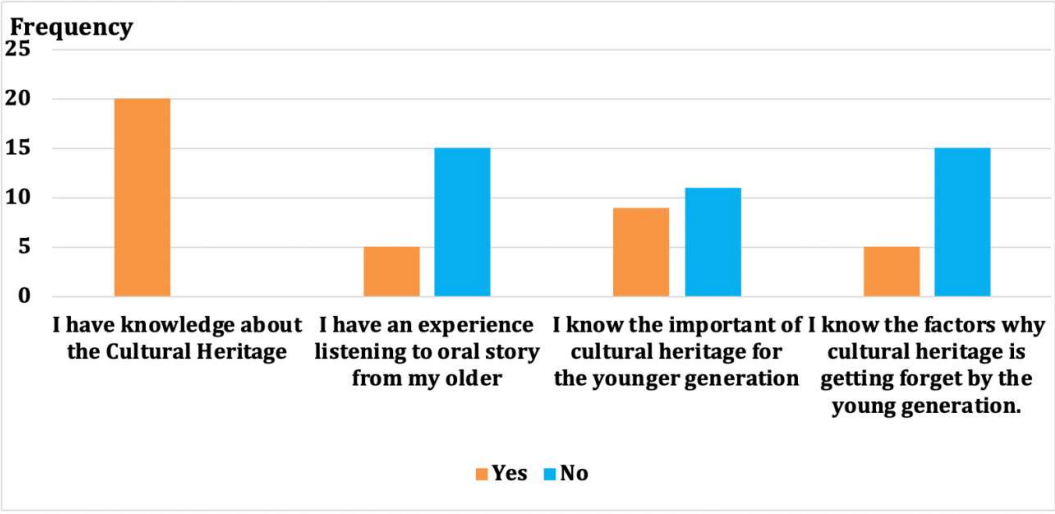


Figure 4. Evaluation for the Second Section Questionnaire

The third section was divided into three subsections that looked into the evaluation of the GUI (Figure 5), the content provided forth by the E-GDH system (Figure 6) and the third subsection looked at testing for effectiveness of the E-GDH system (Figure 7). This question items were meant to test for ease of use and overall appearance of the system from the target respondents.

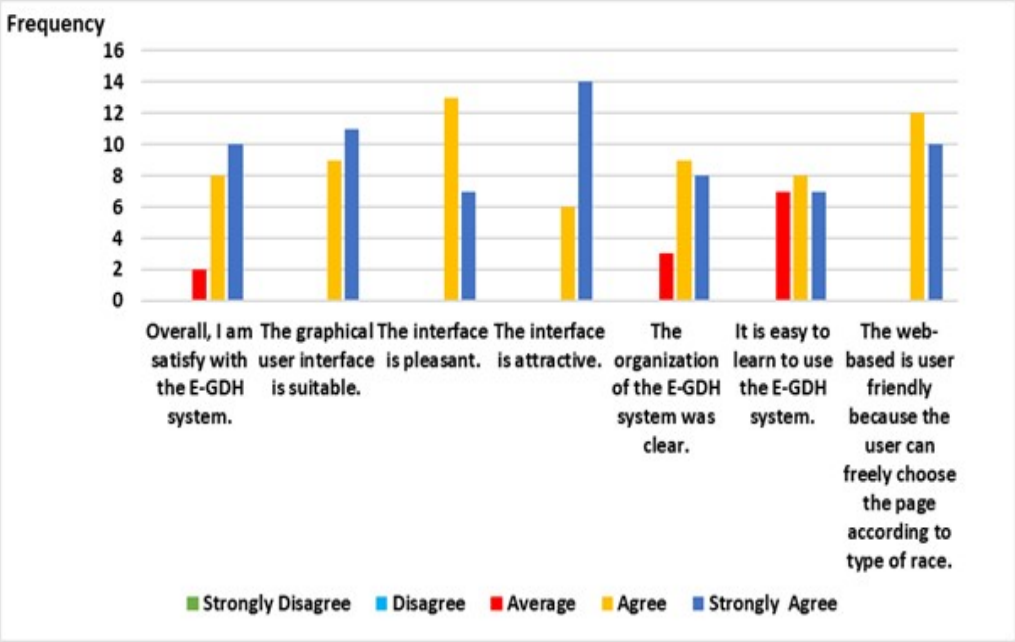


Figure 5. Evaluation of the Interface of the E-GDH System

From the question item that asked the respondents overall view on system (Figure 5), majority of the respondents either strongly agreed (10) or agreed (8) with the question item as compared to only two respondents who were recorded as average or neutral on the question item. For the question item that asked if the GUI was suitable for this purpose, majority of the all respondents either strongly agreed (11) or agreed (9) that the GUI was indeed suitable. The same holds true for the question item that asked if the respondents found the interface pleasant as most respondents either strongly agreed (7) or agreed (13) with the question item. The interface of the system was shown to be attractive as most respondents either strongly agreed (14) or agreed (6) with the corresponding question item. Similarly, organisation of the E-GDH system was said to be clear as shown by the responses (strongly agreed, 8; agreed 9). Majority of the target respondents (7 strongly agreed, 8 agreed) also indicated that it was easy to learn when it comes to using this system.

As for the second subsection that seek to evaluate the content provided forth by the E-GDH system, it was reported (Figure 6) that the content provided on this website was suitable for use by people of all ages as majority of the target respondents either strongly agreed (5) or agreed (8). The question item that asked if it was easy to understand the content on the E-GDH website showed that the content was easily understood as majority of the respondents either strongly agreed (8) or agreed (7) with the question item. The same holds true for the question item that asked if the E-GDH was a good source of knowledge acquisition, majority of the responses either strongly agreed (12) or agreed (8). Additionally, it was shown by the majority of responses (14 strongly agreed, 6 agreed) that the type of data on the site was suitable with the E-GDH system. The majority of target respondents (13 strongly agreed, 7 agreed) that the introduction of this system was helpful in helping the understanding of the cultural heritage.

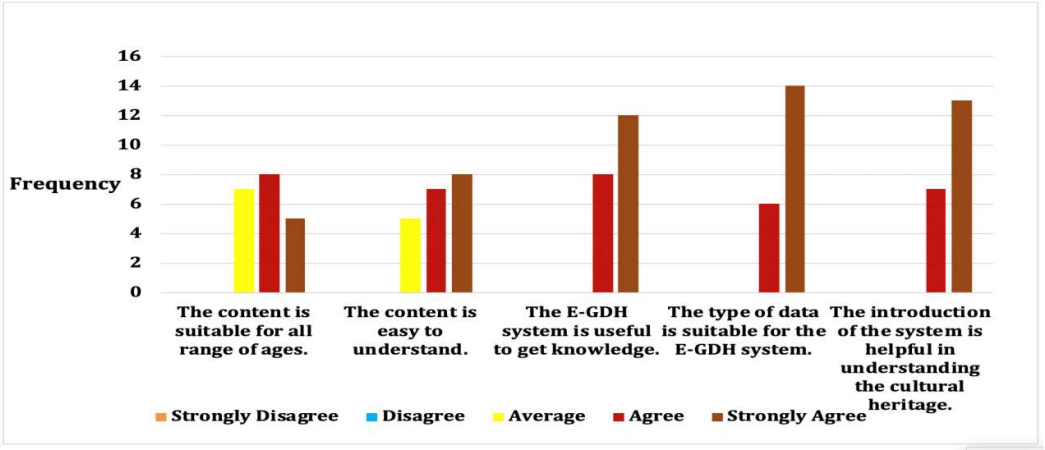


Figure 6. Evaluation of the Content on the E-GDH System

The third and final subsection of the three subsections under the evaluation category was concerned with evaluation of the effectiveness of the E-GDH system (Figure 7). As for the question item that asked if after using the system, respondents understood more about cultural heritage or not, majority of the target respondents either strongly agreed (6) or agreed (9) with the question item while 5 respondents were neutral or average on this question item. The same trend in responses was observed with the question item that asked if the E-GDH system was able to give knowledge about cultural heritage or not, majority of respondents to this effect either strongly agreed (8) or agreed (9) with the question item. Majority of the respondents were also able to differentiate facial differences between the various races as they either strongly agreed (13) or agreed (7) with the question item. Similarly, they were also able to distinguish between these races in terms of the attire they wore as majority of them either strongly agreed (7) or agreed (9) with the question item. As for the question that asked if respondents could differentiate between these races in terms of the type of work they did on a daily basis, majority of respondents either strongly agreed (11) or agreed (12). The last question item under this subtopic asked if respondents found the faces in synchronization with the type of attire, work done by each race or not and majority of responses to this effect either strongly agreed (11) or agreed (9) with the question item.

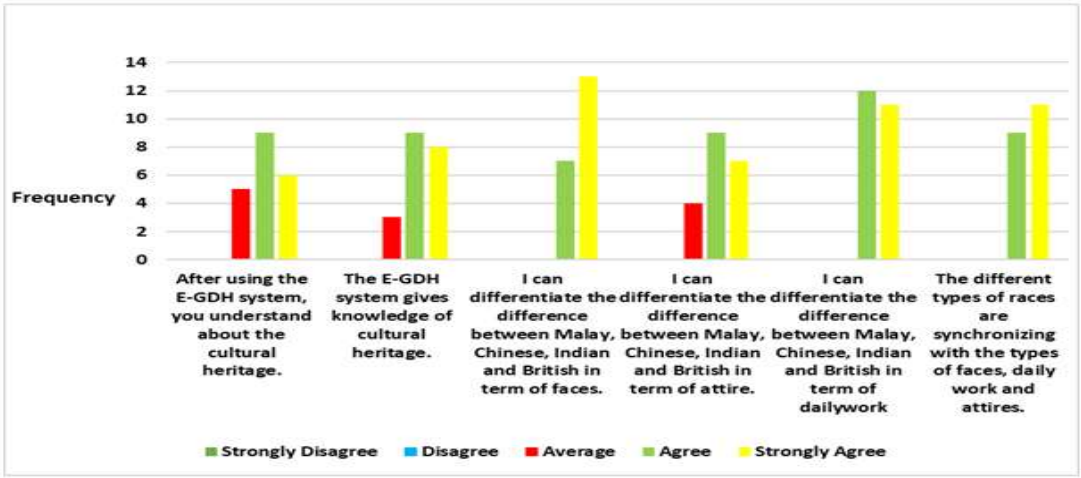


Figure 7. Evaluation of the Effectiveness of E-GDH System

In summation of the evaluation herein, it can be seen from the responses that overall respondents had general knowledge about cultural heritage, however, majority of the respondents (15) had no prior experience in listening to oral stories taught by their parents. From the overall question items that asked about the GUI of the E-GDH website the majority of responses in all these subsection items revealed that the graphic user interface (GUI) was satisfactory and that they had a pleasant time using the website as a result. It was also mentioned that due to overall good GUI design, it was then easy for them to learn with ease on how to use this system. As for the question items that asked respondents about the content used on the website majority of the respondents from all the sub question items indicated that they found the content suitable for use across all age groups and that the content was easily understandable. Most importantly the evaluation found that the E-GDH website was indeed a useful place for them to get knowledge about cultural heritage. The question items that asked if the E-GDH website was effective in educating the masses, responses indicated that overall the website was effective in educating them about cultural heritage as they were able to distinguish between various attires, type of work done by various races as depicted in the website.

7. CONCLUSIONS AND FUTURE WORK

In conclusion, digital archiving or digital documentary initiative is a long-term planning as one great alternative for cultural education. The development of the E-GDH system has enabled for the preservation of the cultural heritage in Georgetown. As the results from this study have indicated, the E-GDH website was designed to optimise learning and the content on the website was proved to be easy to learn and suitable. Results from this study have also shown that the respondents were able to differentiate between the different races based on their cultural heritage as portrayed in the website. This goes on to indicate that learning has taken place in a smooth and efficient manner that was both beneficial to the users of this online system as well as to the overall cultural heritage in Georgetown. As with any other academic studies, there were however limitations as far as the system functionality or target population are concerned. Due to the limitation in acquiring extensive data about the cultural heritage across Malaysia, the E-GDH system thus only placed focus on the people of Georgetown, Penang and the other limitation was the time required to analyse and select the most suitable data for display on the website. It is suggested that for future studies, development of E-GDH system or similar systems incorporate the availability of such systems on mobile platform seeing that recent research has shown that majority of web access is usually done through the use of mobile phones and mobile devices.

ACKNOWLEDGEMENTS

The authors would like to thank Nursyazwani binti Mohd Junik for developing the website.

REFERENCES

[1] Addica, K. (2017). Descriptive Metadata: An Analysis of British Pathé Newsreel Collections from World War Two. *Libri*.*67*(2). pp. 141-148.

[2] Al-Barakati, A., White, M., & Patoli, Z. (2014). The Application of Workflow Management to Digital Heritage Resources. *International Journal of Information Management*.34(5). pp. 660-671.

[3] Anderson, T. (2015). Streaming the Archives: Repurposing Systems to Advance a Small Media Digitization and Dissemination Program. *Journal of Electronic Resources Librarianship*.*27*(4). pp. 221-231.

[4] Blake, J. (2008). UNESCO's 2003 Convention on Intangible Cultural Heritage: The implications of community involvement in 'safeguarding'. In Intangible heritage (pp. 59-87). Routledge.

[5] Dallas, C. (2016). Digital Curation Beyond The "Wild Frontier": A Pragmatic Approach. *Archival Science*.16(4). pp.421-457.

[6] Love, P. E., Zhou, J., Matthews, J., & Luo, H. (2016). Systems Information Modelling: Enabling Digital Asset Management. *Advances in Engineering Software*. *102*.pp.155-165.

[7] Matusiak, K., Tyler, A., Newton, C., & Polepeddi. P. (2017). Finding Access and Digital Preservation Solutions for A Digitized Oral History Project. *Digital Library Perspectives*. 33(2). pp. 88-99.

[8] McGovern, M. (2013). Digital Asset Management: Where to Start. *Curator: The Museum Journal*. *56*(2). pp.237-254.

[9] Montoya, R. D. (2016). Advocating for Sustainability: Scaling-Down Library Digital Infrastructure. *Journal of Library Administration*.*56*(5). pp.603-620.

[10] Olson, B., Placchetti, R., Quartermaine, J., & Killebrew, A. (2013). The Tel Akko Total Archaeology Project (Akko, Israel): Assessing The Suitability of Multi-Scale 3d Field Recording in Archaeology. *Journal of Field Archaeology*. 38(3). pp. 244-262.

[11] Ross, S. (2012). Digital preservation, archival science and methodological foundations for digital libraries. New Review of Information Networking, 17(1), 43-68.

[12] Schumacher, J., Thomas, L. M., VandeCreek, D., Erdman, S., Hancks, J., Haykal, A., ... & Spalenka, D. (2014). From theory to action: "good enough" Digital preservation solutions for under-resourced cultural heritage institutions. The Digital POWRR (Preserving Digital Objects with Restricted Resources) White Paper. Retrieved from: http://commons.lib.niu.edu/handle/10843/13610

[13] Subramanyam, A. V., Emmanuel, S., & Kankanhalli, M. S. (2012). Robust Watermarking of Compressed and Encrypted JPEG2000 Images. *IEEE Transactions On Multimedia*. 14(3). pp. 703-716.

[14] Tebeau, M. (2013). Listening to The City: Oral History and Place in The Digital Era. *Oral History Review*.40 (1). pp 25-35.

[15] Vajda, S., Plötz, T., & Fink, G. A. (2015). Camera-Based Whiteboard Reading for Understanding Mind Maps. *International Journal of Pattern Recognition and Artificial Intelligence*.*29*(03). pp. 1553003.

AUTHORS

C. K. LIM has earned her doctorate from the School of Computer Sciences at Universiti Sains Malaysia (USM) in 2014 with the attachment in Université Joseph Fourier (UJF), France under the MFUC scholarship (2012-2014) at INRIA Grenoble. Currently, she is a senior lecturer in the Computing Department, FSKIK, Sultan Idris Educational University. She previously served as a deputy director II for assessment and instrumentation at UPSI Educational Research Lab (UERL). She specializes in computer graphics, digital heritage, visual informatics, crowd simulation and rendering. She has published various proceedings and journal articles.

K. L. Tan has worked as a Research Scientist at Universite Grenoble Alpes focusing in the area of lifelogging and tourism by using the technology of deep learning and social network in 2016. He was partially sponsored by MFUC to pursue a Joint-Ph.D. program at Universite Grenoble Alpes (UGA) and Universiti Sains Malaysia (USM) and graduated in 2014. In 2008, he worked for Intel as a software engineer for almost 2 years and successfully deployed several automation software solutions in the United States, India, and China. His research interests spams across Information Retrieval (IR), Big Data Analysis (BDA) and mobile game. Tan is presently employed as a senior lecturer at Sultan Idris Education University. He has published more than 35 proceedings and journal articles. He also acts as an advisor in the National Child Data Centre (NCDC).

N. Hambira is an aspiring young researcher with an immense interest in the area of information technology and its integration in the education sector. He obtained his honours degree from Sultan Idris Education University in 2017 and has gone on to publish over 5 conference proceedings and journal papers. He is presently pursuing his master's degree in education with specialisation in information technology. He has extensive experience in research and publication gained while working as a research assistant during his undergraduate as well as postgraduate studies at Sultan Idris Education University.

AN ENVIRONMENT-VISUALIZATION SYSTEM WITH IMAGE-BASED RETRIEVAL AND DISTANCE CALCULATION METHOD

Yuka Toyoshima[1], Yasuhiro Hayashi[2] and Yasushi Kiyoki[1]

[1]Graduate School of Media and Governance, Keio University, Kanagawa, Japan
[2]Faculity of Economics, Musashino University, Tokyo, Japan

ABSTRACT

Many environmental issues have occured in this world and these issues are common to all human beings. It is considered that environmental issues caused by humans exist in the "border" between nature and human society. In other words, there is the possibility that finding the "border" leads to determine the cause of environmental issues and discover the solution. This paper presents an environment-visualization system with image-based retrieval and distance calculation method as the first step of research for finding the "border". We focused on the plastic garbage issue which is related to SDGs14 and this study was made to find the "border", source of the plastic garbage which is scattered on the coast area. In addition, we aim to realize the system which enables people to share the knowledge about the plastic issue in order to acquire knowledge of the environment issues and to promote concrete action to realize sustainable nature and society. In this system, there are 3 features: (1) Composition-Based Image Retrieval Function, (2) Spatio-Temporal-Based Mapping Function, and (3) Coast-area Location-Checking for Selected Images Function. (1) is the image retrieval function for detecting highly-related images to a query-image with dividing one image to three images for separating with nature, human society, and "border". We used euclid calculation to calculate the similarity and show the results in the ranking format. (2) is the mapping function with using the spatio-temporal information which is accompanied with images. (3) is the location-checking function to judge whether the photographing spot is near the ocean or not with image processing metric and select the images only which are near the ocean.We present several experimental results to clarify the feasibility and effectiveness of our method.

KEYWORDS

Environment-Visualization, Image-based Retrieval, Image Processing, Distance Calculation, SDGs (Sustainable Development Goals)

1. INTRODUCTION

Many environmental issues have occured in this world and these issues are common to all human beings. It is considered that environmental issues caused by humans exist in the "border" between nature and human society. In other words, there is the possibility that finding the "border" leads to determine the cause of environmental issues and discover the solution.

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 131–146, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90312

In this paper, we present an environment-visualization system with image-based retrieval and distance calculation method as the first step of research for finding the “border”. This method retrieve the images which is divided with 3 parts: nature, human society, and the border between nature and human society. This system enables to recognize the “meaning” of the envrionmental situations and phenomena from each images and searches the images which is similar with a query-image as a human with adopting the method of division of images with 3 parts based on composition. Therefore, this system enables to detect the specific environmental situation and phenomenon corresponding to the user’s imagination. In addition, this method calculates the distance between the photographing spot and object with metadata accompanied with images and reflects the calculation results to the result of image retrieval. Therefore, it is possible to search more similar images than previous image retrieval.

In this paper, we focused on the issue of plastic garbage as the environmental issue to work on. Plastic garbage issue is getting worse year by year. According to the report about marine debris which was announced in World Economic Forum Annual meeting in Davos [1], at least 8 million ton of plastics end up in the ocean. This is equal to one garbage truck of plastics ends up in the ocean every minute. If we do not do anything, or do not try to solve this problem, the amount of plastics which flow into the ocean will be 2times in 2030 and 4 times in 2050. Figure 1 shows the yearly comparison of the amount of plastics flowing into the ocean in one year.

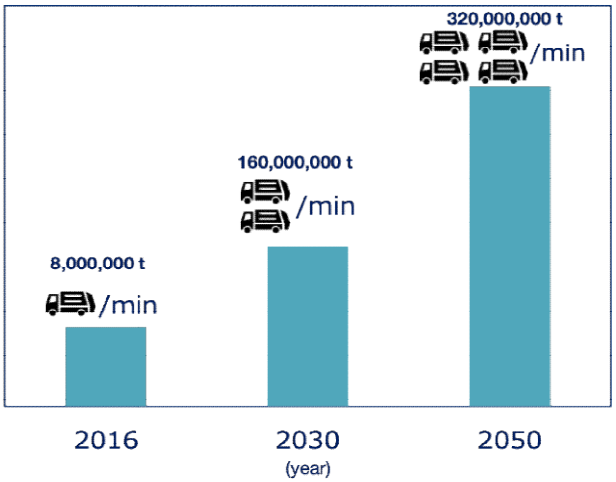


Figure 1. The amount of plastics which flow into the ocean in one year [1]

For this reason, in this paper, we set the issue as "how to reduce marine garbage and microplastic". This issue is corresponding to SDGs 14. Sustainable Development Goals (SDGs) [2] are summarized about environment issues. Future Learning should support SDGs in order to acquire knowledge of the environment issues and to promote concrete action to realize sustainable nature and society. SDGs are 17 goals defined by United Nations to solve the issues about society and environment until 2030. Figure 2 shows the goals of SDGs.

1. No Poverty
2. Zero Hunger
3. Good Health and Well-Being
4. Quality Education
5. Gender Equality

6. Clean Water and Sanitation
7. Affordable and Clean Energy
8. Decent Work and Economic Growth
9. Industry, Innovation and Infrastructure
10. Reduced Inequalities
11. Sustainable Cities and Communities
12. Responsible Composition and Production
13. Climate Action
14. Life Below Water
15. Life On Land
16. Peace, Justice and Strong Institutions
17. Partnerships for the Goals

Figure 2. SDGs: Sustainable Development Goals [2]

In this paper, we propose the method leading to the solution of the problem by the method of image processing. In addition, we aim to realize the system which enables people and organization to share the seriousness and knowledge of plastic garbage issue and to discuss and take some action for solving or improving the issue.

2. RELATED WORK

In this section, we would introduce the researches related our study.

2.1. 5D WORLD MAP SYSTEM

There is the system of image retriever and visualization about environment: "5D World Map" [3][4][5][6][7]. 5D World Map system [3][4][5][6][7] is a collaborative knowledge sharing system that enables to create and share knowledge by visualizing with a map. This system analyses multimedia such as images, videos, audio, documents, etc., by semantic, temporal and spatial information. In addition, this system integrates and visualizes the analyzed results as a 5- dimentional dynamic historic atlas (5D World Map Set). The main feature of this system is to create various context–dependent patterns of environmental/historical/cultural stories according to a user's viewpoints dynamically. For example, when users would like to find the images about forest fire and understand the place and time of the phenomenon of forest fire, they can search by keywords, images, or setting the target databases.

The composition of images is promising to be supported in 5D World Map, as the image retriever of this system. This system enables to realize that users find out the images that show the particular situation of environment such as "forest fire happened in the night".

2.2. IMAGE QUERY CREATION METHOD

There is an image query creation method of using multiple images [8][9]. This method creates the query image to express user's intentions by combining the multiple images based on colors and shapes of objects. This method enables users to express their imagination and intention by multiple images and create the exact image query for each user.
Our method focuses on the particular environmental situation and phenomenon as user's imagination and intention.

3. THE CONCEPT AND STRUCTURE OF OUR SYSTEM

In this section, we explain about the concept and structure of our system. Figure 3 shows the concept of our system.

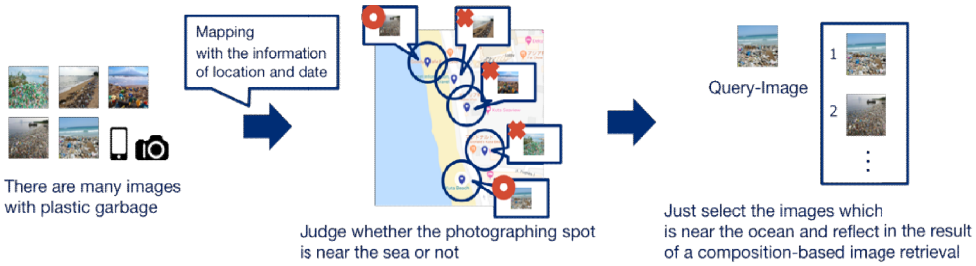


Figure 3. The concept of our system

There are many images with plastic garbage and also these images have the information of plastic such as the information of location and date. With these information, this system mapping the images like this and judging whether the distance between the photographing spot and the sea is close or not. The blue pin is the photographing spot and it is the center of this circle with a radius *r*. If the color of ocean is included in the circle, this system judge "This spot is near the sea". If not, "This spot is far from the sea".

This system just selects the images which is near the ocean and reflects in the result of a composition-based image retrieval.

In addition, Figure 4 shows the system structure.

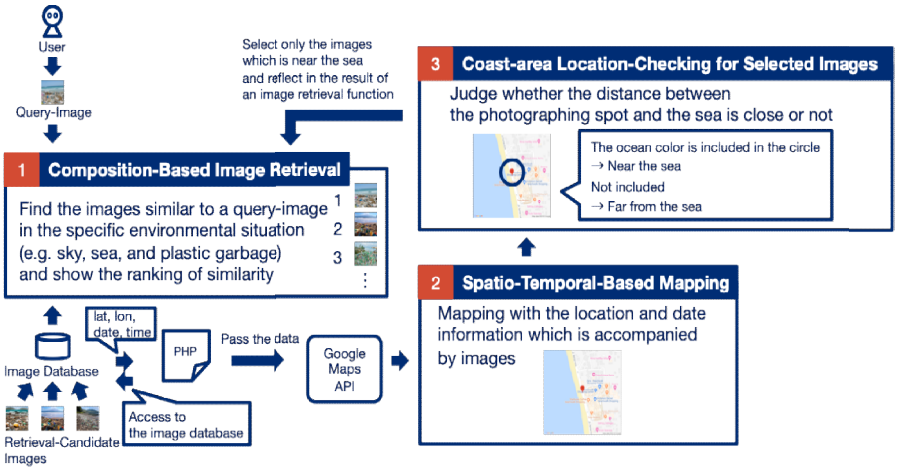


Figure 4. System Structure

In this system, there are 3 features: (1) Composition-Based Image Retrieval Function, (2) Spatio-Temporal-Based Mapping Function, and (3) Coast-area Location-Checking for Selected Images Function.

(1) is the image retrieval function for detecting highly-related images to a query-image with dividing one image to three images for separating with nature, human society, and "border". We used euclid calculation to calculate the similarity and show the results in the ranking

format. (2) is the mapping function with using the spatio-temporal information which is accompanied with images. (3) is the location-checking function to judge whether the photographing spot is near the ocean or not with image processing metric and select the images only which are near the ocean.

4. AN ENVIRONMENT-VISUALIZATOIN SYSTEM WITH IMAGE-BASED RETRIEVAL AND DISTANCE CALCULATION METHOD

In this section, we would mention the implement of our method.

In this method, each "retrieval-candidate image" is divided into three parts, Top-Part, Middle- Part and Bottom-Part. A "query-image" is also constructed with three parts as Top-Part, Middle- Part and Bottom-Part. When the image retrieval is applied, this method compares each corresponding part between query-image and retrieval-candidate image by correlation computing, and makes ranking of retrieval-candidate images according to correlations. In addition, our method creates the map image in which Google Maps [10] makes a pin with the spatio-temporal information of the image and to judge whether the photographing spot is near the ocean or not with image processing metric.

The steps of detailing our method are as follows:

1) **Decide retrieval-candidate images (m)**
We decided retrieval-candidate images (m) as the retrieval objects.

2) Decide the number (n) of colors which we handle and make a cluster of n colors
We defined the number (n) of colors and created the cluster of n colors to standardize colors of m images.

3) Divide one image into three images
We divided one image (300×300 pixels) into three images to consider the composition of each image. Figure 5 shows the example of divided-part-images in a single image.

Figure 5. Divided-part-images in a single image

In this method, each image is divided into Top-Part, Middle-Part, and Bottom-Part.

4) Create a color histogram with each divided-part-image
We created a color histogram with each divided-part-image. We extracted each pixel color from each image and put it in RGB space. Then, we converted RGB space to HSVspace. Equation

(1)~(5) show the formula to convert RGB into HSV. The range of RGB values is 0 to 1. In addition, max and min are the maximum and minimum values of RGB

$$H = \frac{G-B}{\max - min} \times 60, IF \max = R \quad (1)$$

$$H = \frac{B-R}{\max - min} \times 60 + 120, IF \max = G \quad (2)$$

$$H = \frac{R-G}{\max - min} \times 60 + 240, IF \max = B \quad (3)$$

$$S = \frac{\max - min}{max} \quad (4)$$

$$V = max \quad (5)$$

After that, for the distance calculation in HSV space, we used the Godlove color difference formula [11]. Equation (6) shows Godlove color difference formula.

$$\Delta godlove\ (H_1, S_1, V_1, H_2, S_2, V_2) = \frac{2S_1S_2\left(1-\cos\left(2\pi\frac{|H_1-H_2|}{100}\right)\right)+(|S_1-S_2|)^2+(4|V_1-V_2|)^2}{2} \quad (6)$$

5) Use Euclid Distance Calculation

We used the Euclid distance calculation to show the color distance between a query-image and retrieval-candidate images (m). Equation (7) is the Euclid distance calculation.

$$h1, h2, h3 = \sum_{i=1}^{n}\sqrt{(C_{qi} - C_{mi})^2} \quad (7)$$

Each value of color histograms for divided-part-images (*h1, h2, h3*) is calculated by the sum of the color distance between a query-image (c_{qi}) and retrieval-candidate images (m) (c_{mi}). The smaller result is the higher similarity.

6)Use the method to integrate each value of color histogram for divided-part-images

We used the method to integrate each value of color histogram for divided-part-images. Equation (8) is a calculation to integrate them.

$$Th = h1 + h2 + h3 \quad (8)$$

Total value of color histograms (*Th*) is calculated by the sum of each value of color histograms for divided-part-images (*h1, h2, h3*).

7) Use the method of weighting

We used the method of weighting to emphasize the feature part (plastic garbage part) in images. Equation (9) is the calculation to multiple the weight of each part.

$$Th\ with\ weight = h1 \times w1 + h2 \times w2 + h3 \times w3 \quad (9)$$

Each value of color histogram for divided images (h1, h2, h3) is multiplied by each weight (w1,w2, w3).

8) Mapping images onto a map and capturing the screen image
The retrieved images that are included latitude and longitude information as image metadata are mapped onto a map. In this system, Google Maps is applied for mapping. Afterwards, this system captures a screen image of Google Maps which makes a pin in a map.

9) Judging a coastal area with location information by image-processing
We extracted the color of each pixel from the map image created in the step 8) and judged the image was taken in the coastal area which is near the sea or not. As a judging method, this system judge based on whether the color of the ocean is included in the circle whose radius is r and center point is the photographing spot.

5. EXPERIMENTS AND DISCUSSION

In this section, we would mention experiments and discussion.

We conducted three experiments. Experiment 1 has two experiments. Experiment 1-1 is the experiment to compare the previous system without composition and our system with composition. Experiment 1-2 is the experiment to verify the weighting. Experiment 2 is the basic experiment for judgment of image composition automatically. Experiment 3 is the basic experiment for verifying whether the distance sense of the human being when they see the map image and the system judgment are close or not.

5.1. Experiment 1

In this experiment, we constructed an own image database and retrieve images by following the steps as mentioned above. We decided 50 retrieval-candidate images. Figure 6 shows all retrieval-candidate images.

Figure 6. Retrieval-candidate images in Experiment 1

We selected these images by Flickr [12] and Google Images [13]. We treated 5 images within 50 retrieval-candidate images as expected images to be retrieve in this experiment. We defined expected images as the images which are composed with sky, sea, and plastic garbage scattered on the beach. From figure 6, the image surrounded by red rectangle is expected image.

We selected other images according to the following conditions:

- The image with only plastic garbage
- The image excluded plastic garbage
- The image excluded plastic garbage with different color of sky
- The image which has multiple colors in top or center region

Additionally, we handled 130 colors based on Color Image Scale [14].

There are two parts in this experiment.

5.1.1. Experiment 1-1: Comparing With Composition and Without Composition

In this experiment, we compare the previous system without composition and our system with composition.

Table 1 and 2 show the results of the top-ten ranking of 50 retrieval-candidate images. We used the id number 25 as a query-image and make the best similarity on purpose. In addition, we did not use the weighting method in this experiment.

Table 1 shows the result of divided-part-images.

Table 1. Result of divided-part-images

img_id	*Th*
25	0
10	0.715803303
1	0.833790638
6	0.841656171
5	0.86161129
7	0.895032013
3	0.989218226
37	0.991339542
50	0.991339542
42	1.024072877

There are 2 columns, img_id and Th. Img_id is the id number of each image, and Th is the sum of color histogram values of divided-part-images. Blue color is the id number of expected images. From Table 1, we can understand that the most similar image is 25. There are also 10 in the second, 1 in the third, and 3 in the seventh.

Table 2 shows the result of non-divided images.

Table 2. Result of non-divided-images

img_id	*Th*
25	0
10	0.156998147
1	0.163591752
4	0.208711166
37	0.218768187
6	0.220653618
7	0.223305956
3	0.225188325
5	0.227740912
27	0.24074918

There are 2 columns, img_id and H. Img_id is the id number of each image, and H is the sum of color histogram values of a non-divided image. Blue color is the id number of expected images. From Table 2 we can understand that the most similar image is 25. There are also 10 in the second, 1 in the third, and 3 in the eighth.

Comparing with Table 1 and Table 2, it is able to say that the images which has similar color variance value be reflected to the result because it is possible to detect the volume of the color variance value for each part to divide with 3 parts based on the image composition.

5.1.2. Experiment 1-2: Verifing Weighting

In this experiment, we verify the method of weighting to show how the ranking is going to be changed.

We changed the ratio of weight for each part. The ratio of weight for each part is 1:1:N and 1:2:N (N > 2) to emphasize the bottom part (the plastic garbage part).

We have 11 ways to weighting. However, we show the 4 results as Table 3~6 in this paper. Table 3~6 show the results of the top-ten ranking of 50 retrieval-candidate images. We used the id number 25 as a query-image and make the best similarity on purpose.

Table 3 shows the result of the case of 1:1:2.

Table 3. Result of the case of 1:1:2

img_id	*Th*
25	0
10	0.20612656
6	0.234365182
5	0.246331578
7	0.261774284
1	0.263186217
37	0.292710457
42	0.297620847
3	0.299138796
50	0.30366452

There are 2 columns, img_id and Th with weight. Img_id is the id number of each image, and Th with weight is the sum of color histogram values of divided-part-images with weight. Blue color is the id number of expected images. From Table 3, we can understand that the most similar image is 25. There are also 10 in the second, 1 in the third, and 3 in the nineth.

Table 4 shows the result of the case of 1:1:6.

Table 4. Result of the case of 1:1:6

img_id	*Th*
25	0
10	0.15741475
6	0.16508487
5	0.1850233
7	0.206919704
40	0.228159007
42	0.232015678
37	0.236106372
1	0.241070224
50	0.247229191

There are 2 columns, img_id and Th with weight. Img_id is the id number of each image, and Th with weight is the sum of color histogram values of divided-part-images with weight. Blue color is the id number of expected images. From Table 4, we can understand that the most similar image is 25. There are also 10 in the second and 1 in the nineth.

Table 5 shows the result of the case of 1:2:3.

Table 5. Result of the case of 1:2:3

img_id	*Th*
25	0
10	0.289174828
6	0.293847877
5	0.333508466
7	0.38134804
40	0.428124782
42	0.428125157
37	0.441694257
50	0.456333111
47	0.462797156

There are 2 columns, img_id and Th with weight. Img_id is the id number of each image, and Th with weight is the sum of color histogram values of divided-part-images with weight. Blue color is the id number of expected images. From Table 5, we can understand that the most similar image is 25. There are also 10 in the second.

Table 6 shows the result of the case of 1:2:7.

Table 6. Result of the case of 1:2:7

img_id	*Th*
25	0
10	0.151764309
6	0.157147815
5	0.175362075
7	0.198395799
42	0.223592992
40	0.226733782
37	0.229116097
1	0.235441637
50	0.235641094

There are 2 columns, img_id and Th with weight. Img_id is the id number of each image, and Th with weight is the sum of color histogram values of divided-part-images with weight. Blue color is the id number of expected images. From Table 6, we can understand that the most similar image is 25. There are also 10 in the second and 1 in the nineth.

From the experimental results, it is able to say that the part to be emphasized is emphasized properly by weighting in our system. In addition, we found the maximum value of weighting. In the case of 1:1:N, N=6 is the maximum value, and in the case of 1:2:N, N=7 is the maximum value.

5.2. Experiment 2

This experiment is the basic experiment for judgment of image composition automatically. In this experiment, we constructed an own image database and retrieve images by following the steps as mentioned above. We decided 75 retrieval-candidate images which are adding 25 images to the retrieval-candidate images of Experiment 1. Figure 7 shows all retrieval-candidate images.

Figure 7. Retrieval-candidate images in Experiment 2

We selected these images by Google Images [13] and test images [15] which are usually used for the image processing research. We treated 6 images within 75 retrieval-candidate images as expected images to be retrieve in this experiment. We defined expected images as the images which are composed with sky, sea, and plastic garbage scattered on the beach and taken from left diagonal. From figure 7, the image surrounded by red rectangle is expected image. We selected other images according to the conditions which excluded plastic garbage.

In addition, we defined the 16 composition patterns. Figure 8 shows the all patterns of image composition.

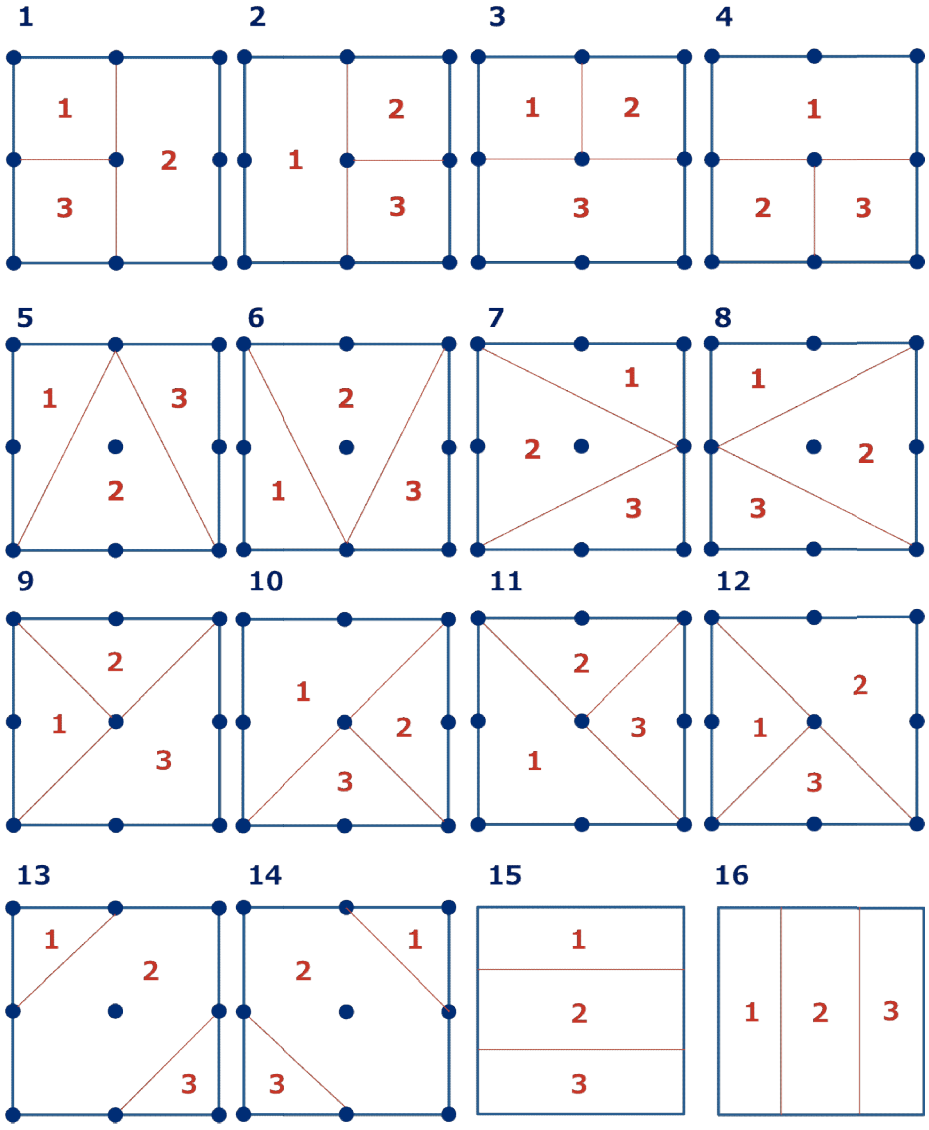


Figure 8. 16 composition patterns

Top part is 1, middle part is 2, and bottom part is 3. Pattern 15 and 16 are the exception case. However, we included the composition patterns because they are basic one as image composition.

We have 16 results. However, we show the 3 results which have higher precision as Table 7~ 9 in this paper. Table 7~ 9 show the results of the top-ten ranking of 75 retrieval-candidate images. We used the id number 51 as a query-image and make the best similarity on purpose. In addition, we define composition pattern 7 is the best way to divide and search the expected images.

Table 7 shows the results of composition pattern 7.

Table 7. Result of composition pattern 7

img_id	*Th*
51	0
7	0.54080188
5	0.618459439
6	0.620064148
2	0.641678675
3	0.643087056
50	0.696883611
65	0.737541973
56	0.761567507
37	0.772267028

There are 2 columns, img_id and Th. Img_id is the id number of each image, and Th is the sum of color histogram values of divided-part-images. Blue color is the id number of expected images. From Table 7, we can understand that the most similar image is 51. There are also 56 in the nineth.

Table 8 shows the results of composition pattern 13.

Table 8. Result of composition pattern 13

img_id	*Th*
51	0
6	0.597303954
7	0.700849855
2	0.710111445
3	0.784299714
5	0.81356392
50	0.821740533
37	0.874531216
55	0.878133988
42	0.882020918

There are 2 columns, img_id and Th. Img_id is the id number of each image, and Th is the sum of color histogram values of divided-part-images. Blue color is the id number of expected images. From Table 8, we can understand that the most similar image is 51. There are also 55 in the nineth.

Table 9 shows the results of composition pattern 14.

Table 9. Result of composition pattern 14

img_id	*Th*
51	0
7	0.705883744
5	0.726226913
3	0.733439106
70	0.75626755
65	0.787828013
2	0.815442295
6	0.818843453
56	0.842761986
37	0.844634256

There are 2 columns, img_id and Th. Img_id is the id number of each image, and Th is the sum of color histogram values of divided-part-images. Blue color is the id number of expected images. From Table 9, we can understand that the most similar image is 51. There are also 56 in the nineth.

Comparing the results of each composition pattern, it is able to say that our system does not detect the features of images with high precision. To improve this, it is necessary to acquire feature quantity from line component with hough transform technology.

5.3. Experiment 3

This experiment is the basic experiment for verifying whether the distance sense of the human being when they see the map image and the system judgment are close or not.

We defined the photographing spot as Kuta Beach in Bali island and set the longtitude and latitude as {-8.71669, 115.168583}. We obtained the map image of 2560×1380 pixels whose scale is 1/100 meters. Figure 9 shows the map image we used in this experiment.

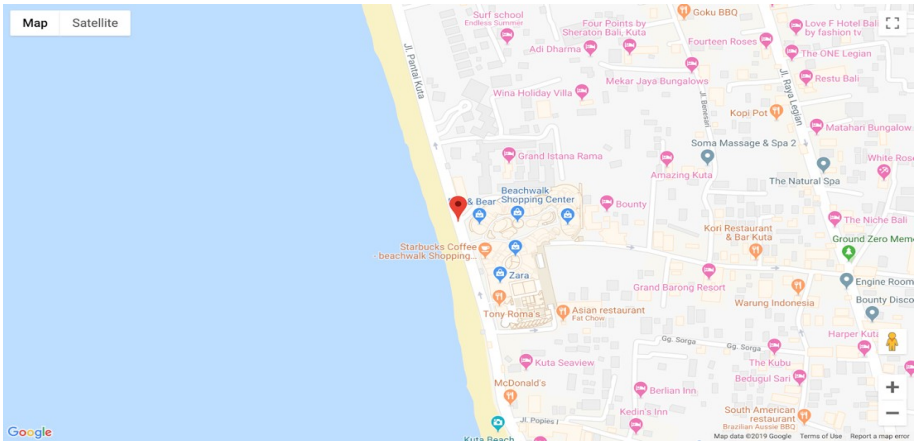


Figure 9. Map image of Kuta Beach

In addition, we defined radius r within 200 pixels. We set the color to detect based on RGB as RGB = (191, 222, 254). If this color is included in the circle, the console shows “Ocean!”. If not, the console shows “Not Ocean!”.

We judge the distance between photographing spot and ocean is close and we verify whether the distance sense of the human being when they see the map image and the system judgment are close or not.

The console showed “Ocean!”. In other words, our system judge “the photographing spot is near the ocean”. Figure 10 shows the result of console.

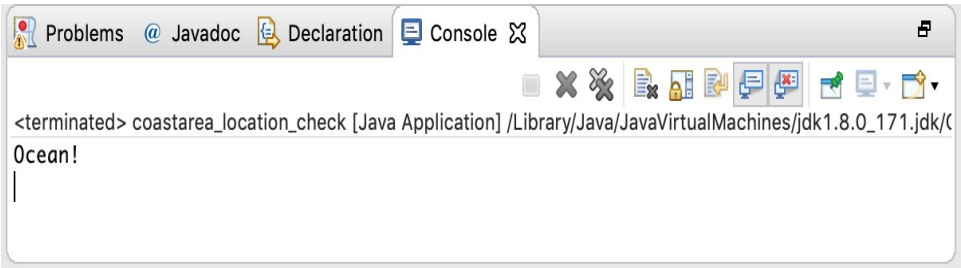


Figure 10. Result of Console

From the experiment result, it is able to say that the distance sense of the human being when they see the map image and the system judgment are close.

However, we have 2 points to improve the precision. The first point is to make it possible to deal with other scales of map images for changing the pixels of radius r. The second point is to acquire automatically metadata such as location information from the image database.

6. CONCLUSION AND FUTURE WORKS

In this paper, we have presented an environment-visualization system with image-based retrieval and distance calculation method. We have shown about 3 functions of this system: (1) Composition-Based Image Retrieval Function, (2) Spatio-Temporal-Based Mapping Function, and (3) Coast-area Location-Checking for Selected Images Function. In addition, we have presented several experimental results about 3 functions to clarify the feasibility and effectiveness of our method.

From the experimental results and discussion, we obtained 4 opinions as follows:

- Our method using image processing metrics enables to discover the plastic garbage in coastal area

- the images which has similar color variance value be reflected to the result because it is possible to detect the volume of the color variance value for each part to divide with 3 parts based on the image composition

- the part to be emphasized is emphasized properly by weighting in our system

- there is roon for improvement regarding experiment 2 and 3

As future works, we will make this system be able to acquire feature quantity from line component with hough transform technology. In addition, we will make this system be able to deal with other scales of map images for changing the pixels of radius r and to acquire automatically metadata such as location information from the image database to improve the precision.

REFERENCES

[1] The new plastic Economy Rethinking the future of plastics, http://www3.weforum.org/docs/WEF_The_New_Plastics_Economy.pdf

[2] 17 Sustainable Development Goals (SDGs), https://www.un.org/sustainabledevelopment/sustainable-development-goals/

[3] Kiyoki, Y., Kitagawa, T., and Hayama, T., "A metadatabase system for semantic image search by a mathematical model of meaning," ACM SIGMOD Record, vol.23 no.4 pp.34-41, 1994.

[4] Kiyoki, Y., Chen, X., Sasaki, S., and Koopipat, C., "Multi-Dimensional Semantic Computing with Spatial-Temporal and Semantic Axes for Multi-spectrum Images in Environment Analysis", Information Modelling and Knowledge Bases, Vol. XXVII, IOS Press, pp.14-30, 2016.

[5] Sasaki, S., Takahashi, Y., and Kiyoki, Y., "The 4D World Map System with Semantic and Spatiotemporal Analyzers", Information Modelling and Knowledge Bases, Vol.XXI, IOS Press, pp.1 - 18, 2010.

[6] Sasaki, S, and Kiyoki, Y., "Analytical Visualization Function of 5D World Map System for Multi-Dimensional Sensing Data", Information Modelling and Knowledge Bases, Vol. XXIX, IOS Press, pp. 71-89, 2018.

[7] Nguyen, D. T. N, Sasaki, S., and Kiyoki, Y., "5D World PicMap: Imagination-based Image Search System With Spatiotemporal Analyzer", Proceeding of The IASTED e-society 2011 Conference, Avila, Spain, pp. 271- 278, 2011.

[8] Hayashi, Y., Kiyoki, Y., and Chen, X., "An Image-Query Creation Method for Expressing User's Intentions by Combining Multiple Images", Information Modelling and Knowledge Bases, Vol.XXI, IOS Press, pp. 188 – 207, 2010.

[9] Hayashi, Y., Kiyoki, Y., and Chen, X., "A Combined Image-Query Creation Method for Expressing User's Intentions with Shape and Color Features in Multiple Digital Images", Information Modelling and Knowledge Bases, Vol. XXII, IOS Press, pp. 258 – 277, 2011.

[10] Google Maps, https://www.google.com/maps

[11] Godlove, I. H., "Improved Color Difference Formula, with Applications to the Perceptibility and Acceptability of Fadings",The Journal of the Optical Society of America (JOSA), Vol.41, Num.11, pp.760-772, 1951.

[12] Flickr, https://www.flickr.com/

[13] Google Image, https://images.google.com/

[14] Kobayashi, S., "Color Image Scale", The Nippon Color & Design Research Institute ed.,Kodansha International, 2001.

[15] Content Based Image Retrieval / Image Database Search Enging, http://wang.ist.psu.edu/docs/related/

INTERSECTION TYPE SYSTEM AND LAMBDA CALCULUS WITH DIRECTOR STRINGS

Xinxin Shen and Kougen Zheng

Department of Computer Science and Technology,
Zhejiang University, Hangzhou, China

ABSTRACT

The operation of substitution in λ-calculus is treated as an atomic operation. It makes that substitution operation is complex to be analyzed. To overcome this drawback, explicit substitution systems are proposed. They bridge the gap between the theory of the λ-calculus and its implementation in programming languages and proof assistants. λ_o-calculus is a name-free explicit substitution. Intersection type system for various explicit substitution calculi are studies. In this paper, we put our attention to λ_o-calculus. We present an intersection type system for λ_o-calculus and show it satisfies the subject reduction property.

KEYWORDS

Intersection type, Lambda calculus, Director strings, Subject reduction

1. INTRODUCTION

In λ-calculus [1], the operation of substitution is treated as an atomic operation. But in the presence of variable binding, substitution is a complex operation to define and implement and may cause size explosion. Therefore, substitutions are delayed and explicitly recorded, in practice. Contrast to the λ-calculus, explicit substitution decomposes the higher-order substitution operation into more atomic steps. In these last years, several explicit substitution systems have been proposed [2-5]. They are divided into two kinds: *named*, such as λx_{gc}, and *unnamed*, such as λs_e, $\lambda\sigma$. Director strings were introduced by Kennaway and Sleep [6] and generalized by Sreedhar and Taghva [7] to capture strong reduction. Fernández et al. [8-9] present the open calculus λ_o which can fully simulate the β-reduction. λ_o-calculus offers an alternative to de Bruijn notation [10] for unnamed calculi. Terms are annotated by director strings which indicate how the substitutions should do. All these explicit substitution calculi provide bridges between formal calculus and their concrete implementations. They lead to a more pertinent analysis of the correctness and efficiency of compilers, theorem proves, and proof-checkers.

Intersection type was introduced in [11-12] to overcome the limitations of Curry's type assignment system and to provide a characterization for the solvable terms of the λ-calculus. It extends simple types to include intersections and added corresponding rules to the type assignment system. It has been used to characterize strongly (weakly or head) normalizing or solvable terms in many variants of the λ-calculus [13-15] and to prove properties in λ-calculus, such as termination [16]. Moreover, approximation theorem, which is an important result in λ-calculus, also can be proved by intersection types [17].

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 147–158, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90313

Related works. Several intersection types for explicit substitution were studied. Dougherty and Lescanne [18] studied the relationship between intersection types and reduction (left reduction and head reduction) of λx.Lengrand [13] characterized strongly normalizing terms of $\lambda x_{gc}$with intersection types.Ventura et al. [19] presented an intersection type system for $\lambda_{db}$and show the subject reduction property. Ventura et al. [20] introduced intersection type systems for $\lambda s_e, \lambda\sigma, \lambda\nu$-calculus and prove the subject reduction property for them. The intersection type system in [20] cannot directly adapted to λ_o-calculus because there are no number indexes for variables. We cannot find the type of the variable from the type environment considered by searching the index. So, the difficulty for the intersection type system for $\lambda_o$is to build the correspondence from variables to the type environment.

To our knowledge the intersection type system for λ_o-calculus has not been studied. In this paper, we introduce an intersection type system for λ_o-calculus and prove the subject reduction property. The rest of this paper is structured as follows. In Section 2, we provide the term syntax of λ_o-calculus. We present the intersection type system for λ_o- calculus and show the subject reduction property in Section 3. We conclude in Section 4.

2. LAMBDA CALCULUS WITH DIRECTOR STRINGS λ_o

2.1. Term Syntax

We recall some definitions and properties of the λ_o from [8], adding some notations.

Definition 1. (**λ-calculus with Director Strings** [8])

Four syntactic categories are defined:

- Directors: We use five special symbols, called directors, ranged over by α, γ, δ:

1. '↘' indicates that the substitution should be propagated only to the right branch of a binary construct (application or substitution, as given below).
2. '↙' indicates that the substitution should be propagated only to the left branch of a binary construct.
3. '⇄' indicates that the substitution should be propagated to both branches of a binary construct.
4. '↓' indicates that the substitution should traverse a unary construct (abstraction and variables, see below).
5. '−' indicates that the substitution should be discarded (when the variable concerned does not occur in a term).

- Strings: A director string is either empty, denoted by ϵ, or built from the above symbols (so is of the form $\alpha_1\alpha_2\cdots\alpha_n$ where the α_i 's are directors). We use Greek letters such as ρ, σ ... to range over strings.

- The length of a string σ is denoted by |σ|. If α is a director, then α^n denotes a string of α 's of length n. If σ is a director string of length n and $1 \leq i \leq j \leq n$, σ_i denotes the i th director of σ and $\sigma_{\setminus i} = \sigma_1 \cdots \sigma_{i-1}\sigma_{i+1} \cdots \sigma_n$ is σ where the i th director has been removed. $\sigma_{i..j} = \sigma_i \cdots \sigma_j$ is our notation for substrings. We use σ_+ to represent the σ where all - have been removed.

$|\sigma|_l$ denotes the number of ↙ and ⇄ occurring in σ , $|\sigma|_r$ the number of ↘ and ⇄, $|\sigma|_{lr}$ the number of ⇄, and $|\sigma|_+$ the number of directors that are not $-$. $|\sigma|_-$ is the number of directors that are $-$.

- Preterms: Let σ range over strings, k be a natural number and $\boldsymbol{t}, \boldsymbol{u}$ range over preterms, which are defined by the following grammar: $t::=\square^{\sigma}|\ (\lambda\ t)^{\sigma}|(tu)^{\sigma}|\ t[k/u]\hat{}\sigma$.

- Terms: Well-formed terms are preterms that recursively satisfy the conditions in **Error! Reference source not found.**, where $\mathcal{U} = (\downarrow\ |-)^*$ and $\mathcal{B} = (\swarrow\ |\ \rightleftarrows\ |\ \searrow\ |-)\hat{}*$.

Name	Term	Constraints
Variable	$\square^{\sigma}$	$\sigma \in \mathcal{U}, \lvert\sigma\rvert_+ = 1$
Abstraction	$(\lambda t^{\rho})^{\sigma}$	$\sigma \in \mathcal{U}, \lvert\rho\rvert = \lvert\sigma\rvert_+ + 1$
Application	$(t^{\rho}u^{\nu})^{\sigma}$	$\sigma \in \mathcal{B}, \lvert\rho\rvert = \lvert\sigma\rvert_l, \lvert\nu\rvert = \lvert\sigma\rvert_r$
Substitution	$(t^{\rho}[k/u^{\nu}])^{\sigma}$	$\sigma \in \mathcal{B}, \lvert\rho\rvert = \lvert\sigma\rvert_l + 1, \lvert\nu\rvert = \lvert\sigma\rvert_r, 1 \le k \le \lvert\rho\rvert$

Figure 1.Term Condition

A variety of different term constructs [8]:

- □presents variables,
- $(\lambda t^{\rho})^{\sigma}$ is an abstraction,
- $(\boldsymbol{tu})^{\sigma}$ is an application,
- $\boldsymbol{t}[k/\boldsymbol{u}]^{\sigma}$ is an explicit substitution, meaning that the variable corresponding to the k^{th} director in $\boldsymbol{t}$'s string is to be replaced by $\boldsymbol{u}$.

Remark 1.

- Given a term t^{ρ}, $|\rho|_+$ is equal to the number of free variables in t.
- In an abstraction $(\lambda t^{\rho})^{\sigma}$, the last director in ρ corresponds to the bound variable.
- Parentheses will be dropped whenever we can, and omit the empty string ϵ unless it is essential.

We give an example:

Example 1. If we consider variables $\{y, z, w\}$, and the pure λ -term $t = \lambda x.xyw$. The term in λ_o-calculus corresponding to (use the function defined in [8]) t is $((\lambda\ (\square^{\downarrow}\square^{\downarrow})^{\searrow\swarrow}\square^{\downarrow})^{\swarrow\ \searrow\swarrow})^{\downarrow-\downarrow}$.

2.2. Reduction rules

The *Beta* rule is aimed at eliminating β-redexes and introduce an explicit substitution. It is defined by

$$\lambda t^{\rho}\boldsymbol{u} \rightsquigarrow_o (\boldsymbol{t}[|\ \rho\ |_+ + 1/\boldsymbol{u}])^{\tau}$$

where $\tau = \psi_b(\sigma, \rho)$with φ_b defined in Fig. 2.

Definition 2. (Reduction Rules [8])

Reduction rules in λ_o contain the *Beta* rule and propagation rules in Figure 2 with:

$$\begin{array}{llll}
\pi(\epsilon, \epsilon) = \epsilon & & \pi'(\epsilon,\epsilon) = \epsilon \\
\pi(\searrow \sigma, \rho) = -\pi(\sigma,\rho) & & \pi'(\searrow \sigma, \alpha\nu) = -\pi \\
\pi(\swarrow \sigma, \alpha\rho) = \alpha\,\pi(\sigma,\rho) & & \pi'(\swarrow \sigma, \nu) = -\pi \\
\pi(\rightleftarrows \sigma, \alpha\rho) = \alpha\ \pi(\sigma,\rho) & & \pi'(\rightleftarrows \sigma, \alpha\nu) = \alpha\pi' \\
\pi(-\sigma, \rho) = -\pi(\sigma,\rho) & & \pi'(-\sigma, \nu) = -\pi
\end{array}$$

Other functions used in the propagation rules are defined in Figure 3. The functions used in the propagation rules just compute the *ad hoc* director strings. They are generated recursively in the same way as above from the tables Figure 3 [8].

Name	*Reduction*			*Cond.*
Var	$(\Box^{\rho}[i/v^{\nu}])^{\sigma}$	$\rightsquigarrow_o$	v^{τ} where $\tau = \pi'(\sigma, \nu)$	$\rho_i = \downarrow$
App_1	$((\mathbf{t}\ \mathbf{u})^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	$((\mathbf{t}[j/\mathbf{v}])^{\upsilon}\ \mathbf{u})^{\tau}$	$\rho_i = \swarrow$
	where $\upsilon = \phi_l(\sigma, \rho_{\setminus i}), \tau = \psi_1(\sigma, \rho_{\setminus i}), j = \lvert\rho_{1..i}\rvert_l$			
App_2	$((\mathbf{t}\ \mathbf{u})^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	$(\mathbf{t}\ (\mathbf{u}[k/\mathbf{v}])^{\omega})^{\tau}$	$\rho_i = \searrow$
	where $\omega = \phi_r(\sigma, \rho_{\setminus i}), \tau = \psi_2(\sigma, \rho_{\setminus i}), k = \lvert\rho_{1..i}\rvert_r$			
App_3	$((\mathbf{t}\ \mathbf{u})^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	$((\mathbf{t}[j/\mathbf{v}])^{\upsilon}\ (\mathbf{u}[k/\mathbf{v}])^{\omega})^{\tau}$	$\rho_i = \rightleftarrows$
where $\upsilon = \phi_l(\sigma, \rho_{\setminus i}), \omega = \phi_r(\sigma, \rho_{\setminus i}), \tau = \psi_3(\sigma, \rho_{\setminus i}), j = \lvert\rho_{1..i}\rvert_l, k = \lvert\rho_{1..i}\rvert_r$				
Lam	$((\lambda\mathbf{t})^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	$\left(\lambda(\mathbf{t}[i/\mathbf{v}])^{\upsilon\cdot\swarrow}\right)^{\tau}$	$\rho_i = \downarrow$
	where $\upsilon = \phi_d(\sigma, \rho_{\setminus i}), \tau = \psi_d(\sigma, \rho_{\setminus i})$			
Comp	$((\mathbf{t}[j/\mathbf{u}])^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	$(\mathbf{t}[j/(\mathbf{u}[k/\mathbf{v}])^{\omega}])^{\tau}$	$\rho_i = \searrow$
	where $\omega = \phi_r(\sigma, \rho_{\setminus i}), \tau = \psi_2(\sigma, \rho_{\setminus i}), k = \lvert\rho_{1..i}\rvert_r$			
Erase	$(t^{\rho}[i/\mathbf{v}])^{\sigma}$	$\rightsquigarrow_o$	t^{τ} where $\tau = \pi(\sigma, \rho_{\setminus i})$	$\rho_i = -$

Figure 2. Propagation Rules $\mathcal{P}$

σ_1	ρ_1	ϕ_l	ϕ_r	ψ_1	ψ_2	ψ_3	π
$\searrow$	ϵ	$\searrow$	$\searrow$	$\swarrow$	$\searrow$	$\rightleftarrows$	$-$
$\swarrow$	$\searrow$	ϵ	$\swarrow$	$\searrow$	$\searrow$	$\searrow$	$\searrow$
$\swarrow$	$\swarrow$	$\swarrow$	ϵ	$\swarrow$	$\swarrow$	$\swarrow$	$\swarrow$
$\swarrow$	$\rightleftarrows$	$\swarrow$	$\swarrow$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$
$\rightleftarrows$	$\searrow$	$\searrow$	$\rightleftarrows$	$\rightleftarrows$	$\searrow$	$\rightleftarrows$	$\searrow$
$\rightleftarrows$	$\swarrow$	$\rightleftarrows$	$\searrow$	$\swarrow$	$\rightleftarrows$	$\rightleftarrows$	$\swarrow$
$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$	$\rightleftarrows$
$-$	ϵ	ϵ	ϵ	$-$	$-$	$-$	$-$
$\swarrow$	$-$	ϵ	ϵ	$-$	$-$	$-$	$-$
$\rightleftarrows$	$-$	$\searrow$	$\searrow$	$\swarrow$	$\searrow$	$\rightleftarrows$	$-$

σ_1	ρ_1	ψ_d	ϕ_d	ψ_b
$\searrow$	ϵ	$\downarrow$	$\searrow$	$\searrow$
$\swarrow$	$\downarrow$	$\downarrow$	$\swarrow$	$\swarrow$
$\rightleftarrows$	$\downarrow$	$\downarrow$	$\rightleftarrows$	$\rightleftarrows$
$-$	ϵ	$-$	ϵ	$-$
$\swarrow$	$-$	$-$	ϵ	$-$
$\rightleftarrows$	$-$	$\downarrow$	$\searrow$	$\searrow$

Figure 3. Functions used in the Propagation Rules

3. THE TYPE SYSTEM

3.1 Intersection Types

We take the idea from [20]. Environments are sequences of types instead of type assignments and types are non-idempotent intersection types [21].

Definition 3.

1. Intersection types are defined by:

$$\varrho, \varsigma \in \mathbb{T} ::= \mathcal{A} \mid \mathbb{U} \to \mathbb{T}$$
$$\mu, \xi \in \mathbb{U} ::= \Omega \ \mathbb{U} \cap \mathbb{U} \mid \mathbb{T}$$

where $\mathcal{A}$ is a denumerable infinite set of type variables.

2. Environments are ordered lists of types $\mu \in \mathbb{U}$, defined by $\Gamma ::= nil \mid \mu.\Gamma$. *nil* is the empty environment. We use Γ, Δ..to denote environments.

3. A term **t** is typable if there are some Γ, ϱ such that $\Gamma \vdash \boldsymbol{t}: \varrho$.

$|\Gamma|$ is the length of Γ and $|nil| = 0$. Γ_i is the i th type in Γ. $\Gamma_{<i}$ is the first $i-1$ types in Γ and $\Gamma_{\leq i}$, $\Gamma_{>i}$, $\Gamma_{\geq i}$ are similarly defined. If $i = 0$, then $\Gamma_{\leq 0}.\Gamma = \Gamma_{<0}.\Gamma = \Gamma$. If i is equal to the length of Γ, then $\Gamma.\Gamma_{\geq i} = \Gamma.\Gamma_{>0} = \Gamma$.

$\Gamma_{\setminus i} = \Gamma_{<i}.\ \Gamma_{>i}$, is Γ where the i th type has been removed. $\Gamma_{i..j} = \Gamma_i.\cdots.\Gamma_j$, is a sub-environment of Γ. Γ_+ is Γ where all $-$ has been removed. Ω^n denotes the environment $\Omega.\cdots.\Omega$ of length n.

3.2 Type System for λ_o

We first define some functions which will be used in defining the typing rules. It gets a new environment from two environments according to a director string.

Definition 4. The function Intersection(Γ, Δ, σ) is defined by Algorithm 1.

Algorithm 1 Intersection(Γ, Δ, σ)

```
Let n = |σ|; Γ' = Ω^n;
for i = 1; i ≤ n; i++ do
    if σ_i = ⇌ then
        Γ'_i = Γ_{|σ_{1..i}|_l} ∩ Δ_{|σ_{1..i}|_r};
    else if σ_i = ↙ then
        Γ'_i = Γ_{|σ_{1..i}|_l};
    else if σ_i = ↘ then
        Γ'_i = Δ_{|σ_{1..i}|_r};
    end if
end for
return Γ'.
```

Definition 5. The function *AddErase* (Γ, σ) is defined by Algorithm 2.

Algorithm 2 AddErase (Γ, σ)

Let $n = |\sigma|; \Gamma' = \Omega^n$;
for $i = 1; i \leq n; i++$ **do**
  if $\sigma_i \neq -$ **then** $\Gamma'_i = \Gamma_{|\sigma_{1..i}|_+}$;
  end if
end for
return Γ'.

Definition 6. The function *Drop*(Γ, σ) is defined by Algorithm 3.

Algorithm 3 Drop (Γ, σ)

Let $n = |\sigma|; \Gamma' = \Omega^n$;
for $i = 1; i \leq n; i++$ **do**
  if $\sigma_i = \swarrow$ or $\sigma_i = \rightleftarrows$ **then** $\Gamma'_i = \Gamma_{|\sigma_{1..i}|_l}$;
  end if
end for
return Γ'.

Definition 7. (Typing Rules)

Typing rules for λ_o are defined by Figure where (*) is $\rho = \sigma_+$ and functions are defined in Definition 4, Definition 7 and Definition 6.

$$\text{(var)}\quad \frac{}{\varrho.nil \vdash \square^{\downarrow} : \varrho} \qquad \text{(abs)}\quad \frac{\Gamma.\mu \vdash t^{\rho} : \varrho}{AddErase(\Gamma, \sigma) \vdash (\lambda t^{\rho})^{\sigma} : \mu \to \varrho}$$

$$\text{(-i)}\quad \frac{\Gamma \vdash t^{\rho} : \varrho}{\text{AddErase}(\Gamma, \sigma) \vdash t^{\sigma} : \varrho}(*)$$

$$\text{(app)}\quad \frac{\Gamma \vdash \boldsymbol{t} : \wedge\varsigma_k \to \varrho \quad \Delta^k \vdash \boldsymbol{u} : \varsigma_k \quad \forall k \in \{1, ..., n\}}{\text{Intersection}(\Gamma, \cap\Delta^k, \sigma) \vdash (\boldsymbol{tu})^{\sigma} : \varrho}$$

$$\text{(cut)}\quad \frac{\Gamma_{<i}. \wedge \varsigma_k.\Gamma_{>i} \vdash \boldsymbol{t} : \varrho \quad \Delta^k \vdash \boldsymbol{u} : \varsigma_k \quad \forall k \in \{1, ..., n\}}{\text{Intersection}(\Gamma_{\setminus i}, \cap\Delta^k, \sigma) \vdash (\boldsymbol{t}[i/\boldsymbol{u}])^{\sigma} : \varrho} \quad \rho_i \neq -$$

$$\text{(drop)}\quad \frac{\Gamma \vdash \boldsymbol{t} : \varrho \quad \Delta \vdash \boldsymbol{u} : \varsigma}{\text{Drop}(\Gamma_{\setminus i}, \sigma) \vdash (\boldsymbol{t}[i/\boldsymbol{u}])^{\sigma} : \varrho} \quad \rho_i = -$$

Figure 4. Typing Rules

Lemma 1. (Generation Lemma)

1. $\Gamma \vdash t\hat{}\sigma : \varrho$ and $\rho = \sigma_+$, then $\Gamma = \text{AddErase}(\Gamma', \sigma)$ and $\Gamma' \vdash t^{\rho} : \varrho$.
2. $\Gamma \vdash \square\hat{}\sigma : \varrho$, if $\sigma_i \neq -$ then $\Gamma_i = \varrho$.
3. $\Gamma \vdash (\lambda \boldsymbol{t})^{\sigma} : \varrho$, then $\varrho = \mu \to \varsigma$ for some $\mu \in \mathbb{U}$ and $\varsigma \in \mathbb{T}$, where $\Gamma_+.\mu \vdash t^{\rho} : \varrho$.

4. $\Gamma \vdash (\boldsymbol{tu})^{\sigma} : \varrho$, then $\Gamma = \text{Intersection}(\Gamma', \Delta', \sigma)$ such that $\Gamma' \vdash \boldsymbol{t}: \wedge \varsigma_k \to \varrho$, $\Delta' = \cap_1^n \Delta^k$ and $\forall\, 1 \leq k \leq n, \Delta^k \vdash \boldsymbol{u}: \varsigma_k$.
5. $\Gamma \vdash (t^{\rho}[i/\boldsymbol{u}])^{\sigma}: \varrho$ and $\rho_i \neq -$, then $\Gamma = \text{Intersection}(\Gamma'_{\setminus i}, \Delta', \sigma)$ such that $\Gamma'_{<i}.\wedge \varsigma_k.\Gamma_{>i} \vdash \boldsymbol{t}: \varrho, \Delta' = \cap_1^n \Delta^k$ and $\forall\, 1 \leq k \leq n, \Delta^k \vdash \boldsymbol{u}\}: \varsigma_k$.
6. $\Gamma \vdash (t^{\rho}[i/\boldsymbol{u}]))^{\sigma}: \varrho$ and $\rho_i = -$, then $\left(\Gamma_{\leq|\rho|}\right)_{\setminus i} \vdash t^{\rho}: \varrho$ and $\Delta \vdash \boldsymbol{u}: \varsigma$.

Proof. By induction on the structure of derivations.

Lemma 2. $\boldsymbol{\Gamma \vdash t^{\rho}: \varrho}$ **,** then $|\Gamma| = |\rho|$ and the Γ_i is the type of the variable which ρ_i corresponds to.

Proof. Induction on the structure of t^{ρ}.

1. $\square^{\rho}$, it is immediately.
2. $(\lambda t^{\rho})^{\sigma}$. By **Lemma 1** $\Gamma_+.\mu \vdash t^{\rho}: \varrho$ for some $\mu \in \mathbb{U}$ and $\varsigma \in \mathbb{T}$, where $\varrho = \mu \to \varsigma$. By induction hypothesis, $|\Gamma_+.\mu| = |\rho|$ and Γ_{+i} is the type of ρ_i. By term conditions, $|\rho| = |\sigma|_+ + 1$. From the definition of *AddErase*, we can get $|\Gamma| = |\sigma|$ and Γ_i is the type ofσ_i.
3. $(t^{\rho}u^{\nu})^{\sigma}$. By **Lemma 1**, $\Gamma = \text{Intersection}(\Gamma', \Delta', \sigma)$ such that $\Gamma' \vdash t^{\rho}: \wedge \varsigma_k \to \varrho$, $\Delta' = \cap_1^n \Delta^k$ and $\forall\, 1 \leq k \leq n, \Delta^k \vdash u^{\nu} : \varsigma_k$. By induction hypothesis, $|\Gamma'| = |\rho| = |\sigma|_l$, Γ'_i is the type of the variable which ρ_i corresponds to and $|\Delta'| = |\Delta^k| = |\nu| = |\sigma|_r$, Δ'_i is the type of the variable which ν_i corresponds to. From the definition of the function *Intersection*, we easily get $|\text{Intersection}(\Gamma', \Delta', \sigma)| = |\sigma|$ and Γ_i is the type of the variable whichσ_i corresponds to.
4. $\Gamma \vdash (t^{\rho}[i/\mathbf{u}])^{\sigma}: \varrho$ and $\rho_i \neq -$, it is similar to the last case.
5. $\Gamma \vdash (t^{\rho}[i/\mathbf{u}])^{\sigma}: \varrho$ and $\rho_i = -$, it is immediately by induction hypothesis and the reduction rule.

Theorem 1. (Subject Reduction)

Let $\Gamma \vdash \boldsymbol{t}:\varrho$. If $\boldsymbol{t} \rightsquigarrow_o \boldsymbol{u}$, then $\Gamma \vdash \boldsymbol{u} : \varrho$.

Proof. By the verification of subject reduction for each reduction rule of the λ_o-calculus.

• *(Beta)*: Let $\Gamma \vdash ((\lambda\, \mathbf{t})^{\rho}\mathbf{u})\hat{}\sigma : \varrho$. We want to prove that $\Gamma \vdash (\mathbf{t}[|\rho|_+ + 1 / \mathbf{u}])^{\tau}: \varrho$. By **Lemma 1**, we have the following derivation:

$$\frac{\dfrac{\Gamma_2.\wedge \varsigma_k \vdash \boldsymbol{t} : \varrho}{\Gamma' = AddErase(\Gamma_2, \rho) \vdash (\lambda \boldsymbol{t})^{\rho} : \wedge \varsigma_k \to \varrho} \quad \forall 1 \leq k \leq n, \Delta^k \vdash \boldsymbol{u} : \varsigma_k}{\Gamma = \text{Intersection}(\Gamma', \cap \Delta^k, \sigma) \vdash ((\lambda \boldsymbol{t})^{\rho}\boldsymbol{u})^{\sigma} : \varrho}$$

By Lemma 2, $|\Gamma'| = |\rho|$ and Γ'_i is the type of the variable which ρ_i corresponds to. $\Gamma_1 = \text{Intersection}\}(\Gamma_2, \Delta', \tau) \vdash (\mathbf{t}[|\rho|_+ + 1 / \mathbf{u}])^{\tau}: \varrho$ by rule (cut). Suppose the variable σ_i corresponding to is x, observing the procedure and the definition of φ_b:

1. $\sigma_i = \searrow$. Then $\tau_i = \searrow$; $\Gamma_{1_i} = \Delta'_{|\tau_{1..i}|_r}$. $\Gamma_i = \Delta'_{|\sigma_{1..i}|_r}$. They are the type of $\boldsymbol{u}_x$ (variable x in $\boldsymbol{u}$). $\Gamma_{1_i} = \Gamma_i$.
2. $\sigma_i = \swarrow$. σ_i indicates the substitution is propagated to the left branch of $((\lambda \boldsymbol{t})^{\rho} \boldsymbol{u})$. If $\rho_i \neq -$, $\tau_i = \swarrow = \sigma_i$. Γ_i is the type of $\boldsymbol{t}_x$. Let $m = |\sigma_{1..i}|_l$ and $n = |\rho_{1..m}|_+$. $\Gamma_i = \Gamma_{2_n}$. τ_i indicate the substitution is propagated to the left branch $\boldsymbol{t}$. It is the type of $\boldsymbol{t}_x$. $\Gamma_{1_i} = \Gamma_{2_{|\tau_{1..i}|_l}}$. If $\rho_i = -$, then $\tau_i = -$. The variable does not occur in the two terms. Then $\Gamma_i = \Omega = \Gamma_{1_i}$.
3. $\sigma_i = \rightleftarrows$ indicates the substitution should be propagated into both branches. If $\rho_i \neq -$, $\tau_i = \rightleftarrows$.Let $m = |\sigma_{1..i}|_l$ and $n = |\sigma_{1..i}|_r$. $p = |\rho_{1..m}|_+$. $\Gamma_i = \Gamma_{2_p} \cap (\cap \Delta^k)_n$. It is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{u}_x$.$\tau_i = \rightleftarrows$ indicate the substitution should be propagated into both branches. Let $m' = |\tau_{1..i}|_l$ and $n' = |\tau_{1..i}|_r$. $\Gamma_{1_i} = \Gamma_{2_{m'}} \cap (\cap \Delta^k)_{n'}$. It is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{u}_x$. If $\rho_i = -$, $\tau_i = \searrow$. Then $\Gamma_{1_i} = \Delta'_{|\tau_{1..i}|_r}$, it is the type of $\boldsymbol{u}_x$. $\Gamma_i = \Delta'_{|\sigma_{1..i}|_r}$, it is also the type of $\boldsymbol{u}_x$.
4. $\sigma_i = -$, then $\tau_i = -$. So $\Gamma_{1_i} = \Gamma_i$.

Therefore, $\Gamma = \Gamma_1$.

- (Var): Let $\Gamma \vdash (\square^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho$ and $\rho_i = \downarrow$. We want to prove that $\Gamma \vdash \boldsymbol{v} : \varrho$.By **Lemma 1**, $\Gamma = \text{Intersection}(\Gamma'_{\setminus i}, \Delta', \sigma)$, $\Gamma'_{<i} . \wedge \varsigma_k . \Gamma'_{>i} \vdash \square^{\rho} : \varrho$ and $\forall\, 1 \leq k \leq n \Delta^k \vdash \boldsymbol{v} : \varsigma_k$.By term definition, $\rho_j = -$ for all$j \neq i$. So $\Gamma'_{\setminus i} = \Omega^{|\rho|-1}$.Observing the procedure of function Intersection.

  - $\sigma_i = \rightleftarrows$ or $\sigma_i = \searrow$. $\Gamma_i = \Omega \cap \Delta'_{|\sigma_{1..i}|_r} = \Delta'_{|\sigma_{1..i}|_r}$. i.e. Γ_i is type of the $\{|\sigma_{1..i}|_r$th variable.
  - $\sigma_i = -$ or $\sigma_i = \swarrow$. $\Gamma_i = \Omega$.

By **Lemma 2**, it coincides with the definition of π'.

- (*Erase*): Let $\Gamma \vdash (t^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho$ and $\rho_i = -$. We want to prove that $\Gamma \vdash t^{\tau} : \varrho$.By **Lemma 1**, and **Lemma 2**, it is easily to get that Drop(Γ, σ) is coincides with π.

- (*Lam*): Let $\Gamma \vdash ((\lambda \boldsymbol{t})^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho$ and $\rho_i = \downarrow$. We want to prove that $\Gamma \vdash (\lambda(\boldsymbol{t}[i/\boldsymbol{v}])^{\nu.\swarrow})^{\tau} : \varrho$.By **Lemma 1**, we have the following derivation:

$$\frac{\dfrac{\Gamma'.\mu \vdash \boldsymbol{t} : \zeta}{\Gamma' \vdash (\lambda \boldsymbol{t})^{\rho} : \varrho} \quad \forall 1 \leq k \leq n\ \Delta^k \vdash \boldsymbol{v} : \varsigma_k \quad \Gamma'_i = \wedge \varsigma_k}{\Gamma = \text{Intersection}(\Gamma'_{\setminus i}, \cap \Delta^k, \sigma) : ((\lambda \boldsymbol{t})^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho}$$

where $\mu \to \zeta = \varrho$. Hence

$$\frac{\dfrac{\Gamma'.\mu \vdash \boldsymbol{t} : \zeta \quad \forall 1 \leq k \leq n\ \Delta^k \vdash \boldsymbol{v} : \varsigma_k \Gamma'_i = \cap_1^n \varsigma_k}{\Gamma_1 = \text{Intersection}(\Gamma'_{\setminus i}, \Delta', \nu.\swarrow).\mu \vdash (\boldsymbol{t}[i/\boldsymbol{v}])^{\nu.\swarrow} : \mu \to \zeta}}{\Gamma_2 = \text{AddErase}(\Gamma_1, \nu.\swarrow)) \vdash (\lambda(\boldsymbol{t}[i/\boldsymbol{v}])^{\nu.\swarrow})^{\tau} : \varrho}$$

Observing the definition of ϕ_d, there are no $-$'s in ν. Suppose the j th variable is x.

1. $\sigma_j = \swarrow, \rho_j = -,\ \nu_j = \epsilon, \tau_j = -.\ \Gamma_{2_j} = \Omega = \Gamma_j.$

2. $\sigma_j = \swarrow, \rho_j = \downarrow,\ \nu_j = \swarrow,\ \tau_j = \downarrow. \sigma_j = \swarrow$ indicates the substitution is propagated to the left branch $(\lambda\, \boldsymbol{t})^{\rho}$. Let $m = \left|\sigma_{1..j}\right|_l$. Γ_j is type of the variable ρ_m corresponding to, indicated by $\boldsymbol{t}_x$ (variable x in $\boldsymbol{t}$). $\Gamma_j = \Gamma'_{\backslash i_m}. \nu_j = \swarrow$ indicates the substitution is propagated to the left branch $\boldsymbol{t}$. Γ_{2_j} is the the type of $\boldsymbol{t}_x$. Let $m' = \left|\tau_{1..j}\right|_+$ and $n' = |\nu.\swarrow|_l$. $\Gamma_{2_j} = \Gamma_{1_{m'}} = \Gamma'_{\backslash i_{n'}}$.

3. $\sigma_j = \searrow, \rho_j = \epsilon, \nu_j = \searrow, \tau_j = \downarrow$. Let m be $\left|\tau_{1..j}\right|_+$. The type of $\Gamma_{2_j} = \Gamma_{1_m} = \Delta'_{|\nu_{1..m}|_r}$. $\Gamma_j = \Delta'_{\left|\sigma_{1..j}\right|_r}. \sigma_j = \searrow$ indicates the substitution is propagated to the right branch $\boldsymbol{v}$. Let $m = \left|\sigma_{1..j}\right|_r$. Γ_j is the type of $\boldsymbol{v}_x$. $\Gamma_j = \left(\cap \Delta^k\right)_m$. $\nu_j = \searrow$ indicates the substitution is propagated to the right branch $\boldsymbol{v}$. Γ_{2_j} is the type of $\boldsymbol{v}_x$. Let $m' = \left|\tau_{1..j}\right|_+$ and $n' = \left|\nu.\swarrow_{1..j}\right|_l$. $\Gamma_{2_j} = \Gamma_{1_{m'}} = \Gamma'_{\backslash i_{n'}}$.

4. $\sigma_j = \rightleftarrows, \rho_j = \downarrow,\ \nu_j = \rightleftarrows, \tau_j = \downarrow. \sigma_j = \rightleftarrows$ indicates the substitution is propagated to both branches $\boldsymbol{t}$ and $\boldsymbol{v}$. Let $m = \left|\sigma_{1..j}\right|_l, n = \left|\sigma_{1..j}\right|_r$. Γ_j is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{u}_x$. $\Gamma_j = \Gamma'_{\backslash i_m} \cap \left(\cap \Delta^k\right)_n$. $\nu_j = \rightleftarrows, \tau_j = \downarrow$ indicates the substitution is propagated to both branches $\boldsymbol{t}$ and $\searrow$. $\Gamma_{2_j}$is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{v}_x$. Let $m' = \left|\tau_{1..j}\right|_+, n' = |\nu.\swarrow_{1..m}|_l, p' = |\nu.\swarrow_{1..m}|_r$. $\Gamma_{2_j} = \Gamma_{1_m} = \Gamma'_{\backslash i_{n'}} \cap \left(\cap \Delta^k\right)_{p'}$.

5. $\sigma_j = \rightleftarrows, \rho_j = -,\ \nu_j = \searrow, \tau_j = \downarrow.\ \sigma_j = \rightleftarrows$ indicates the substitution is propagated to both branches $\boldsymbol{t}$ and $\boldsymbol{v}$. Γ_j is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{v}_x$. $\rho_j = -$, the type of $\boldsymbol{t}_x$ is Ω. So Γ_j is the type of $\boldsymbol{v}_x$. Let $n = \left|\sigma_{1..j}\right|_r$, $\Gamma_j = \left(\cap \Delta^k\right)_n$.Let $m' = \left|\tau_{1..j}\right|_+, n' = |\nu_{1..m}|_r$. The substitution is propagated to right branch $\boldsymbol{v}$. $\Gamma_{2_j}$is the type of $\boldsymbol{v}_x$. $\Gamma_{2_j} = \Gamma_{1_m} = \left(\cap \Delta^k\right)_{n'}$.

6. $\sigma_j = -, \rho_j = \epsilon,\ \nu_j = \epsilon, \tau_j = -.\ \Gamma_{2_j} = \Omega = \Gamma_j$.
   So $\Gamma_2 = \Gamma$.

- App_1, App_2, App_3 and Comp are similar. We just show the case App_1. Let $\Gamma \vdash ((\boldsymbol{tu})^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho$ and $\rho_i = \swarrow$. We want to prove that $\Sigma \vdash ((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})^{\tau} : \varrho$. By Lemma 1, we have the following derivation.

$$\frac{\dfrac{\Gamma_1 \vdash \boldsymbol{t} : \wedge\vartheta_k \to \varrho \quad \forall 1 \le k \le n\ \Gamma^k \vdash \boldsymbol{u} : \vartheta_k}{\Gamma' = \text{Intersection}(\Gamma_1, \cap\Gamma^k, \rho) : (\boldsymbol{tu})^{\rho} : \varrho} \quad \forall 1 \le k \le n\ \Delta^k \vdash \boldsymbol{v} : \varsigma_k \quad \Gamma_i = \wedge\varsigma_k}{\Gamma = \text{Intersection}(\Gamma'_{\backslash i}, \cap\Delta^k, \sigma) : ((\boldsymbol{tu})^{\rho}[i/\boldsymbol{v}])^{\sigma} : \varrho}$$

Hence

$$\frac{\dfrac{\Gamma_1 \vdash \boldsymbol{t} : \wedge\vartheta_k \to \varrho \quad \forall 1 \le k \le n\ \Delta^k \vdash \boldsymbol{v} : \varsigma_k \quad \Gamma_{1j} = \wedge\varsigma_k}{\Gamma_2 = \text{Intersection}(\Gamma_{1\setminus j}, \cap\Delta^k, \nu) \vdash (\boldsymbol{t}[j/\boldsymbol{v}])^{\nu} : \wedge\vartheta_k \to \varrho} \quad \forall 1 \le k \le n\ \Gamma^k \vdash \boldsymbol{u} : \vartheta_k}{\Gamma' = \text{Intersection}(\Gamma_2, \cap\Gamma^k, \tau) \vdash ((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})^{\tau} : \varrho}$$

For better reading, we use variable names. Suppose the mth variable is x.

1) $\sigma_m = -, \rho_m = \epsilon,\ \nu_m = \epsilon, \tau_m = -$. The m th variable does not occurs in $((\boldsymbol{tu})^{\rho}[i/\boldsymbol{v}])^{\sigma}$ and $((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})^{\tau}$. $\Gamma'_m = \Gamma_m = \Omega$.

2) $\sigma_m = \swarrow$. The m th substitution is only considered on the left branch of $((\boldsymbol{tu})^{\rho}[i/\boldsymbol{v}])^{\sigma}$. It is the $n = |\sigma_{1..m}|_l$th in ρ.

   a. $\rho_m = \searrow,\ \nu_m = \epsilon, \tau_m = \searrow$. The n th substitution is only considered on the right brach of $\boldsymbol{tu}$. It is the $p = |\rho_{1..n}|_r$ th in $\boldsymbol{u}$'s director string. The variable ρ_p corresponding to is denoted by $\boldsymbol{u}_x$ (variable x in $\boldsymbol{u}$). So $\Gamma_m = \left(\cap\,\Delta^k\right)_p$. It is the type of $\boldsymbol{u}_x$. The m th substitution is only considered on the right branch of $((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})$. It is the $n' = |\tau_{1..m}|_r$th in $\boldsymbol{u}$'s director string. So $\Gamma'_m = \left(\cap\,\Delta^k\right)_{n'}$. It is the type of $\boldsymbol{u}_x$.

   b. $\rho_m = \swarrow,\ \nu_m = \swarrow, \tau_m = \swarrow$. The nth substitution is only considered on the left brach of $\boldsymbol{tu}$. It is the $p = |\rho_{1..n}|_l$th in $\boldsymbol{t}$'s director string. The variable is denoted by $\boldsymbol{t}_x$. $\Gamma_m = \Gamma_{1_p}$ and it is the type of $\boldsymbol{t}_x$. The m th substitution is only considered on the left branch of $((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})$. It is the $n' = |\tau_{1..m}|_l$th in $(\boldsymbol{t}[j/\boldsymbol{v}])$. Then it is the $p' = |\nu_{1..n'}|_l$ in $\boldsymbol{t}$'s string. So $\Gamma'_m = \Gamma_{1_{p'}}$ and it is the type of $\boldsymbol{t}_x$.

   c. $\rho_m = \rightleftarrows,\ \nu_m = \swarrow, \tau_m = \rightleftarrows$. The n th substitution is considered both on the left and right braches of $\boldsymbol{tu}$. It is the $p = |\rho_{1..n}|_l$th in $\boldsymbol{t}$'s string and $q = |\rho_{1..n}|_r$th in $\boldsymbol{u}$'s string. $\Gamma_m = \Gamma_{1_p} \cap \left(\cap\,\Gamma^k\right)_q$. It is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{u}_x$. The m th substitution is considered both on the left and right branches of $((\boldsymbol{t}[j/\boldsymbol{v}])^{\nu}\boldsymbol{u})$. It is the $n' = |\tau_{1..m}|_l$th in ν. Then it is the $p' = |\nu_{1..n'}|_l$ in $\boldsymbol{t}$'s string. $q' = |\tau_{1..m}|_r$ in $\boldsymbol{u}$'s string. So $\Gamma'_m = \Gamma_{1_{p'}} \cap \left(\cap\,\Gamma^k\right)_{q'}$. It is intersection of the type of $\boldsymbol{t}_x$ and $\boldsymbol{u}_x$.

   d. $\rho_m = -,\ \nu_m = \epsilon, \tau_m = -$. $\Gamma_m = \Gamma'_m = \Omega$.
   So $\Gamma = \Gamma'$.

3) $\sigma_m = \rightleftarrows$ or $\sigma_m = \searrow$. These two cases are similar to $\sigma = \swarrow$. We can get $\Gamma = \Gamma'$ from the sketch of the last case.

3. CONCLUSIONS

λ_o-calculus is an unnamed explicit substitution calculus. Director strings are added to indicate how substitutions should do. It offers an alternative to de Bruijn notation. It can be used in theorem prover implementation. λ_o-calculus fully simulates the β-reduction in classical λ-calculus and it preserves the PSN property. In this paper, we propose an intersection type system for λ_o and prove the type system satisfy the subject reduction property. If a term *M* can reduce to *N,* if *M* is typed by ϱ, then *N* is also typed byϱ. In the future work, we will try to prove that a

typable term in this type system is strongly normalizing and try to show a term is strongly normalizing if and only if it is typable in a certain intersection type.

REFERENCES

[1] Barendregt, H., (1984) The Lambda Calculus: Its Syntax and Semantics, Elsevier

[2] Lescanne, P. (1994) "From λσ to λv a journey through calculi of explicit substitutions", In: Proceedings of POPL, pp 60–69.

[3] Bloo, R. & Rose, K., (1995) "Preservation of strong normalisation in named lambda calculi with explicit substitution and garbage collection", In: Computing Science in the Netherlands, pp 62–72.

[4] Bloo, R. & Geuvers, J., (1999) "Explicit substitution: on the edge of strong normalization", Theoretical Computer Science, Vol. 211, pp 375–395.

[5] Ayala-Rincón,M. & Kamareddine, F., (2001) "Unification via the λse-style of explicit substitution", Logic Journal of the IGPL, Vol 9, pp 489–523.

[6] Kennaway, R. & Sleep, R., (1988) "Director strings as combinators", ACM Transactions on Programming Languages and Systems, Vol. 10, No. 4, pp 602–626.

[7] Sreedhar.V. &Taghva, K.,(1993)"Capturing strong reduction in director string calculus",Theoretical Computer Science, Vol. 107, No. 2, pp 333–347.

[8] Fernández,M., Mackie,I. &Sinot,F.R.,(2005)"Lambda-calculus with director strings", Applicable Algebra in Engineering, Communication and Computing, Vol. 15, No.6, pp 393–437.

[9] Sinot, F.R.,Fernández,M. & Mackie,I.,(2003)"Efficient Reductions with Director Strings", In: Nieuwenhuis R. (eds) Rewriting Techniques and Applications. RTA 2003. Lecture Notes in Computer Science, Vol 2706. Springer, pp 46-60.

[10] De Bruijn,N., (1972)"Lambda calculus notation with nameless dummies",IndagationesMathematicaeVol 34, pp 381–392.

[11] Coppo,M. &Dezani-Ciancaglini,M.,(1978)"A new type assignment for lambda-terms",Archivfu¨r mathematischeLogik und Grundlagenforschung, Vol 19, pp 139–156.

[12] Coppo, M. &Dezani-Ciancaglini,M.,(1980)"An extension of the basic functionality theory for the λ-calculus", Notre Dame J. Formal Logic, Vol. 21, No.4, pp 685–693.

[13] Lengrand, S.,Lescanne, P.,Dougherty,D.,Dezani-Ciancaglini,M. & van Bakel.S.,(2004)"Intersection types for explicit substitutions", Information and Computation, Vol 189, No.1, pp 17 – 42.

[14] Dezani-Ciancaglini, D., Honsell, F. &Motohama, Y.,(2005)"Compositional characterisations of λ-terms using intersection types", Theoretical Computer Science, Vol. 340, No. 3, pp 459 – 495.

[15] Santo,J. E. &Ghilezan, S.,(2017)"Characterization of strong normalizability for a sequent lambda calculus with co-control", In: Proceedings of the 19th International Symposium on Principles and Practice of Declarative Programmin, pp 163–174.

[16] Koletsos,G.,(2012)"Intersection Types and Termination Properties",FundamentaInformaticae, Vol. 121, No.1-4, pp 185–202.

[17] Barendregt, H., Coppo, M. &Dezani-Ciancaglini,M. (1983) A filter lambda model and the completeness of type assignment, The Journal of Symbolic Logic, Vol. 48, No.4,pp 931–940.

[18] Dougherty, D. &Lescanne, P.,(2003) Reductions, intersection types, and explicit substitutions, Mathematical Structures in Computer Science, Vol. 13, No.1, pp 55-85.

[19] Ventura, D. L., Ayala-Rincón,M., &Kamareddine,F., (2009) Intersection Type System with de Bruijn Indices, The Many Sides of Logic, Sudies in Logic Vol. 21, W., Carnielli,Coniglio, M. E. and D'Ottaviano, I. M. L., eds. pp. 557-576.

[20] Ventura,D. L.,Kamareddine,F.& Ayala-Rincón,M., (2015) Explicit substitution calculi with de bruijn indices and intersection type systems, Logic Journal of the IGPL, Vol. 23, No. 2, pp 295–340.

[21] De Carvalho,D.,(2009)"Execution time of lambda-terms via denotational semantics and intersection types", Mathematical Structures in Computer Science Conference

AUTHORS

Xinxin Shen received the BE degree in 2011 from Henan University. She is currently a Master student at College of Computer Science and Technology, Zhejiang University, China. Her area of interests are logic, lambda calculus and type theory.

Kougen Zheng received the B.E. degree in 1986 and the PhD in 1990 from Warwick University, UK. He has 23 years of teaching experience and 7 years of industry experience. He is presently working as professor in Zhejiang University, China. His areas of interest include Artificial Intelligence, Rail Traffic Signal Processing, Logic, Theory of Computation.

ARRAY FACTOR IN CURVED MICROSTRIPLINE ARRAY ANTENNA FOR RADAR COMMUNICATION SYSTEMS

Putu Artawan[1,2] Yono Hadi Pramono[1], Mashuri[1] and Josaphat T. Sri Sumantyo[3]

[1]Physics Department, Faculty of Natural Sciences, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia
[2]Physics Department, Faculty of Mathematics and Natural Sciences, Ganesha University of Education, Singaraja, Bali, Indonesia
[3]Josaphat Microwave Remote Sensing Laboratory, Center for Environmental Remote Sensing (CEReS) Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522-Japan

ABSTRACT

This paper presents the designed of varians *array in curved microstripline* antenna *for radar communication. The antenna geometry comprises of three varians in matrics 2x2, 2x4 and 4x4 dimensions. The several array operates in C-Band frequencies (4GHz – 8GHz) and X-Band frequencies (8GHz-12GHz) with a 1.82 VSWR, -18.72dB Return loss,* 0.29 *reflection coefficient, and 5.8dB gain for 2x2 array, 1.64 VSWR, -16.17dB Return loss,* 0.24 *reflection* coefficient*, and 5.4dB gain for 2x4 array, 1.04 VSWR, -37.70dB Return loss,* 0.19 *reflection* coefficient*, and 7.6dB gain for 4x4 array. All of the varians in array elements are feed using a direct feeding technique. This array antenna is suitable developed for use in radar communication systems.*

KEYWORDS

Array, Curved microstripline, Radar communication, C-Band, X-Band

1. INTRODUCTION

Researches on Design of Antenna had conducted previously. The variation of design, the analysis, and the result obtained were not be optimal to be conducted in its application especially in radar communication systems. In this research, 2x2, 2x4, and 4x4 varian arrays in curved microstripline antenna was designed. In radar communications, there is a emphasizing need to minimize the size, weight and power requirments of antenna in array varians. The concept of array antennas with widely separated frequencies bands. Thus, in this research multiband frequencies antenna elements have been designed and fabricated that can potentially to develope radar communication systems. A novel design in microstrip antenna is one type of antenna that is designed using a PCB (Printed Circuit Board). Microstrip antennas are physically components that are designed to emit and or receive electromagnetic waves. In this antenna design also considerated the size of feeding stripline to reach the maximum results in antenna indicator

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019
pp. 159–172, 2019. © CS & IT-CSCP 2019 DOI: 10.5121/csit.2019.90314

parametric. As a goal of this research is to find the optimum indicator parametric results in S_{11}, Voltage Standing Wave Ratio (VSWR), Return Loss, Reflection Coefficient and gain that can be develope in radar application communications.

2. THEORIES

2.1 Microstrip Antenna

Microstrip antennas are electrically thin, lightweight, comformable, low cost, easily fabricated and can be connected to Microwave Integrated Circuits (MICs) at various frequencies [1]. There are various types of microstrip antenna designs on the taper section. There is a rectangular, circular, triangle shape according to the empirical analysis of antenna design. The design of the antennas varies with the single side and the double side. This study designed novel curved microstripline antenna with 2x2, 2x4 and 4x4 array, to produce greater gain so that it could be more optimally applied to radar communication systems.

2.2 Array Factor

Microstrip antennas arranged in Array are not only useful for widening bandwidth but also have an impact on the radiation pattern produced. The radiation pattern in the Antenna is generally written with the equation:

$$R\ (\theta,\emptyset) \text{ with i element in the position of } r_i = (x_i, y_i, Z_i)$$

The relationship with the wave emitted from the antenna array (Y) with the multiplier of complex numbers (wi) in the function (θ, ∅), is obtained:

$$Y = R\ (\theta,\emptyset) w_1 e^{-j\,k\,.\,r_1} + R\ (\theta,\emptyset) w_2 e^{-j\,k.r_2} + \cdots R(\theta,\emptyset) w_N e^{-j\,k\,r_N}$$

With k is the wave vector in the incoming wave.
Next can be written:

$$Y = R\ (\theta,\emptyset) \sum_{i=1}^{N} w_i e^{-j\,k\,.r_i}$$

$$Y = (\theta,\emptyset)\ AF$$

$$; AF = \sum_{i=1}^{N} w_i e^{-j\,k\,.r_i}$$

AF = Array Factor (as an Antenna position function).

2.3 Design Overview

Calculation of Antenna Dimensions uses the following steps:

The first calculating is to find the total electricity permittivity (ε_{rtot}) using the capacitor equation:

$$\frac{1}{c_{tot}} = \frac{1}{c_1} + \frac{1}{c_2}$$

$$\frac{1}{\varepsilon_o \varepsilon_{r\ tot}\ A/d\,tot} = \frac{1}{\varepsilon_o \varepsilon_{r_1}\ A_1/d_1} + \frac{1}{\varepsilon_o \varepsilon_{r_2}\ A_2/d_2}$$

where ε_{r1} is ε_r for air ($\varepsilon_{r1} = 1$), ε_{r2} is ε_r for substrate (ε_r $FR_4 = 4.3$), d_1 thick of substrate and d_2 distance of substrate to the reflector, with dtot is d_1+d_2.

And then using the following equation:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + 10\frac{h}{w}\right)^{-0,555}$$

To calculate the effective permittivity electricity (ε_{eff}). Where ε_r is the same with εr_{tot}, h is d_{tot} and w is the various wide for patch and stripline side.

The following equation is to know the maximum dimension in the patch side (w_1):

$$f = \frac{2c}{3w\sqrt{\varepsilon_r}}$$

$$w_1 = \frac{2c}{3\sqrt{\varepsilon_r}\,f}$$

where c is lightspeed in air, ε_r is electricity permittivity and f is frequency.

And to calculate the effective width stripline side ($W_{2,3}$), using the following figure:

$$W_{2,3} = \frac{1}{2f\sqrt{\mu_0 . Z_o}}\sqrt{\frac{2}{\varepsilon_r + 1}}$$

Where f is frequency, μo is permeability constant and Zo is characteristic impedance.

The calculation wavelength of the substrate (λg), using the following equation:

$$\lambda g = \frac{\lambda}{\sqrt{\varepsilon_{eff}}}$$

From the analysis above we find to fix the parameter of antenna fabrication.

The following figure is the Curved Microstripline Antenna Design in array variant.

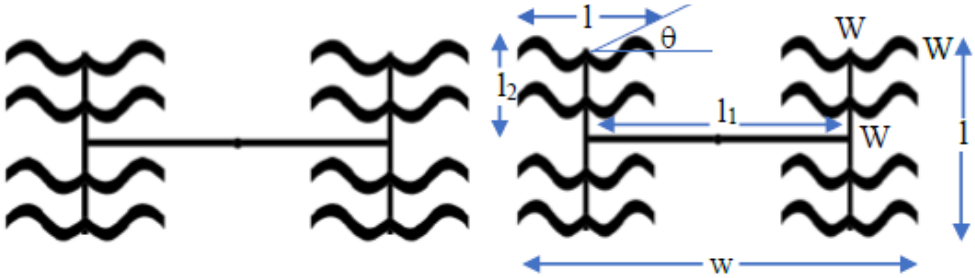


Figure 1. (a) Curved Microstripline 2x2 Array Antenna Design (b) Dimension of Curved Microstripline 2x2 Array Antenna Design.

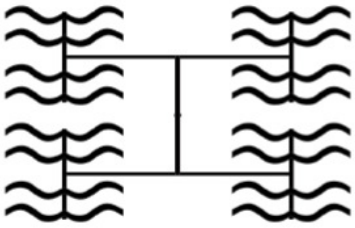


Figure 2. Curved Microstripline 2x4 Array Antenna Design.

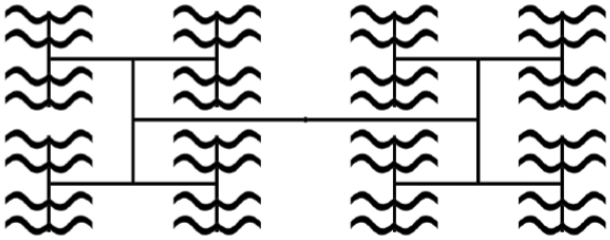
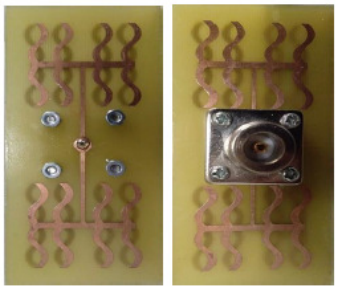


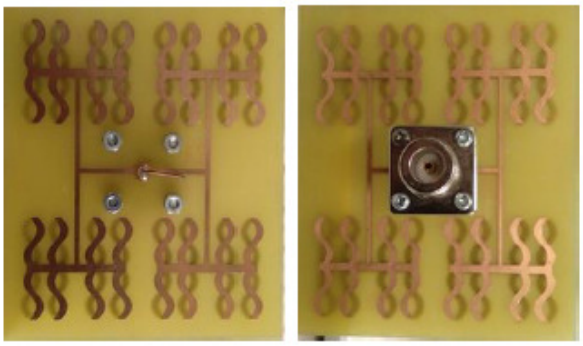
Figure 3. Curved Microstripline 4x4 Array Antenna Design.

3. FABRICATION, SIMULATION AND MEASUREMENTS RESULT

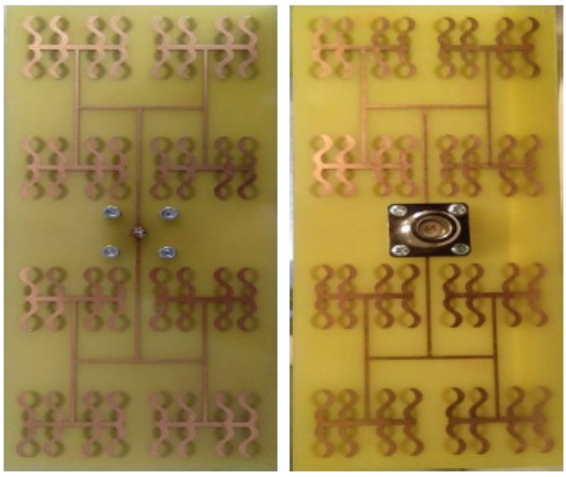
The following figure is the result of fabrication in curved microstripline array varians antenna.



(a) 2x2 Array Varians.



(b) 2x4 Array Varians.



(c) 4x4 Array Varians.

Figure 4. Curved microstripline array varians antenna prototype. (a). 2x2 Array Varians. (b) 2x4 Array Varians. (c) 4x4 Array Varians.

Curved microstripline array varians antenna prototype was fabricated by UV photoresist laminate. In our work, the antenna prototypes are fabricated on Flame Retardant 4 (FR_4) material with 4.3 dielectric constant. The first step in the fabrication process is to generate the photo mask artwork by printing on stabline or rubylith negative film of the desired geometry on butter sheet. Using the precision cutting blade of a manually operated co-ordino graph the opaque layer of the stabiline or rubylith film is cut to the proper geometry and can be removed to produce either a positive or negative film representation of the antenna sketches. The design dimensions and tolerances are verified on a cordax measuring instrument using optical scanning. Enlarged artwork should be photo reduced using a high precession camera to produce high resolution negative, which is later used for exposing the photo resist. The photographic negative must be now held in very close contact with the polyethylene cover sheet of the applied photo resist using a vacuum frame copy board or other technique, to assure the fine line resolution required. With exposure to proper wavelength of light, polymerization of the exposed photo resist occurs making it insoluble in the developer solution. Now, it is then coated with a negative photo resist and exposed to UV-radiation and it is immersed in developer solution up to two minutes through the

mask. The exposed photo resist hardens and those in the unexposed areas are washed off using a developer. The unwanted copper portions are now removed using Ferric Chloride ($FeCl_3$) solution. $FeCl_3$ dissolves the copper coating on the laminate except which is underneath the hardened photo resist layer after few minutes. Finally, the laminate is then washed with water and cleaned in acetone solution to remove the hardened negative photo resist. The fabrication process has shown in the following figure.

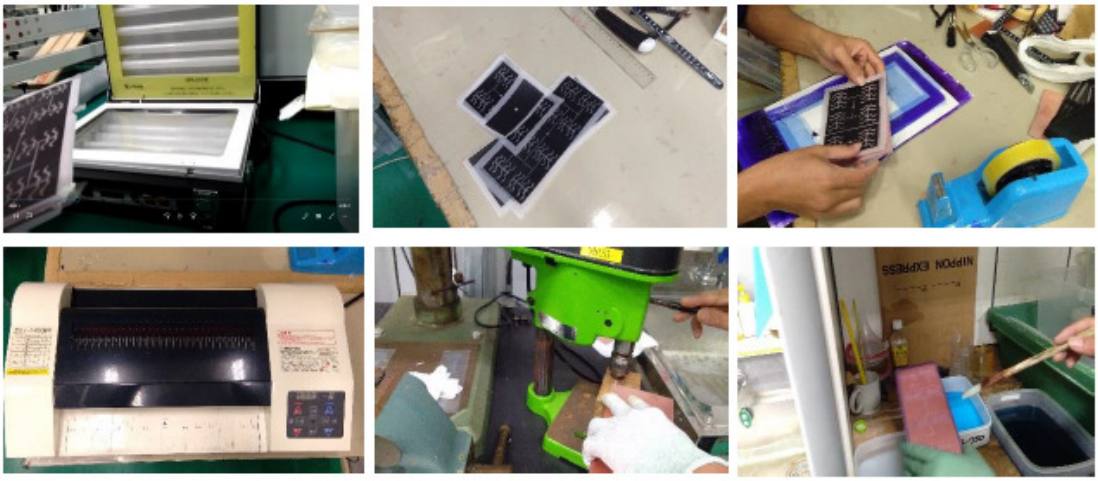
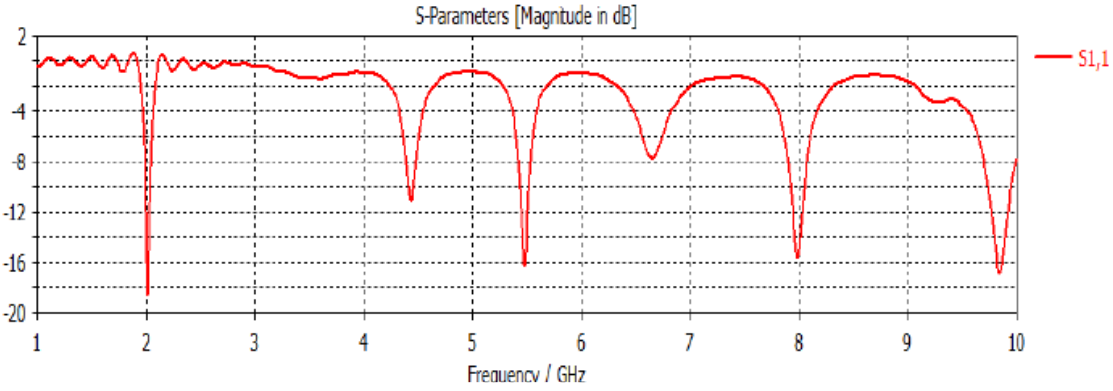
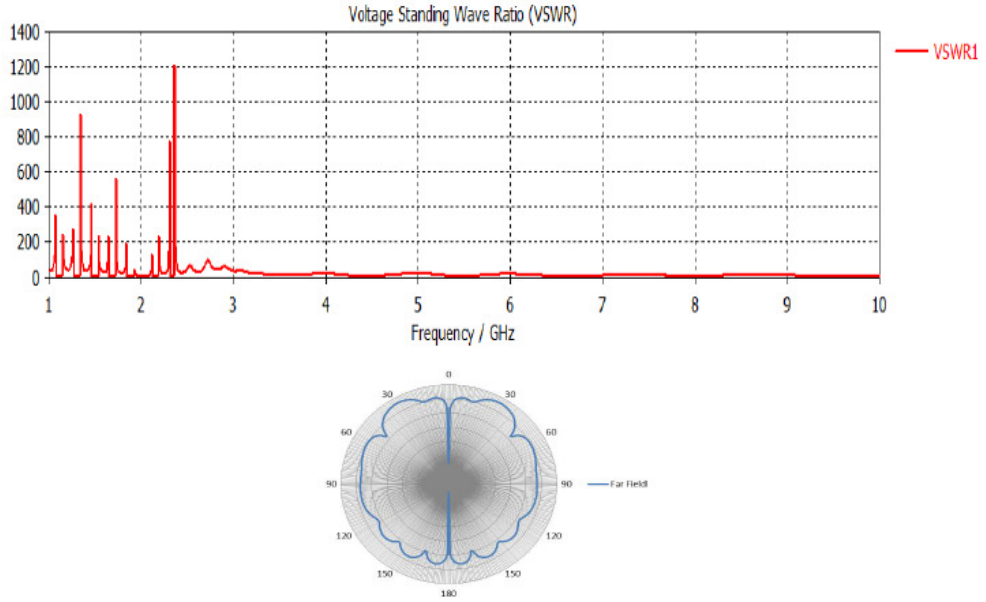


Figure 5. Fabrication Process.

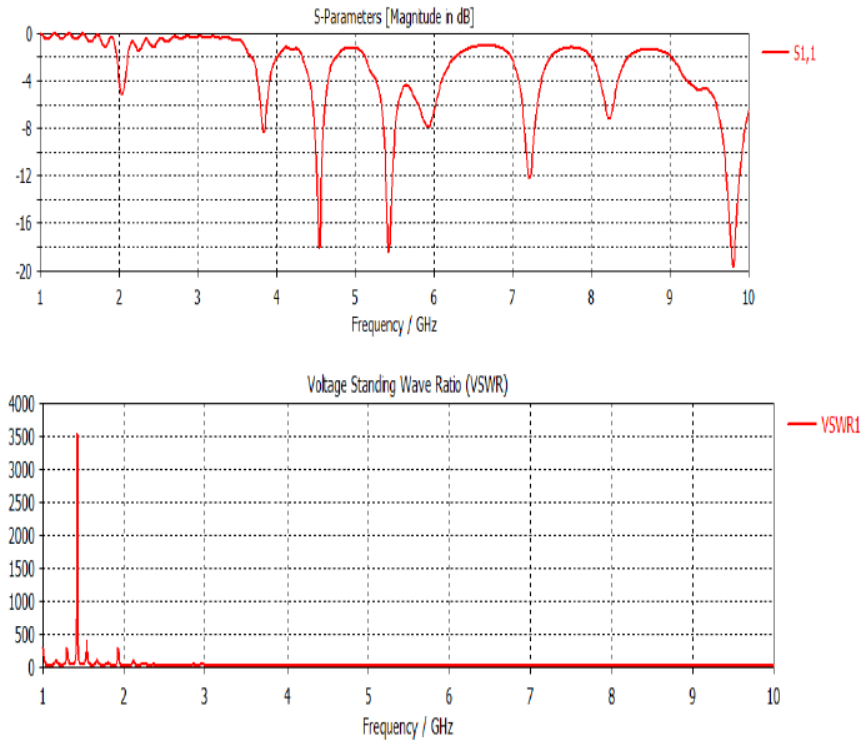
The curved microstripline array varians antenna shown in Figure 6 has been modeled in CST programme to determine S_{11} parameter, Voltage Standing Wave Ratio (VSWR), Return Loss, Reflection Coefficient and Radiation Pattern.

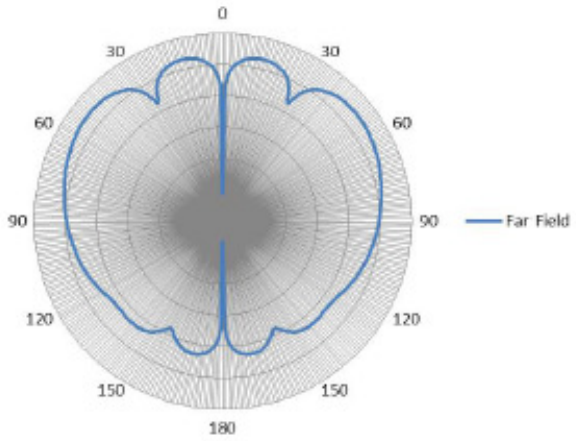




(a) S_{11} Parameter. (b) VSWR. (c) Radiation Pattern.

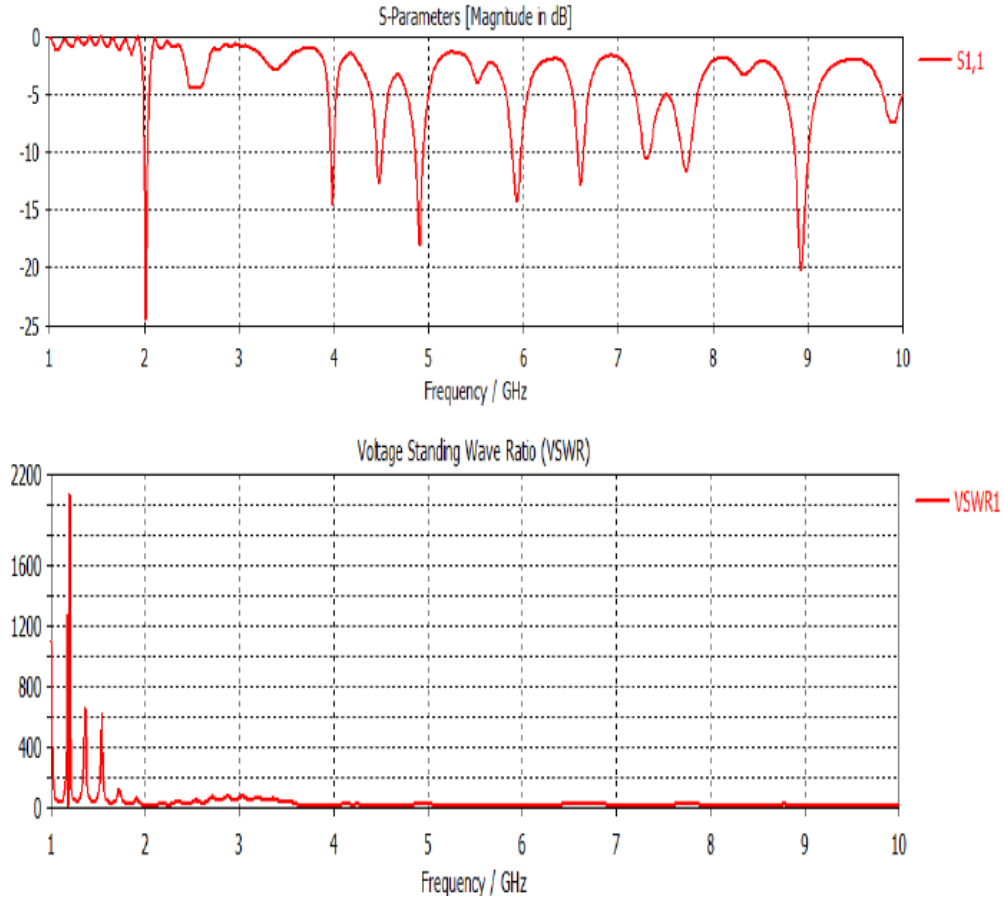
Figure 6. Curved microstripline 2x2 array varians antenna simulations. (a) S_{11} Parameter Simulation. (b) VSWR Simulation. (c) Radiation Pattern Simulation.

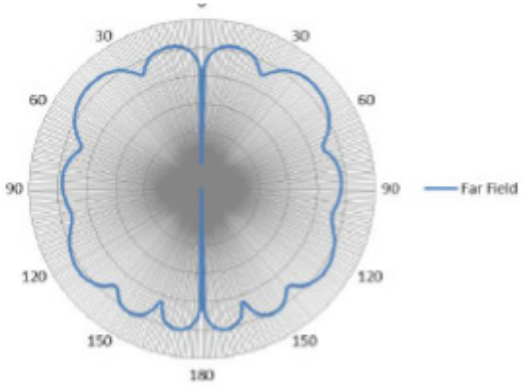




(a) S11 Parameter. (b) VSWR. (c) Radiation Pattern.

Figure 7. Curved microstripline 2x4 array varians antenna simulations. (a) S_{11} Parameter Simulation. (b) VSWR Simulation. (c) Radiation Pattern Simulation.

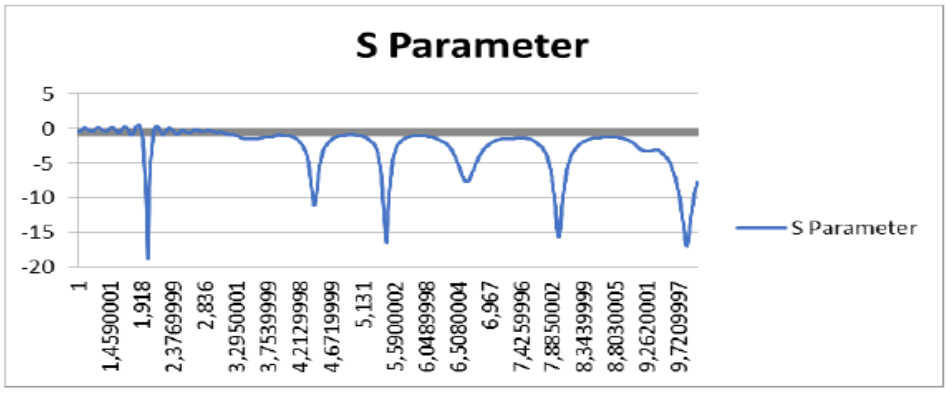




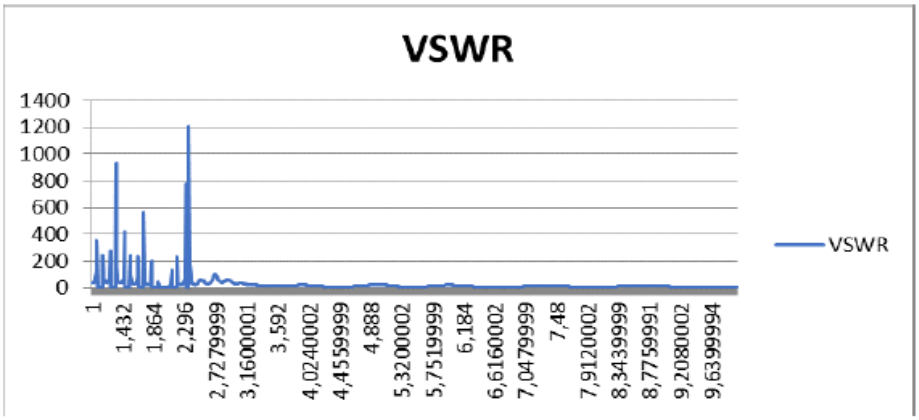
(a) S_{11} Parameter. (b) VSWR. (c) Radiation Pattern.

Figure 8. Curved microstripline 4x4 array varians antenna simulations. (a) S_{11} Parameter Simulation. (b) VSWR Simulation. (c) Radiation Pattern Simulation.

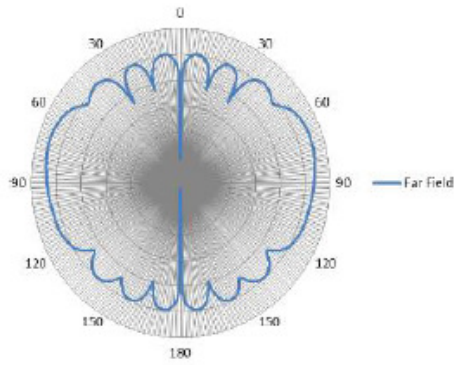
Antenna measurement using a Network Analyzer type device the Agilent 8510 Vector Network Analyzer. With result:



(a)



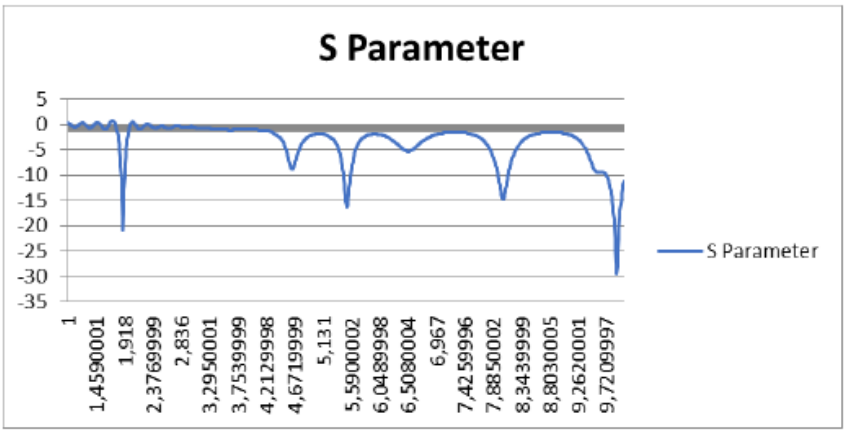
(b)



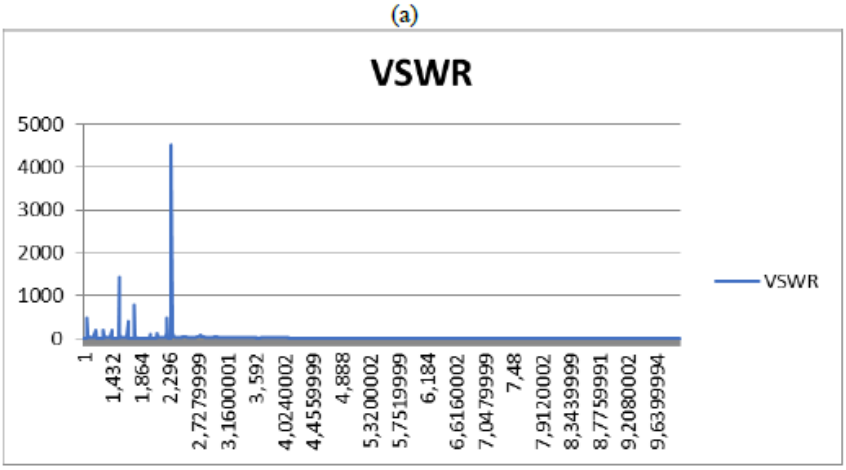
(c)

(a) S_{11} Parameter. (b) VSWR. (c) Radiation Pattern.

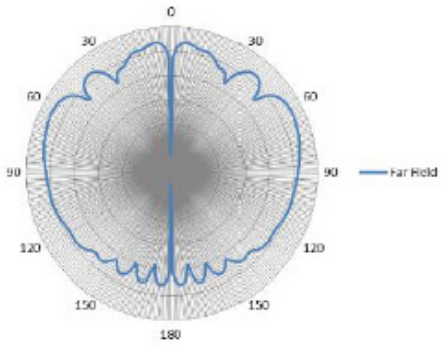
Figure 9. Curved microstripline 2x2 array varians antenna measurements. (a) S_{11} Parameter Measurement. (b) VSWR Measurement. (c) Radiation Pattern Measurement.



(a)



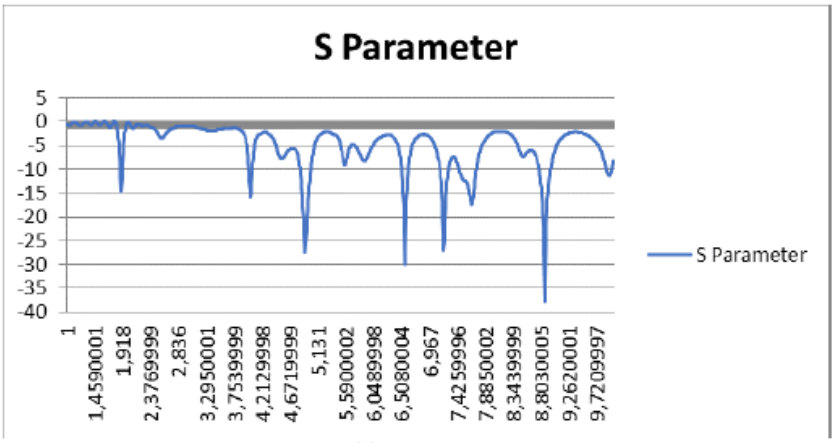
(b)



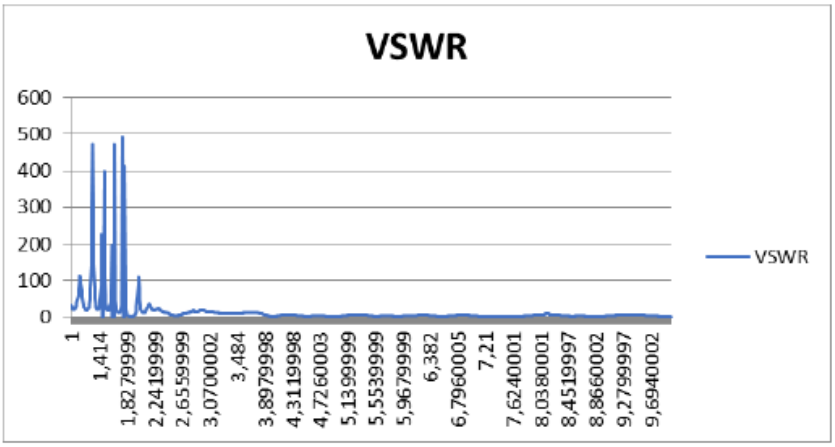
(c)

(a) S_{11} Parameter. (b) VSWR. (c) Radiation Pattern.

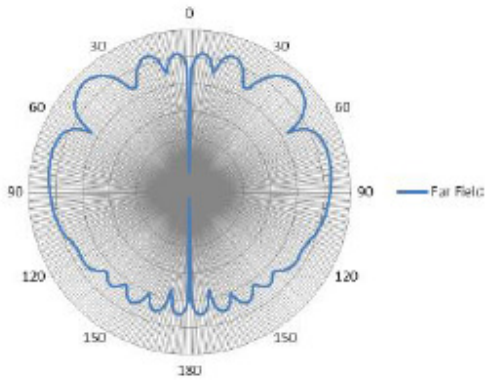
Figure 10. Curved microstripline 2x4 array varians antenna measurements. (a) S_{11} Parameter Measurement. (b) VSWR Measurement. (c) Radiation Pattern Measurement.



(a)



(b)



(c)

(a) S_{11} Parameter. (b) VSWR. (c) Radiation Pattern.

Figure 11. Curved microstripline 4x4 array varians antenna measurements. (a) S_{11} Parameter Measurement. (b) VSWR Measurement. (c) Radiation Pattern Measurement.

In general, base on the simulation and the measurement results, curved microstripline array varians antenna has the optimal parametric characteristic as a good requirements antenna that can be develop in radar communication applications. The characteristics of Curved microstripline array varians antenna can be described in the following table.

Indicator	Value	Standard Parametric
VSWR	Array 2 x 2 1.82 Array 2 x 4 1.64 Array 4 x4 1.04	1 < VSWR < 2
Return Loss	Array 2 x 2 -18.72 Array 2 x 4 -16.17 Array 4 x4 -37.70	≤ - 10 dB
Reflection Coefficient	Array 2 x 2 0.29 Array 2 x 4 0.24 Array 4 x4 0.19	Close to 0
Gain	Array 2 x 2 5.8 Array 2 x 4 5.4 Array 4 x 4 7.6	> 4dB

Table 1. Curved microstripline array varians antenna characteristics.

4. CONCLUSION

Adding Arrays to the Curved Microstripline Antenna gives more optimal results especially in the range of working frequencies (bandwidth) and also the resulting gain. The results indicate that the antenna is able to apply in multiband frequency. The radiation pattern produced in this design is Omnidirectional in linear polarization. The band frequencies array in this design is capable to develope in radar application communication systems.

ACKNOWLEDGMENT

The authors would like to thank the Indonesian Ministry of Research, Technology and Higher Education through LPDP and PKPI (Sandwich-like) scholarships, Center for Environmental Remote Sensing (CEReS), Josaphat Tetuko Sri Sumantyo (JMRSL Chiba University), Promotor Yono Hadi Pramono and Mashuri (Physics Department, ITS Surabaya).

REFERENCES

[1] Artawan. Fabrikasi dan Karakterisasi Antena Mikrostrip Tapered Patch Untuk Aplikasi Antena Panel Pada Frekuensi 2,4GHz. Tesis Magister, Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, 2011.

[2] Artawan, Hadi Pramono, Yono. Perancangan Antena Panel Mikrostrip Horn Array 2x2 Untuk Komunikasi Wi-Fi Pada Frekuensi 2,4GHz. Prosiding Simposium Fisika Nasional (SFN), ITS, Surabaya, 2010.

[3] Balanis, C.A. Antena Theory Analysis and Design. Second Edition, John Wiley and Sons, New York, 1997.

[4] Edward, Terry. Foundation For Microstrip Circuit Design. Knaresborough England, 1991.

[5] Shafai. Microstrip Antena Design Handbook. Profesor University Of Manitoba, Wimmipeg, Canada, 2001.

[6] Kraus, John, D. Electromagnetics. Third Edition, McGraw-Hill, New York, 1984.

[7] Ohri, V, Amin, O, Gebremariam, H Dubois, B. Microwave Mikrostrip Horn Antena Design and Test System. San Jose State University, 2003.

[8] Masduki, K. Desain, Fabrikasi dan Karakterisasi Antena Mikrostrip Biquad dengan CPW (Coplanar Waveguide) pada Frekuensi Kerja 2,4GHz. Program Magister Bidang Keahlian Optoelektronika Jurusan Fisika, FMIPA-ITS: Surabaya. 2009.

[9] Hund, E. Microwave Communications, Component and Circuit. McGraw Hill, New York, 1989.

[10] Hadi Pramono, Yono. Karakterisasi Antena Mikrostip Patch 3GHz Secara Simulasi FDTD (Finite Difference Time Domain) Dan Eksperimen. Jurnal Fisika. Institut Teknologi Sepuluh Nopember. Surabaya, 2005.

[11] Hadi Pramono, Yono. Prototipe Antenna Bi-Mikrostrip Tapered Patch dengan Dua Arah Pola Radiasi Dan Satu Feeding Monopole Beroperasi Pada Freq.2,4GHz. Prosiding T. Informatika, UPN. Yogyakarta, 2009.

[12] Hidayah, Ifa. Desain dan Fabrikasi Antena Bi-Mikrostrip Tapered Patch dengan Dua Arah Radiasi dan Satu Feeding Monopole Untuk Komunikasi Wi-fi. Tesis Magister. Institut Tekologi Sepuluh Nopember. Surabaya, 2009.

[13] Naqiah, Hawaun. Fabrikasi dan Karakterisasi Antena Mikrostrip Loopline untuk Komunikasi Wireless Local Area Network (WLAN). Program Magister Bidang Keahlian Optoelektronika Jurusan Fisika FMIPA-ITS: Surabaya, 2009.

[14] Risfaula, Erna. Antena Mikrostrip Panel Berisi 5 Larik Dipole dengan Feedline Koaksial Waveguide untuk Komunikasi 2,4GHz. Program Keahlian Optoelektronika Jurusan fisika FMIPA-ITS: Surabaya, 2011.

[15] S. Gao, Q. Luo, F. Zhu, Circularly Polarized Antenna, John Wiley & Sons, Ltd, 2014.

[16] S. Murugan, V. Rajamani, "Study of Broadband Circularly Polarised Microstrip Antenna" Science Engineering and Management Research (ICSEMR), 2014 International Conference on. IEEE, 2014.

[17] Haider Raad, "An UWB Antenna Array for Flexible IoT Wireless System," Progress In Electromagnetics Research, Vol. 162, 109-121, 2018.

[18] Kurniawan Farohaji, Sri Sumantyo, J. T, Gao Steven, Ito Koichi, Edi Santosa C. "Square-Shaped Feeding Truncated Circularly Polarised Slot Antenna". IET Microwaves, Antenna & Propagation Journals. ISSN 1751-8725, 2018.

AUTHORS

Putu Artawan received his Bachelor Degree (Physics Education, IKIP N Singaraja, Bali,Indonesia) Master degree in (Physics Department, Faculty of Mathematics and Natural Sciences, ITS Surabaya, Indonesia) and received his doctoral degree in Physics Department, Faculty of Natural Sciences, ITS Surabaya, Indonesia), presently he working as lecturer in Ganesh Universitu of Education Sinagaraja bali. He has published several articles in reputed Journals And conferences.

AUTHOR INDEX