Faculty of Information and Communication Technology

CLASSIFICATION OF GENDER USING GLOBAL LEVEL FEATURES IN FINGERPRINT FOR MALAYSIAN POPULATION

Siti Fairuz Binti Abdullah

Master of Science in Information and Communication Technology

2016

CLASSIFICATION OF GENDER USING GLOBAL LEVEL FEATURES IN FINGERPRINT FOR MALAYSIAN POPULATION

SITI FAIRUZ BINTI ABDULLAH

A thesis submitted
In fulfilment of the requirements for the degree of Master of Science
in Information and Communication Technology

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2016

DECLARATION

I declare that this thesis entitled "Classification of Gender using Global Level Features in Fingerprint for Malaysian Population" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :

Name : Siti Fairuz Binti Abdullah

Date :

APPROVAL

I hereby declare that I have read this thesis and in my opinion, this thesis is sufficient in terms of scope and quality as a partial fulfilment of Master of Science in Information and Communication Technology.

Signature	:	
Supervisor Name	:	Mr. Ahmad Fadzli Nizam bin Abdul Rahman
Date	:	

DEDICATION

Sabda Rasulullah (SAW):
"Sesungguhnya para malaikat membentangkan sayapnya kepada penuntut ilmu tanda reda (suka) dengan apa yang dilakukannya itu" (Riwayat Ahmad)

Ya Allah, dengan segala limpah rahmatMu dan izinMu terhasilnya semua ini.
Khas buat yang tersayang;
Abdullah Baharom dan **Siti Meryam Jiwa**,
Penyelia merangkap sahabat dan saudara;
Encik Ahmad Fadzli Nizam Abdul Rahman, **Dr. Zuraida Abal Abas,** dan **Dr. Wira Hidayat Mohd Saad** yang banyak memberi tunjuk ajar, membimbing, membantu, menasihat dan sabar melayan kerenah seorang pelajar.
Seterusnya adik-beradik yang tercinta **Norazaliza**, **Siti Rokhiyah**, **Siti Muniroh**, **Muhammad Firdaus** dan **Siti Farahnasihah**, sahabat-sahabat, saudara mara dan beberapa orang yang sentiasa ada di sisi dulu, kini dan selamanya.
Sesungguhnya, kejayaan ini adalah milik kita bersama.

Semoga kejayaan ini memberikan kebaikan dan menjadi pemangkin untuk kita terus berjaya dan cemerlang di dunia mahupun di akhirat. Insya-Allah.

"Ya Allah, sembuhkanlah jiwa-jiwa yang berduka,
Kembali berbicara dan mengukir senyuman."

"Ya Allah, tetapkanlah jiwa hati kami semua,
Sehingga bertemuMu menatapMu,
Nanti di sana"

"Ya Allah, temukanlah pengakhiran yang terindah,
Sehingga bertemuMu,
MenghadapMu di Syurga"

Amin Ya Rabb.

ABSTRACT

A new approach of algorithm based on the Mark Acree's theory, focusing on fingerprint global extracted features is proposed and implemented for enhancing gender classification method. This proposed method can automatically execute the ridge calculation process from the 25mm^2 fingerprint and enhance the forensic gender classification process. In this study, a relationship between fingerprint global features and a gender of person in Malaysian population is also explored, enhanced and improved by exploiting another five additional fingerprint features. A sample of 3000 fingerprints from 300 respondents of random selection are carefully taken before any relationship can be determined. For the classification part, five extracted features of the fingerprint are used which are Ridge Density (RD), Mean Ridge Count (RC), Ridge Thickness to Valley Thickness Ratio (RTVTR), White Lines Count (WLC) and Mean Pattern Types (PT). Two classification approaches which are the descriptive statistical and data mining are used in order to examine the classification of the gender by using the five extracted features. For data mining classification part, there are four popular machine learning classifiers used which are Bayesian Net.work (Bayes Net.), Multilayer Perceptron Neural Network (MLPNN), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). These four classifiers are used in the data mining task with five test cases each in order to find the accuracies of the gender classification. The accuracy of the results from the proposed method is compared to the Acree Method is shown in terms of relative error. For statistical approach using Ridge Density (RD), the relative error is 3.7% for male respondent and 4.1% for female respondent. Meanwhile, the overall performance of the result from the proposed method achieved more than 90% classification rate for all the classifiers. SVM emerges as the best classifier for all the different cases in order to classify the gender using the results from the proposed method.

ABSTRAK

Satu algoritma baru berdasarkan teori Mark Acree yang memberi tumpuan kepada pengekstrakan ciri-ciri global cap jari dicadangkan dan dilaksanakan untuk meningkatkan kaedah dalam pengelasan jantina. Kaedah yang dicadangkan ini boleh melaksanakan kaedah pengiraan rabung jari di dalam kawasan 25mm^2 secara automatik dan membantu kaedah pengelasan jantina di dalam bahagian forensik. Dalam kajian ini, hubungan ciri-ciri global cap jari dan jantina dalam kalangan penduduk Malaysia juga diterokai, dipertingkatkan dan diperbaiki dengan mengeksplotasi lima ciri-ciri tambahan cap jari. Sampel kajian terdiri terdiri daripada 3000 cap jari yang diambil secara rawak dan teliti daripada 300 orang responden. Untuk bahagian pengelasan, lima ciri-ciri cap jari telah diekstrakkan iaitu Ketumpatan Rabung (RD), Bilangan Rabung (RC), Nisbah Ketebalan Rabung dan Ketebalan Lembah (RTVTR), Bilangan Lembah Putih (WLC) dan purata Jenis Corak Cap Jari (PT). Dua jenis pendekatan dalam pengelasan yang digunakan iaitu deskritif statistik dan perlombongan data digunakan untuk pengelasan jantina. Empat jenis pengelas telah digunakan dalam kaedah mengklasifikasi iaitu Bayesian Network (Bayes Net.), Multillayer Perceptron Neural Network (MLPNN), K-Nearest Neighbor (KNN), dan Support Vector Machine (SVM) dan setiap pengelas akan diuji dengan lima ujian yang berbeza bagi mendapatkan ketepatan yang terbaik. Hasil ketepatan dari kaedah yang dicadangkan dibandingkan dengan kaedah Mark Acree ditunjukkan dari segi ralat relatif. Untuk pendekatan statistik, ketumpatan rabung (RD) mempunyai nilai ralat relatif sebanyak 3.7% dari responden lelaki dan 4.1% dari responden perempuan. Sementara itu, hasil prestasi keseluruhan daripada pengelasan perlombongan yang digunakan adalah melebihi 90% kadar pengelasan untuk kesemua pengelas. SVM muncul sebagai pengelas terbaik dalam pengelasan jantina untuk kaedah yang dicadangkan.

ACKNOWLEDGEMENTS

"In the name of Allah, the Most Gracious, Most Compassionate".

All the praises and thanks be to **Allah SWT** for His uncountable blessings, and best prayers and peace be unto His best messenger **Nabi Muhammad SAW**, his pure descendent, and his family and noble companions.

First, I would like to thank my family, especially my lovely mom and dad, **Mr Abdullah Baharom** and **Mrs. Siti Meryam Jiwa**. Without their support and love over the years, none of this would have been possible. They have always been there for me and I am thankful for everything they have helped me to achieve.

Next, I would like to thank my super duper diligent and supportive supervisor, **Mr**. **Ahmad Fadzli Nizam Abdul Rahman**. Your helps and guidances over the years, which are immeasurable and without it, I would not be where I am today. I would like to thank my co-Supervisor too, **Dr. Zuraida Abal Abas** and not forget **Dr. Wira Hidayat Mohd Saad** and **all members and lecturers from Optimization, Modelling, Analysis, Simulation and Scheduling (OptiMASS)** for their assistance and continuous support in many aspects of this research. This work would not have been possible without their help and input.

I would also like to express my thanks to the **all staff and lecturers of Faculty of Information and Communication Technology** for their support and also thanks to **MyBrain UTeM** for the financial support during the research was conducted.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

A	Arch
ANN	Artificial Neural Network
DL	Dual Loop
DCT	Discrete Cosine Transform
DLRT	Dual Loop Right Tentarch
DLLT	Dual Loop Left Tentarch
DWT	Discrete Wavelet Transform
FO	Favored Odd
FCM	Fuzzy C-Means
FFT	Fast-Fourier Transform
GMM	Gaussian Mixture Model
KNN	K-Nearest Neighbor
L	Loop
LL	Left Loop
LR	Likelihood Ratio
LDA	Linear Discriminat Analysis
LPL	Left Pocket Loop
LLRT	Left Loop Right Tentarch
MLPNN	Multilayer Perceptron Neural Network

NB	Naïve Bayes
PCA	Principle Component Analysis
PSD	Power Spectral Density
RC	Ridge Count
RD	Ridge Density
RL	Right Loop
RPL	Right Pocket Loop
RLLT	Right Loop Left Tentarch
RTVTR	Ridge Thickness to Valley Thickness Ratio
SD	Standard Deviation
SVD	Singular Value Decomposition
T	Tentarch
W	Whorl
WA	Whorl Arch
WLC	White Lines Count
WRT	Whorl Right Tentarch
WLT	Whorl Left Tentarch
WRTLT	Whorl Right Tentarch Left Tentarch

LIST OF PUBLICATIONS

1. Abdullah, S. F., Rahman, A. F. N. A., & Abas, Z. A. 2015. Classification Of Gender By Using Fingerprint Ridge Density In Northern Part Of Malaysia. *ARPN Journal of Engineering and Applied Sciences*, 10(22).
Published : ARPN Journal of Engineering and Applied Sciences (Indexed Scopus)

2. Abdullah, S. F., Rahman, A. F. N. A., Abas, Z. A., & Saad, W. H. M. 2016. Multilayer Perceptron Neural Network in Classifying Gender using Fingerprint Global Level Features. *Indian Journal of Science and Technology*, *9*(9).
Published : Indian Journal of Science and Technology (Indexed ISI)

3. Abdullah, S. F., Rahman, A. F. N. A., Abas, Z. A., & Saad, W. H. M. 2016. Development of a Fingerprint Gender Classification Algorithm Using Fingerprint Global Features. *International Journal of Advanced Computer Science and Applications(IJACSA), 7(6).*
Published : International Journal of Advanced Computer Science and Applications (IJACSA) (Indexed ISI)

4. Abdullah, S. F., Rahman, A. F. N. A., Abas, Z. A., & Saad, W. H. M. 2016. Support Vector Machine, Multilayer Perceptron Neural Network, Bayes Net and k-Nearest Neighbor in Classifying Gender using Fingerprint Features. *International Journal of Computer Science and Information Security (IJCSIS), 14(7).*
Published : International Journal of Computer Science and Information Security (IJCSIS) (Indexed ISI)

5. Abdullah, S. F., Rahman, A. F. N. A., Mahadhir, M., & Abas, Z. A. 2016. Fingerprint Based Gender Classification Using DWT and KNN. *International Journal of Computer Science and Information Security (IJCSIS), 14(8).*
Published : International Journal of Computer Science and Information Security (IJCSIS) (Indexed ISI)

6. Abdullah, S. F., Rahman, A. F. N. A., Abas, Z. A., & Saad, W. H. M. 2016. Fingerprint Gender Classification Using Univariate Decision Tree (J48). *International Journal of Advanced Computer Science and Applications(IJACSA), 7(9).*
Published : International Journal of Advanced Computer Science and Applications (IJACSA) (Indexed ISI)

CHAPTER 1

INTRODUCTION

1.1 Introduction

Gender details is among an important data needed in order to do a further investigation that leads to finding an unknown person (Gornale et. al., 2015; Dongre et. al., 2015). Existing methods for gender classification have a limited use for crime scene investigation because they depend on the availability of teeth, bones, or other identifiable body parts having the physical features that allow gender determination by conventional methods (Gnanasivam and Muttan, 2012). Several studies have investigated the use of biometric modalities to determine gender, including the face (Makinen and Raisamo, 2008), gait (Shan and McOwan, 2007), iris (Thomas et. al, 2007), hand shape (Amayeh et.. al, 2008), fingertip (Wang et. al, 2008), finger length (Brown et. al., 2002) and fingerprint (Cummin et. al., 1961; Acree, 1999; Sudesh et. al., 2007; Gutierrez et. al., 2011, Kaur et.. al., 2011; Ping et. al., 2012; Kumar et. al., 2013; Nanakorn et. al., 2013; Rajan et. al., 2014; Ceyhan et. al., 2014; Rajesh et. al., 2014; Gupta et. al., 2014; Gornale et. al., 2015; Dongre et. al., 2015; Agrawal et. al., 2014).

For this research, the gender of a person is identified from the fingerprint because of the fingerprint is one of the ways in gender classification or recognition that used to minimize the criminals suspect list in forensic anthropology (Gomale and Kruthi, 2014). Fingerprint analysis plays a role in convicting the person responsible for an audacious crime and it has been considered as legitimate proofs of evidence in court of law in 1911.

Fingerprint is always associated with the criminology especially in forensics (Nte, 2012) and it has been used and accepted since 1975 as an important way to recognise a person's gender (Reddy, 1975). Fingerprint is always used for identification or verification of a person and for official documentation.

Nowadays, the fingerprint has been used as a biometric verification for the gender identification because of its unique nature and does not change throughout the life of an individual (Maltoni et. al., 2003). This is due to its their high acceptability, immutability, and uniqueness of the fingerprint itself. Immutability of the fingerprint refers to the pattern that remains unchanged over time whereas uniqueness related to the differences of individual ridge details across the whole fingerprint image (Bhuyan, 2012). These two characteristics make fingerprint highly acceptable and trusted. In any official documentation in banking and financial institution, the fingerprint images have been long widely used an identification or verification of a person.

1.2 Research Background

This research emphasises on the gender classification of a person by using fingerprint. Fingerprint data is very important in order to minimize the criminal suspect list and have been accepted and recognized as a classification method in forensic anthropology (Gomale and Kruthi, 2014). An illustration of the ridges from the fingertip epidermis always used as evidence in criminal investigations (Khan et. al., 2012). Forensic anthropology is a special field of physical anthropology that applying analysis and techniques to solving a criminal cases. Fingerprint analysis also plays an important role in convicting a person who commits a crime.

Existing method uses teeth, bones and other identifiable body states of the people to estimate the age and sex (Wadwa et. al., 2013). Existing application of fingerprint is

usually for person identification, but actually, it can be used more than that, for example, to identify age, race, blood group and gender of a person. Unfortunately, the process of latter identifying is hard, complex and takes time to process. A lot of steps involved need to be revised and tested before any conclusion can be made.

Fingerprint based gender classification involves forensic stage processes which are not well understood and lack of the organize classification procedures and more on a conceptual description on how to classify based on the structure fingerprint. Fingerprint classification is an important step in any fingerprint gender identification system because it reduces the time taken in identification of fingerprints, especially where the accuracy and speed are critical (Wadhwa et. al., 2013). Recently, many studies have been carried out the method of storing fingerprint in computer for rapid search and matching of fingerprint, but a few studies available on this method of feature extraction using ridges density (Acree, 1999; Nayak et. al., 2010).

The studies of gender classification by using the fingerprint have been carried out in several countries, but for Malaysian population, from the previous and the current publication of journals and conferences, there is one journal paper with the small population of Malaysian people has been carried out to determine the mean of ridge density (Nayak et. al. 2010).

Nayak reported an observation on gender difference of two population which is 200 Chinese and 100 Malaysian population which the respondent in aged between 18 to 25 years old. All the respondent were student that studying in Manipal University, India and this experiment is carried out with the smaller number of populations and it have been conducted in India. In order to get the actual distribution of Malaysian population, this

study must be conducted in Malaysia with the larger number of respondents, with random race, state and age of Malaysian respondents.

A fingerprint is the representation of the epidermis of a finger. It comprises of an example of interleaved edged and valleys as shown in Figure 1.1. Fingertip edges formed throughout the years to permit people to handle and grasp an object. Unique fingerprint edges structure through a mix of hereditary and ecological components impression shaping is like the development of vessels and veins in angiogenesis (Bharadi, 2014).

Figure 1.1: The ridge and valley

There are two levels of features in fingerprint structure, which are global and local features. The global level structures consist of many ridges to form some specific shape like an arch, loop, and whorl as shown in Figure 1.2 while the local level structures are called minutiae, which further classified as either endpoints or bifurcations as shown in Figure 1.3.

Arch Loop Whorl

Figure 1.2: Fingerprints global level structure

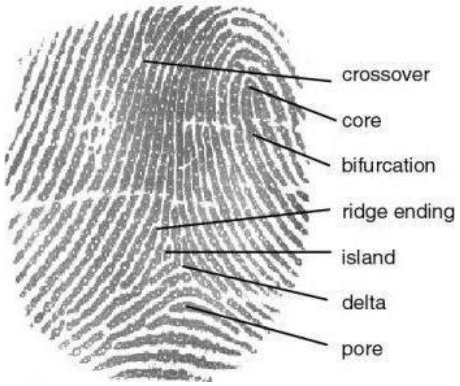

Figure 1.3: Fingerprint local level structures (Barcode, 2015)

Local ridge and valley details carry information about the individuality of a fingerprint, while global pattern features carry information about the fingerprint class (Nigeria, 2012). For this research, the global level features from a fingerprint, which is ridge density (RD), ridge count (RC), ridge thickness to valley thickness ratio (RTVTR), white lines count (WLC) and fingerprint pattern types (PT) are used to classify the gender.

In fingerprint, the ridges are formed during the gestational weeks 12-19 from the day of development of feotus, and the resulting fingerprint ridge configuration is fixed permanently (Mulvihill and Smith, 1969; Babler, 1991). Ridges and their patterns exhibit a number of properties that reflect the biology of individuals. Fingerprints are static and their size and shape changes may vary with age, but the basic pattern of the fingerprint which is