GENDER RECOGNITION BASED ON FACIAL IMAGE EXTRACTION

NURUL ZARINA BINTI MD ISA

This thesis is submitted as partial fulfilment of the requirements for the award of the Bachelor of Electrical Engineering (Hons.) (Electronics)

Faculty of Electrical & Electronics Engineering
Universiti Malaysia Pahang

NOVEMBER, 2010

"I hereby acknowledge that the scope and quality of this thesis is qualified for the award of the Bachelor Degree of Electrical Engineering (Electronics)"

Signature : _____

Name : MR. MOHD HISYAM BIN MOHD ARIFF.

Date : 29 NOVEMBER 2010

Signature : _____

Author : NURUL ZARINA BINTI MD ISA

Date : 29 NOVEMBER 2006

ABSTRACT

In principle, to combine face detection and gender classification methods may seem simple. However, this process is more complex than it appears because requires many aspects for consideration. The gender classification has attracted much attention in psychological literature, relatively few machine vision methods have been proposed. However it has been extensively studied in the context of surveillance applications and biometrics. This project is mainly concern with offline gender classification using purely image processing technique which using a database that was included in the system. The way of doing this is by extracting the differences between male and female facial features. Obviously the classification base on a single feature is not adequate since humans share many facial properties even within different gender group. So multilayer processing is needed. This project is working as expected based on the scope and objective of project. Although not many varieties of facial images have been considered like colored hair the basic techniques should be just the same. For the system classification, Template Matching Technique is used to match image with the database image. The system attempts are made to capture the most appropriate representation of face images as a whole and exploit the statistical regularities of pixel intensity variations. When attempting recognition, the unclassified image is compared with all the database images, returning a vector of matching score. The unknown person is then classified as the one giving the highest cumulative score. This project will be build using the MATLAB software. Overall, the project can be used and developed for various purposes, particularly to expedite the process of searching the database. The refinement of this project in other hand can lead to more accurate and reliable result by considering other facial properties like eyes, nose and eyebrows.

ABSTRAK

Secara prinsipnya, untuk menggabungkan pengecaman muka dan pengkelasan jantina mungkin nampak mudah. Akan tetapi, proses ini menjadi rumit berbanding keadaan sebenar kerana perlu mempertimbangkan banyak perkara. Bidang pengecaman jantina telah menjadi satu topik yang diberikan perhatian dalam pengajian psikologi. Namun begitu, hanya sedikit pendekatan melalui teknik pengelihatan yang telah diperkenalkan. Bidang ini sebenarnya telah dipelajari secara mendalam dalam konteks keselamatan dan biometrik. Projek ini adalah berkisar tentang pengecaman jantina melalui teknik pemprosesan imej secara tertutup semata-mata yang mana menggunakan pengkalan data yang telah dimasukkan di dalam sistem. Ianya dilakukan dengan mengenalpasti perbezaan ciri-ciri di antara muka lelaki dan perempuan. Adalah terbukti bahawa pengkelasan berdasarkan satu ciri sahaja adalah tidak tepat memandangkan manusia mempunyai ciri-ciri muka yang hampir sama walaupun dari kelas jantina yang berbeza. Oleh kerana itu, pengkelasan secara berperingkat diperlukan. Projek ini berjaya sepertimana yang diharapkan berdasarkan skop dan objektif yang telah ditetapkan. Walaupun tidak banyak jenis-jenis muka yang diambil kira seperti warna rambut yang berlainan dari asal, teknik yang digunakan sepatutnya masih lagi sama. Untuk sistem pengecaman, satu teknik yang dinamakan sebagai ‘Template Matching’ digunakan untuk memadankan imej yang dikehendaki dengan imej yang berada didalam pengkalan data. Sistem ini akan mencari imej yang paling hampir dengan imej yang dikehendaki dengan menggunakan statistic yang kebiasaan bagi variasi pixel intensity. Setelah mendapatkan imej yang dikehendaki, ianya akan dikelaskan sebagai satu skor komulatif yang tertinggi. Projek ini akan dibina menggunakan perisian Matlab. Secara keseluruhannya, projek ini boleh digunakan dan dikembangkan kepada pelbagai tujuan terutamanya untuk mempercepatkan process pencarian dalam pangkalan data. Dengan sedikit pengubahsusian, projek ini semestinya akan menghasilkan satu sistem yang lebih tepat; dengan mengambil kira ciri-ciri muka manusia yang lain seperti mata, hidung dan kening.

ACKNOWLEDGEMENTS

Praise to Allah, the Most Gracious and Most Merciful, Who has created the mankind with knowledge, wisdom and power.

First of all, I am heartily thankful to my supervisor, Mr. Mohd Hisyam Bin Mohd Ariff, whose encouragement, guidance and helpful feedback during the writing of this report and project . I would like to acknowledge his comments and suggestions, which was crucial for the successful completion of this study.

Special thanks to Mrs Nurul Wahida Arshad, Ms. Faradila Naim, Mrs. Rohana Abdul Karim, Mrs. Rosyati Hamid for advice, contributing ideas, making a short course on Image Processing for a better understanding of the project and helpful cooperation during the period of this project.

My sincere thanks go to all my lab mates and members of the staff of the Electrical and Electronics Engineering Department, UMP, who helped me in many ways and made my stay at UMP pleasant and unforgettable. Many special thanks go to member engine research group for their excellent co-operation, inspirations and supports during this study.

I acknowledge my sincere indebtedness and gratitude to my parents for their love, dream and sacrifice throughout my life. I cannot find the appropriate words that could properly describe my appreciation for their devotion, support and faith in my ability to attain my goals. Lastly, I offer my regards and blessings to all of those who supported me in any respect and contributed directly or indirectly in the completion of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

ς_i	*i*th Gaussian basis function
c_i	Center
σ^2	Variance
b	Bias term
ω	Weight coefficient
$T(x,y)$	Template of an image
$S(x,y)$	Region within the image
W	Width dimension
H	Height dimension
μ_T	Mean value of the template
μ_S	Mean value of the sub image
M	Mask

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

SVM	Support Vector Machine
PCA	Principle Component Analysis
GUI	Graphical User Interface
RGB	Red Green Blue

CHAPTER 1

INTRODUCTION

1.1 Introduction to Face Recognition

The face is one of the most important biometric features of the human beings and normally used as identification. Each person has their own innate face and mostly a different face. As a human, to recognize the different faces without any difficulty is more easily but it become difficulty to the system to recognize the human faces. Face recognition is an active research area, and they can be used in wide range applications such as surveillance and security, telecommunication and digital libraries, human-computer intelligent interaction, and smart environment.

Yet, it is a challenging task to design and develop a robust computer system for face identification thus classifies the gender of the human. But, nowadays, the lack of automated face recognition systems is especially apparent when compared to our own

innate face recognition ability. Human can perform face recognition, which is an extremely complex visual task, almost instantaneously and our own recognition ability is far better and robust than any computer's can hope to be. Beside, a human also can recognize a familiar individual under very adverse lighting conditions, from varying angles or viewpoints.

In the research for determination of features, such as, gender, race, expression of a person, using facial features goes back to 1960s, it is only very recently that acceptable results have been obtained. However, face recognition system is still an area of active research since a completely successful approach or model has not been proposed to solve the face recognition problem. For the next generation surveillance systems are expected to take human frontal face as input pattern and extract useful information such as gender information, emotion, race or age information from the human face that soon it also can be used for the security system or more than that.

1.2 Problem in Gender and Face Recognition System

The problem of gender is rarely heard and less attention. However, gender identity has been used in identification cards over the years as identification. The probability of errors of gender may have while make the identification cards. Consequently, the identification card shall be remade to rectify the mistakes. This will make the process for preparing the identity card to be longer.

Thus, the existence of gender recognition based on facial image extraction may be able to help the National Registration Department to reduce or overcome these problems and speed up the processing time. This system uses the still image as an input and would classify gender based on the still image. To improve the accuracy of the output, this project will apply the template matching in face recognition and further classify the gender.

The other problem is most face recognition systems had at most a few hundred faces. This becomes a problem when the number of databases increased in a system. Larger database means longer computational and processing time. Therefore, the identification of gender can help to better face recognition system by focusing more on identity-related features, and limit the number of entries to be searched in a large database, thus improving the search speed. In other words, estimation will be done on the input image and recognition of image is done only in the estimation group. Theoretically, it will cut the processing time almost to half.

1.3 Introduction to Gender Recognition

This project is about gender recognition for classification based on the frontal facial images. The classification using the frontal still facial image is not easy. It is difficult mostly because of the inherent variability of the image formation process in terms of image quality, images size, lighting condition, the angle of image and photometry, geometry, and/or occlusion, change, and disguise. A successful gender classification method has many potential applications such as human identification, security system, smart human computer interface, computer vision approaches for monitoring people, passive demographic data collection, etc.

The interest on gender recognition are discussed about the techniques that appearance based methods, that is will learn the differential boundary between gender for male and female classes from example images, without extracting any geometrical features such as distances, face shape, face genetics and etc. Hopefully, in the 21st century now, there is a new system that can implement or develop for solutions to similar face tasks as well.

1.4 Objective

The objective of this project is to;

i. Write a MATLAB code that can recognize the gender of a person.
ii. Compare the extracted feature with facial image in database
iii. Declare classifier a gender matches of the same facial feature
iv. Develop a system that can extract recognize a gender

1.5 Scope of Project

There are several scopes that need to be proposed for the project:

i. Gender classification of a person
ii. Based on only a frontal view image
iii. Using offline test image
iv. Only single face on the image
v. No restriction on wears
vi. Transvestite is not considered in this project

1.6 Project Outline

This project is divided into five chapters. The outlines of the project are as follows;

Chapter 1 - Introduction

This chapter discusses the objectives and scope of the project which focuses on the identification and information of the facial recognition and gender estimation technology.

Chapter 2 - Literature Review

This chapter deals with facial detection of the previous work, facial feature extraction and gender recognition. There are several techniques that associated with the project will be reviewed briefly. The differences between male and female facial features will be exposed and described in this chapter. Finally, some of important part of image processing techniques will be explained in more detail and discussed.

Chapter 3- Methodology

This chapter covers in detail on the techniques used and steps taken to complete the task and finish this project. A few algorithms are proposed to be applied in this project. The flow chart of the project also will be presented and discussed in this chapter.

Chapter 4- Results

The final result of this project are shown and discussed in this chapter. Some analysis of the results and each algorithm applied are also included.

Chapter 5-Conclussion

This chapter contains the conclusion of the project. It also reflects the problems that arise and suggestion solutions for future improvement and works.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

There are several methods proposed for gender recognition. They do not consist of image processing techniques only but also applying other type of approach. All the methods to be used may help to identify the gender of the human. This chapter will be described briefly of the proposed method used.

2.2 Gender Recognition

By considering frontal facial as an image pattern, it is become a challenge to detect faces and segment into its specific features, although faces can be very different but it is still have a few same basic structure and content. Gender classification based on facial images is difficult mostly because of the inherent variability of the image formation process in terms of image quality and facial features itself. There are a few attempts have been made to perform gender classification in the earliest 1990s. The most common way of doing this is through these 2 schemes [1]:

a. Template based approach

For any template based approach, it is very much necessary to obtain a template which is a good representative of the data [2]. Template-based approach is more promising due its ease of implementation and robustness. In holistic template-matching systems, attempts are made to capture the most appropriate representation of face images as a whole and exploit the statistical regularities of pixel intensity variations [3].This simple face recognition technique is based on the whole image gray-level templates. The most direct of the matching procedures is the correlation technique. When attempting recognition, the unclassified image is compared with all the database images, returning a vector of matching score. The unknown person is then classified as the one giving the highest cumulative score [3].

b. Geometrical local feature based approach

The geometrical-based approach performs successfully in accurate facial feature detection scheme. However, it's become unreliability in some cases. The procedure is that utilize various properties of a face (such as facial topology, hair, etc) as the code for face. The idea is to extract relative position and other parameters of distinctive features such as eyes, mouth, nose and chin. [1,4]

2.3 Proposed Processing Techniques

By making use the above approaches various processing techniques have been proposed in previous work. All the methods have their own advantages to give the accuracy in term of some part in this project. Some well-known method will be described briefly.

a. Support Vector Machine (SVM)

A Support Vector Machine is a learning algorithm for pattern classification, regression and density estimation. It is provides a novel means of classification using the principles of structure risk minimization. [5] SVM is primarily designed

for binary classification problems. The popular methods are that multiclass classifications problems are decomposed into many class problems and these binary-class SVMs are incorporated in a certain way. An interesting property of SVM is an approximate implementation of the structure risk minimization induction principle that aims at minimizing a bound on the generation error of the model, rather than minimizing the mean square error over the data set. [5] The basic training principle behind SVM's is finding the optimal linear hyper plane such that the expected classification error for unseen test samples is minimized –i.e., good generalization performance. [1]

b. Radial Basis Function Networks

A Radial Basis Function (RBF) network is also a kernel-based technique for improved generalization, but it is based instead on regularization theory. A typical RBF network with K Gaussian basis functions is given by Eq.(2.1)

$$f(x) = \sum_{i}^{K} \omega_i \zeta(x; c_i, \sigma_i^2) + b \qquad (2.1)$$

Where the ζ_i is the *i*th Gaussian basis function with center c_i and variance σ_i^2 and b is a bias term. The weight coefficient ω_i combine the basis functions into single scalar output value, with b as a bias term.[1]

c. Principle Component Analysis (PCA)

The main idea in this method is getting the features of the face in mathematical sense instead of the physical face feature by using mathematical transforms [6]. It's also involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal

components [6]. In brief, this algorithm will find the principle components of the covariance matrix of a set of the face images. PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences of the face images. Each face image in the database set can be represented exactly in terms of a linear combination of the eigenfaces [7]. This technique allows the system to represent the necessary information for comparing the faces using the little information once the mathematical representation accomplished which it is need to have a lot of faces to be store. On the other hand, it suffers a bit from the fact that facial image have to be normalized that meaning they all have to be the same size and the eyes, nose, and mouth in the sample images must be lined up before the PCA applied [7].

2.4 Physical Differences between Genders

From the medical research and human anatomy it is known that males and females have some different, distinctions and advantages in their physical appearance. However, there also exist some overlapping features of the males and females that sometimes are difficult to recognize.

Table 2.1 show summarizes the basic differences between male and female faces but of course the degree of masculinity or femininity varies from person to person. It is important to remember that no single feature makes a face male or female, it is the number of masculine or feminine features that counts [8, 9].