# Face Emotion Recognition Using Artificial Intelligence Techniques

By

**Karthigayan A/L Muthukaruppan**
**(0540610045)**

A thesis submitted
in fulfillment of the requirements for the degree of
Doctor of Philosophy (Mechatronics Engineering)

**School of Mechatronics Engineering**
**UNIVERSITI MALAYSIA PERLIS**

2008

**GRADUATE SCHOOL**

**UNIVERSITI MALAYSIA PERLIS**

**PERMISSION TO USE**

In presenting this thesis in fulfillment of a post graduate degree from the Universiti Malaysia Perlis, I agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by my supervisor(s) or, in their absence, by the Dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Malaysia Perlis for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or in part should be addressed to:

**Dean of Graduate School**
**Universiti Malaysia Perlis (UniMAP)**
**Jalan Bukit Lagi**
**01000 Kangar**
**Perlis, Malaysia**

## ACKNOWLEDGMENT

The author extends his gratefulness to the supervisors and would like to express his boundless appreciation and warmest gratitude to Assoc. Prof. Dr. Mohd Rizon, Prof. Dr. Sazali Yaacob and Prof. Dr. R. Nagarajan for their valuable supervision, continuous encouragement, inspiring suggestion, guidance in the research and in the preparation of this thesis. They provided the author with great opportunity and allowed him to go in depth in the areas of image processing, genetic algorithm, fuzzy clustering and neural network. Their motivation, advice, encouragement and many discussions have helped the author for the completion of the research and this thesis.

The author would like to express his gratitude to the Vice Chancellor of Universiti Malaysia Perlis, Kol. Prof. Dato Dr. Kamarudin Hussin for his permission to carryout this research. The author also thanks the University for extending a financial support through a Graduate Assistantship.

The author would like to thank to Haniza Yazid and Siti Baizura (South East Asian subjects) who allowed me to collect their personalized expression gallery.

The author would also like to express his sincere thanks to his colleagues for their friendly cooperation.

Finally, the author is also grateful to both of his parents Mr. SM. Muthukaruppan and Mdm. M. Sornavalli, his siblings Dr. M. Somasundaram & Mrs. M. Vasanthi and also his in-laws Mr. T. Sevarcodiyan & Dr. M. Uma for their love, continuous support, patience and encouragement in completing this research work.

Thank GOD for His blessing.

## TABLE OF CONTENTS

## LIST OF TABLES

Table

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

AI – Artificial Intelligence

AU – Action Units

CCD – Charge-Coupled Device

CMOS – Complimentary Metal Oxide Semiconductor

EEG – Electroencephalography

EMG – Electromyography

ERP – Event Related Potential

et al. – and others

FCM – Fuzzy C- Mean

FER – Face Emotion Recognition

FMRI – Functional Magnetic Resonance Imaging

GA – Genetic Algorithm

GSR – Galvanic Skin Reflex

HCI – Human Computer Interaction

ISFER – Integrated System for Facial Expression Recognition

KNN – K- Nearest Neighbor

LOG – Laplacian of Gaussian

RAA – Robot-Assisted Activity

ROI – Region of Interest

SEA – South East Asian

SVM – Support Vector Machines

**PENGENALANPASTIAN EMOSI MUKA DENGAN TEKNIK TEKNIK KECERDIKAN BUATAN**

## ABSTRAK

Sejak kebelakangan ini, pembangunan yang pesat telah berlaku dalam bidang Interaksi Manusia Komputer (Human Computer Interaction: HCI). Banyak aplikasi HCI telah di documentasi dan salah satu daripada yang terkenal adalah Pengenapastian Emosi Muka (Face Emotion Recognision: FER). Tujuh emosi muka telah dikenalpasti dalam penyelidikan iaitu: gembira, sedih, marah, takut, terkejut, jijik dan neutral. FER boleh digunakan di hospital dan di rumah (warga tua, pesakit yang tidak boleh bergerak dan yang cedera parah) serta untuk menganalisa emosi psikologi kendiri. FER mempunyai pelbagai arahtuju dan kaedah membolehkannya memiliki pakej pengenapastian yang baik. Walau bagaimanapun, terdapat banyak sebab bagi kegagalan dalam pakej dan salah satu daripadanya adalah rupa wajah yang berubah dengan umur, warna, keadaan mental dan expresi muka individu.

Dalam penyelidikan ini, masalah telah ditumpukan kepada emosi muka kenderi dan kajian telah dilakukan bagi memperolehi pengenapastian yang lebih baik. FER telah dicapai dalam dua bahagian iaitu bahagian pemprosesan imej dan bahagian klasifikasi. Bahagian pertama menyelidik satu set pemprosesan imej yang sesuai bagi pengenapastian emosi muka. Imej yang diperolehi melalui beberapa kaedah pemprosesan awal. Pengesan penjuru mesti berjaya walaupun kecerahan lampu adalah tidak seimbang. Jadi, untuk mengatasi masalah ini, imej histogram dipisahkan kepada 2 rantau kecenderungan (ROI) – mata dan bibir. Kedua-dua rantau ini telah melalui pemprosesan awal yang sama tetapi dengan bendul yang berlainan. Konfigurasi mata dan bibir lebih cenderung kepada bujur. Dengan objektif untuk mencari perubahan pada bahagian mata dan bibir, satu set fungsi kecergasan bujur dicadangkan. Fungsi kecergasan ini mencari perubahan pada paksi minor pada kedua-dua imej mata dan bibir. Fungsi kecergasan ini digunakan pada algoritma genetik (GA) untuk mencari nilai optimum pada paksi minor. Tiga fungsi kecergasan direkabentuk, satu bagi mata dan dua bagi bibir (bibir atas dan bawah). Fungsi kecergasan ini digunakan pada mata dan bibir imej subjek dari Asia Tenggara, Jepun dan China. Pemerhatian kepada pelbagai emosi dari ketiga-tiga subjek telah membawa kepada satu sifat mata dan bibir yang unik. Hasil dari nilai yang telah dioptima menunjukkan nisbah paksi minor berbanding dengan emosi neutral dari Asia Tenggara, Jepun dan China. Telah ditemui, dari data yang dioptimakan, tidak terdapat corak umum bagi mengenalpasti emosi pada ketiga-tiga subjek. Ketidakhadiran corak umum menyebabkan pengajian dijalankan pada etnik khusus. Bagi memahami pengenapastian emosi muka khusus, fungsi kecergasan diaplikasikan pada dua subjek Asia Tenggara. Walau bagaimanapun, ditemui bahawa terdapat emosi yang bertindih dengan julat emosi yang lain. Bagi mengatasi masalah ini untuk mengenali emosi, dua teknik klasifikasi cerdik buatan (AI) iaitu rangkaian neural dan longgokan fuzi digunakan. Pelbagai bentuk rangkaian neural telah digunakan dan salah satu daripadanya ditemui mencapai kadar kejayaan 91.24% bagi subjek 1 dan 89.76% bagi subjek 2 dari Asia Tenggara. Bagi kes teknik klasifikasi kedua, dua bentuk longgokkan fuzi-c-purata telah dicuba dan prestasinya dibandingkan. Salah satu daripadanya menunjukan perlaksanaan yang baik dengan mencapai kadar kejayaaan 90% bagi kedua-dua subjek dari Asia Tenggara. Dengan ini disimpulkan bahawa analisa emosi khusus melalui ciri wajah dua subjek menunjukkan kadar kejayaan yang lebih tinggi dibandingkan dengan analisa umum yang dilakukan pada banyak muka dari pelbagai personaliti etnik.

## ABSTRACT

*Recently, there has been tremendous growth in the area of Human Computer Interaction (HCI). Many HCI applications were documented, and among them, the Face Emotion Recognition (FER) is one of the well known areas. Seven face emotions are considered universally in FER research: they are happy, sad, angry, fear, surprise, disgust and neutral. The FER can find applications in hospital and in home (for senior citizens, bed ridden persons and severely injured patients) and in analyzing the personal emotion psychology. The FER comes with various approaches and methods in the way to have a good recognition package. However, there are various reasons for the failures in the packages and one of them is due to face features that change with age, color, mental state and individual face expressions.*

*In this research, the problem is focused on the personalized face emotion and some studies are extended for better emotion recognition. FER is achieved in two parts, they are image processing part and classification part. The first part investigates a set of image processing methods suitable for recognizing the face emotion. The acquired images have gone through few preprocessing methods. The edge detection has to be successful even when the intensity of light is uneven. So, to overcome the difficulty of uneven lighting, the histogram equalized image is split into two regions of interest (ROI) – eye and lip regions. These two regions have been applied with the same preprocessing methods but with different threshold values. The human eye and lip configurations are found to be more of towards ellipses. With the objective of finding the changes in eye and lip areas, a set of new forms for ellipse fitness function is proposed. The fitness functions find changes in the minor axes of both eye and lip images. The fitness functions are utilized by genetic algorithm (GA) to find the optimized values of minor axes. Three fitness functions are developed, one for the eye and two for the lip (top and bottom lip). These fitness functions are applied on eye and lip images of South East Asian, Japanese and Chinese subjects. Observation of various emotions of the three subjects leads to a unique characteristic of eye and lip. Outcome of optimized values indicate the ratios of the minor axes with respect to neutral emotion for the SEA, Japanese and Chinese subjects. It is found, from the optimized data, that there is no common pattern to recognize emotions within among the three subjects. The absence of common patterns leads to studies on emotion personalized to an ethnic. In order to understand the personalized face emotion recognition, the developed fitness functions are applied on two SEA subjects. However, it is found that some emotion range overlaps with other emotion ranges. In order to circumvent this problem in recognizing the emotions, two Artificial Intelligence (AI) classification techniques such as neural network and fuzzy clustering are employed. Various forms of neural networks have been applied and one of them is found to perform well in achieving a success rate of 91.42% for SEA1 and 89.76% for SEA2. In the case of second classification technique, two forms of fuzzy c-mean clustering are considered and their performances are compared. One of them performs better by achieving a 90% success rate for both SEA1 and SEA2. It is concluded that the analysis of personalized emotion through facial features of two subjects indicate higher rate of success compared to a general form of analysis that is applied to varieties of faces of several ethnic personalities.*

# CHAPTER 1

# INTRODUCTION

## 1.1 Overview

Recently, there are many research projects which focused on human computer interaction (HCI). It is an interdisciplinary subject, relating computer science with many other fields of study and research such as pattern recognition, image processing, bio-sensors and electronics. The HCI is a discipline concerned with the design, evaluation and implementation of interactive computing system for human use and with the study of major phenomena surrounding them. The area of research can be in gesture recognition, multi-media, 3-D visual, virtual reality, computer supported and cooperative work, natural language and speech [Brad, 1998]. Since HCI relies on human machine communication, it draws supporting knowledge from both the machine and the humans. On the machine side are the techniques in computer graphics, operating systems, programming languages, and development environments. On the human side, communication theory, graphic and industrial design disciplines, linguistics, social sciences, cognitive psychology, and human performance are relevant. A long term goal of HCI is to design systems that minimize the barrier between what the human's cognitive model wants to accomplish and the computer's understanding of the user's task. Computer vision has many advantages as an input modality for multimodal or perceptual interfaces. Visual information is clearly important in HCI, as meaningful information conveyed through identity, facial expression, posture, gestures and other visually observable cues. While vision is one of the possibly several sources of

information about the interaction to be combined multimodal in perceptual interface, it focuses solely on the vision modality. Using computer vision to sense and perceive the user in an HCI context is often called vision based interaction or vision based interface [Medioni and Kang, 2004].

Emotion can be defined as any agitation or disturbance of mind, feeling, passion or any vehement or excited mental state [Goleman, 1996]. According to a researcher, only 7% of our impression is accounted for verbal messages and rest are non-verbal message such as vocal impact for 38% and facial impact for 55%. It shows that the facial impact is the highest for a good emotion expression [Brent and Lea, 2006].

Goleman (1996) has discussed about active ingredients of prevention program, social and emotional learning. In this, the author describes the emotional skills for active ingredients of prevention program. They are identifying and labeling feelings, expressing feeling, assessing the intensity of feelings, managing feelings, delaying gratification, controlling impulses, reducing stress and knowing the difference between feeling and action. The social and emotional learning results for child development project are more responsible, more assertive, more pro-social and helpful in better understanding of others, more pro-social strategies for interpersonal problem solving etc. In the section on intimate enemies, Daniel [Goleman, 1996] expressed his view on the increase on divorce rate every decade. He also discussed that the emotional intelligence can improve the relationship among intimates. It has been learnt that emotion is one of the main factors in our life. To find and interpret face motion is one of the main tasks to bring good harmony in the life. This shows the importance and

necessity of the emotion recognition package in the community. This does help a lot to the matter just mentioned above.

Face emotion recognition (FER) is one of the researches based on HCI. Since last decade, many researches and developments have been carried out on FER. The FER research and development can be achieved through machine vision and Electromyogram (EMG) [Muthukaruppan, 2004]. Both machine vision and EMG have positive and negative research aspects. Vision system has proven record in various fields of research [Chellapa et.al., 1995; Guilherme and Avinash, 2002]. Although various methods are available in the modern ways of research, vision system is still being considered as successful because of its non-invasiveness, where as EMG is derived from human muscles actions. The various methods and techniques implemented in FER using vision, EMG and electroencephalogram (EEG) will be discussed in detail in the following chapter.

**1.2 Problem Statement and Its Significance**

Although, many techniques and methods have been implemented in recognition of face emotion, new techniques are expected to bring good classification of emotion. In order to achieve a good classification, few problems are to be solved. They are acquisition of good image, processing the image for a suitable feature extraction and recognition. Though several methods have achieved the goal in recognizing an emotion, some draw backs still exist. The tools used for emotion recognition are neural network, fuzzy logic and a combination of both neural network and fuzzy logic [Pantic and Leon , 2000].

Apart from the above stated difficulties, the main problem affecting the emotion recognition is the varied ethnic face features. Each ethnic people from one ethnic feature the other with respect to the shape and size of eye, lip, colour and face. Even when the same ethnic group, some difficulty in emotion recognition still exists due to individual ways of expressing emotion. Facial expression of one's emotion differs from that of others within the same ethnic group. In this circumstance, it makes sense to understand that the personalized emotion of an ethnic is needed to be studied. The study of personalized emotion of ethnic can be carried out by observing the changes of the face features (such as both eye and lip). The most expressive or interpretation of a person's emotion features are eyes and lips. Whenever there are changes in the emotion of a person, definitely there are specific changes in the size and shape of the features. In this research, features as eyes and lips are considered in face emotion recognition.

### 1.3 Research Objectives

The objective of this research work is of two fold as given below:

i. To investigate on certain geometrical shape for the human eye feature

The human eye shape tends to have geometrical ellipse. The eye image can be considered as an ellipse with major and minor axis. The major axis of an is more or less fixed in elliptical configuration for a particular person. The image of eye acquired from a camera and processed may not have a perfect form of ellipse. The minor axis is to be obtained optimally through some optimization methods. Genetic algorithm (GA) is proposed as an optimization method. However, the GA needs relevant fitness function. A fitness equation is derived to find a minor axis such that the emotion changes can be

easily recognized. With the assistance of fitness function, the GA optimally computes minor axis, b, such that the emotions can be easily related to the values of b. It is the objective to derive fitness equation and evaluate its performance in determining the changes in minor axis data towards emotion recognition.

ii. To Investigate a form of complex elliptical shape to human lip

The human lip shape is more of towards combination of two ellipses and is referred to as irregular ellipse. The word 'irregular' means that the ellipse has two different minor axes. One minor axis is connected to top lip and the other to bottom lip. Variations in minor axis can categorize the human emotion. Here also the use of GA is proposed. Two fitness functions for this irregular ellipse are to be derived. This is the objective to develop and evaluate the performance of the fitness equations for irregular ellipse to categorize human emotions.

**1.4 Research Methodology**

The objective of this research work is of four fold as given below:

i. To compile a literature survey on face emotion recognition:

There have been quite a number of contributions to the literature discussing various aspects, methods and applications of detection emotion through changes in face features. It has been found that the set of research papers for the last ten years have demonstrated modern approaches in face emotion recognition and applications. This is an objective to organize research papers of the last 10 years to bring out important

aspects connected to human emotion recognition and categorization through face images.

ii. To analyze and determine suitable feature extraction methods for FER

Feature extraction plays one of the vital roles in the emotion recognition part. There are various types of feature extraction methods available in the literature. It is needed to pick up one method suitable to more than one variety of ethnic faces. In this study, three typical faces from ASIAN are to be considered. The performance of the various feature extraction methods need to be evaluated and to derive the one most useful in face emotion recognition

iii. To investigate the application of neural network in solving emotion range overlapping problems

Neural network is one of the good classifiers. Each emotion has its own range based on GA optimized data. Some GA optimized data range can be overlapped with other emotion ranges. This is expected because there can be small changes in emotions, so that the chances of emotion overlapping are high. In order to overcome this overlapping problem between the emotions and at the same time obtain enhanced classification, a neural network approach is to be implemented. Neural network has to be applied to solve overlapping problem among emotion ranges obtained from GA techniques. The neural network application of the optimized data acquired from genetic algorithm is to be investigated

iv. To investigate the application of fuzzy clustering approach to solve emotion range overlapping problems

Fuzzy clustering approach has been widely used in problem of pattern recognition, data classification and data clustering. This approach is also suitable in the overlapping problem of emotion range. The investigation is to evaluate the application of fuzzy clustering of emotion data from eye and lip features and to compare its performance with neural network approach.

## 1.5 Thesis Layout

This thesis is composed of seven (7) chapters:

i. Chapter 1 discusses the introduction of face emotion recognition; human computer interaction introduces the problem statement, research objectives and thesis layout.

ii. Chapter 2 deals with the literature survey of several previous works that are related to face emotion recognition.

iii. Chapter 3 presents the face image processing, morphological filtering method and feature extraction method. In image processing, there are some image transformations such as conversion to gray scale, filtering, edge detection, selection of ROI regions and suppressing the background. The feature extraction methodology is to be applied to extract features from the cropped image. The

feature extraction methods suitable for face images such as profile projection, moments and contour profile are discussed.

iv. Chapter 4 developes equations for face emotion classification of features such as eye and lip using genetic algorithm. The eye is considered to be a regular ellipse whereas the lip is envisaged as a combination of two ellipses. One fitness equation for eye and two fitness equations for lip are derived from equations of original ellipse and that of irregular ellipse respectively.

v. Chapter 5 deals with the application of elliptical configuration of eye and lip. GA is used for obtaining the optimum parameters of the eye and lip configurations. The developed fitness functions for eye and lip are to be applied to the faces of three ethnic groups, namely South East Asian, Japanese and Chinese.

vi. Chapter 6 brings out with personalized face emotion classification for two South East Asian subjects. Both eye and lip fitness functions are used. It is found that the optimized data of one emotion range overlapped with the other emotion ranges. To overcome this problem, some artificial intelligence techniques are applied. Neural network and fuzzy-C mean clustering are discussed and applied to the optimized data. The performances of these AI methodologies are summarized for both South East Asian subjects.

vii. Chapter 7 deals with the contribution and conclusion of the research with an indication of a set of possible future research work.

# CHAPTER 2

## A SURVEY ON FACE EMOTION RECOGNITION AND APPLICATIONS

### 2.1 Introduction

In the twentieth century, the world has come across with many technological innovations. Among the developing technologies the machine vision (vision sensors) and bio sensors play important roles in several areas of applications from industry to medical. While vision sensors are mostly suitable in detecting and categorizing the changes on the surface of human body, the bio-sensors measure and classify the activities from inside.

In recent years, there has been a growing interest in improving all aspects of interaction between human and computers especially in the area of human emotion recognition by observing facial expression. Such an activity of human emotion conveyed through facial expressions is referred, in this chapter, as 'facial emotion'. Human being possesses an ability of communication through facial emotion in day to day interactions with others. There are several categories of facial emotions; some of them are easy to detect and others are difficult to recognize. In addition, facial changes due to an emotion cannot be predicted. That is, changes of facial expressions vary from time to time even for the same emotion. Facial expressions also vary from person to person even for the same emotion. Hence, detecting and classifying the emotions through human facial changes are very challenging. Nevertheless, several researchers