



Investigation of Nonlinear Feature Extraction Techniques for Facial Emotion Recognition

by

**Hasimah Ali
(1140610629)**

A thesis submitted in fulfillment of the requirements for the degree of
Doctor of Philosophy

**School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS**

2016

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : HASIMAH ALI
Date of birth : 03 DECEMBER 1980
Title : INVESTIGATION OF NONLINEAR FEATURE EXTRACTION
TECHNIQUES FOR FACIAL EMOTION RECOGNITION
Academic Session : 2015/2016

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

- CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*
- RESTRICTED** (Contains restricted information as specified by the organization where research was done)*
- OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only. (except during a period of _____ years, if so requested above).

Certified by:

SIGNATURE

SIGNATURE OF SUPERVISOR

(NEW IC NO. / PASSPORT NO.)

NAME OF SUPERVISOR

Date : _____

Date : _____

NOTES : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.

ACKNOWLEDGMENT

First and foremost, I thank God the Almighty, Allah Subhana Wata'ala, for giving His blessing to me to complete this fruitful journey of doctoral study. This research is never solitary and thus, I have many people to thank for making the writing of this thesis bearable and to reach the end of my battle successfully.

I should thank the Ministry of Higher Education Malaysia and my employer, Universiti Malaysia Perlis (UniMAP) for giving this golden opportunity and financial support, without which I may have not pursued this study diligently. Also, I would like to acknowledge the Vice Chancellor of University Malaysia Perlis (UniMAP), Brig. Jeneral Dato' Prof. Dr. Kamaruddin Hussin, the Dean and staffs of Mechatronic Engineering of UniMAP and the Dean and staffs of Centre for Graduate Studies of UniMAP for their support and encouragement throughout my candidature.

Especially, I would like to thank my main supervisor, Dr. Hariharan Muthusamy for supervising my doctorate study. He has been very patient and dedicated with the writing and eventual completion of this thesis. His inputs and comments were tremendous and constructive and his insight proved to be very valuable.

My deepest gratitude also goes out to Prof. Dr. Sazali Yaacob, my co-supervisor, especially during the first year of my candidature, who graciously set aside time to entertain my fearful emotion of the dark ages and guide me through the literature review of this study. I deeply appreciate his genuine concern for my well-being, his true belief in me, his steady support and his kind words of encouragement.

I wish to also thank my second co-supervisor, Prof. Dr. Hamid Adom for his advice and support. Also, many thanks go to fellow members of the Intelligent Signal Processing Cluster of UniMAP, especially to my kind friends, Hazadura, Nadia, who, closely share their knowledge and guide me very well through real examples.

It is no denying that I went through a very stressful time during my study. During these times of enormous difficulty, my family was the most that I sacrificed. My sincerest and utmost gratitude go to my beloved husband, Mohamed for his kindness, allegiance and unconditional love and support by taking finest care of our children, Muriyah and Fatimah and for helping me with the household chores. I am also grateful to my beloved mother, for her patience, love, understanding and endless prayers towards my success.

©This item is protected by original copyright

TABLE OF CONTENTS

ACKNOWLEDGMENT	II
TABLE OF CONTENTS	IV
LIST OF TABLES	VIII
LIST OF FIGURES	X
LIST OF ABBREVIATIONS	XIV
LIST OF SYMBOLS	XVI
ABSTRAK	XVIII
ABSTRACT	XIX
CHAPTER 1 INTRODUCTION	1
1.1 Research Background	1
1.2 Problem Statements	5
1.3 Research Objectives	6
1.4 Research Scope	7
1.5 Thesis Outline	8
CHAPTER 2 LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Facial Expression Recognition Analysis	11
2.3 Face Pre-Processing	12
2.3.1 Face Detection	12
2.3.2 Face Normalization	14
2.4 Facial Feature Extraction Method	14
2.4.1 Feature Type	15
2.4.2 Geometry Based-Feature	17
2.4.3 Appearance-Based Approach (Texture)	18

2.5	Data Dimensionality Reduction Technique (Feature Reduction)	27
2.6	Facial Emotion Classification	29
2.7	Motivation of the Present Studies.	31
2.7.1	HOS	34
2.7.2	EMD	39
2.8	Summary	42
CHAPTER 3 RESEARCH METHODOLOGY		44
3.1	Description of the Proposed Method	44
3.2	Facial Expression Database	46
3.2.1	JAFFE Database	46
3.2.2	CK database	47
3.3	Image Pre-processing	48
3.4	Radon Transform	49
3.5	Feature Extraction: Application of Nonlinear Technique	55
3.5.1	HOS-Based Feature Extraction	55
3.5.2	Applied HOS on Facial Images	56
3.5.3	Applied EMD on Facial Images	59
3.6	Dimensionality Reduction Techniques	60
3.6.1	Principal Component Analysis (PCA)	61
3.6.2	Linear Discriminant Analysis (LDA)	62
3.6.3	PCA + LDA	63
3.6.4	LFDA	64
3.6.5	Kernel LFDA (KLFDA)	66
3.7	Classification	67
3.7.1	K-Nearest Neighbor (K-NN)	67
3.7.2	Support Vector Machines (SVM)	67
3.7.3	Extreme Learning Machines (ELM)	71
3.8	Summary	74

**CHAPTER 4 PERFORMANCE ANALYSIS OF FACIAL EMOTION
RECOGNITION USING HIGHER ORDER SPECTRA 76**

4.1	Introduction	76
4.2	Bispectrum Plot	76
4.3	Dimensionality Reduction on HOS Features	84
4.3.1	Distribution of HOS Features using LDA	84
4.3.2	Distribution of HOS Features Based on LFDA	86
4.3.3	Distribution of HOS Features Based on KLFDA	88
4.4	Statistical Analysis	89
4.5	Experimental Results and Discussion	90
4.5.1	Experiments on the JAFFE database	90
4.5.2	Experiments on the CK database	97
4.5.3	Experiments on Partial Occlusion Images	100
4.5.4	Experiments on Robustness to Noise	103
4.6	Summary	106

**CHAPTER 5 PERFORMANCE ANALYSIS OF FACIAL EMOTION
RECOGNITION USING EMPIRICAL MODE DECOMPOSITION 107**

5.1	Introduction	107
5.2	Analysis of Applied EMD on Facial Images	108
5.3	Dimensionality Reduction on IMF features	112
5.3.1	Distribution of IMF1 features due to PCA + LDA	113
5.3.2	Distribution of IMF1 features due to PCA + LFDA	115
5.3.3	Distribution of IMF1 features due to KLFDA	117
5.4	Statistical Analysis: ANOVA Test	118
5.5	Results and Discussion	119
5.5.1	Experiments on JAFFE database	119
5.5.2	Experiments on CK database	126
5.5.3	Experiments on Noisy Images	133
5.5.4	Experiments on Partial Occlusion Images	137
5.6	Performance Comparison with Previous Work	140
5.7	Summary	141

CHAPTER 6 CONCLUSIONS AND FUTURE WORK	144
6.1 Main Contributions	144
6.2 Recommendation for Future Work	147
REFERENCES	148
APPENDIX A	160
APPENDIX B	161
LIST OF PUBLICATIONS	176

©This item is protected by original copyright

LIST OF TABLES

NO.		PAGE
2.1	Feature extraction methods reported in the literature	16
4.1	Recognition rates of raw HOS features using different classifiers on JAFFE dataset	91
4.2	Confusion matrix of raw features using SVM on JAFFE dataset	92
4.3	Recognition rates based on HOS + PCA features using different classifiers on JAFFE dataset	93
4.4	Confusion matrix of HOS and PCA using SVM on JAFFE dataset	93
4.5	Recognition rates of HOS + LDA by using different classifiers on JAFFE dataset	94
4.6	Recognition rates of HOS + LFDA by using different classifiers on JAFFE dataset	95
4.7	Recognition rates of HOS + KLFDA by using different classifiers on JAFFE dataset	96
4.8	Recognition rates of HOS + LDA by using different classifiers in CK database	97
4.9	Recognition rates of HOS + LFDA by using different classifiers on CK database	98
4.10	Recognition rates of HOS + KLFDA by using different classifiers on CK database	99
5.1	<i>F</i> -value of six components PCA + LDA of extracted IMF1 feature (<i>p</i> -value < 0.05)	119
5.2	Recognition rates of IMF1 + PCA + LDA by using different classifiers on JAFFE dataset	120
5.3	<i>F</i> -value of six components PCA + LFDA of extracted IMF1 feature (<i>p</i> -value < 0.05)	122
5.4	Recognition rates of IMF1 + PCA + LFDA by using different classifiers on JAFFE dataset	122
5.5	<i>F</i> -value of six components KLFDA of extracted IMF1 features (<i>p</i> -value < 0.05)	125

5.6	Recognition rates of IMF1 + KLFDA by using different classifiers on JAFFE dataset	125
5.7	<i>F</i> -value of six components PCA + LDA of extracted IMF1 features (<i>p</i> -value < 0.05) on CK database	126
5.8	<i>F</i> -value of six components PCA + LFDA of extracted IMF1 features (<i>p</i> -value < 0.05) on CK database	127
5.9	<i>F</i> -value of six components KLFDA of extracted IMF1 features (<i>p</i> -value < 0.05) on CK database	127
5.10	Recognition rates of IMF1 + PCA + LDA by using different classifiers on CK database	131
5.11	Recognition rates of IMF1 + PCA + LFDA by using different classifiers on CK database	131
5.12	Recognition rates of IMF1 + KLFDA by using different classifiers on CK database	132
5.13	Confusion matrix of IMF1 + KLFDA features using ELM-RBF classifier on CK database	132
5.14	Performances of recognition rate for different variances and classifiers using IMF2 + PCA + LDA on CK database	136
5.15	Recognition rates of IMF2 + PCA + LFDA by using different classifiers on CK database	136
5.16	Recognition rates of IMF2 + KLFDA by using different classifiers on CK database	137
5.17	Performance comparison with previous work using JAFFE database	140
5.18	Performance comparison with previous work using CK database	141

LIST OF FIGURES

NO.		PAGE
2.1	Generic facial expression recognition frameworks	11
2.2	Example of face detection using hybrid approach	13
2.3	Feature classification	15
2.4	Geometric features with 34 fiducial points	17
2.5	Representation of Gabor wavelet using three different kernels	20
2.6	The basic LBP operator	22
2.7	A face image is divided into small regions which LBP histograms are extracted and concatenated into a single vector	23
2.8	Two-dimensional discrete wavelet transform at first level decomposition	26
2.9	The symmetry regions of bispectrum (denoted by 1). The region denoted 1 contains a unique set of values and those in the other denoted regions can be mapped to these	39
2.10	Step 1, 2, 3 and 4 of EMD algorithm	41
2.11	Example of EMD computation and its result of original signal	42
3.1	Block diagram of the proposed method	45
3.2	Example of seven expressions of emotion in JAFFE database	47
3.3	Example of seven expressions performed by subject in CK database	48
3.4	Example of pre-processing (a) Face model and (b) Cropped images of the subject in JAFFE database	49
3.5	The source and the sensor compartment are rotated about the centre of the object	50
3.6	Illustration of the line in the image at $\theta = 19^\circ$ and the result of Radon transform with $\theta = 19^\circ$	51
3.7	The Radon transform of the facial image $f(x, y)$ results in 1D parallel beam projection, $g(s)$ at specified angle θ	53
3.8	Radon transformation corresponding to seven categories (a) to (g) of facial emotions, calculated at 70 degrees	54

3.9	The principle region or the non-redundant region (Δ) of computation of the bispectrum	56
3.10	Bispectrum magnitude plot in $f_1 - f_2$ plane with principal domain indicated	58
3.11	The resultant of IMF1, IMF2, IMF3 and residue signal derived from EMD technique applied on 1D facial signal of emotion surprise	60
3.12	SVM with linear separable case	68
3.13	The structure of ELM network	74
4.1	1D facial signal of facial images and its bispectrum plot of: (a) anger, (b) disgust, (c) fear, (d) happiness, (e) neutral, (f) sadness and (g) surprise	77
4.2	Scatter plot of feature distributions of HOS features due to LDA on, (a) JAFFE dataset and (b) CK dataset	85
4.3	Scatter plot of feature distribution of HOS features due to LFDA on (a) JAFFE dataset and (b) CK dataset	87
4.4	Scatter plot of feature distribution of HOS features due to KLFDA on (a) JAFFE dataset and (b) CK dataset	88
4.5	Recognition rates of seven facial emotions based on HOS + KLFDA features using ELM-RBF classifier	99
4.6	Illustration of the image in different occlusion modes. Image (a) to (d) are the faces in four occlusion modes ; (a) upper face, (b) lower face, (c) right face and, (d) left face	100
4.7	Performance of the HOS + LDA features under different occlusion modes	101
4.8	Performance of the HOS + LFDA features under different occlusion modes	101
4.9	Performance of the HOS + KLFDA features under different occlusion modes	102
4.10	Illustration of neutral and sad expressions with their corresponding occlusion	102
4.11	Example of noise corrupted images: (a) the original image, (b) corrupted image by variance $\nu = 0.01$, (c) corrupted image by variance $\nu = 0.02$ and (d) corrupted image by variance $\nu = 0.03$	103

4.12	Performance of the HOS + LDA features under different variances	104
4.13	Performance of the HOS + LFDA features under different variances	105
4.14	Performance of the HOS + KLFDA features under different variances	105
5.1	Signal of emotion surprise, IMF1 to IMF3 and residue of a JAFFE subject	109
5.2	Signal of emotion sadness, IMF1 to IMF3 and residue of a JAFFE subject	109
5.3	Signal of neutral, IMF1 to IMF3 and residue of a JAFFE subject	110
5.4	Signal of emotion happiness, IMF1 to IMF3 and residue of a JAFFE subject	110
5.5	Signal of emotion fear, IMF1 to IMF3 and residue of a JAFFE subject	111
5.6	Signal of emotion disgust, IMF1 to IMF3 and residue of a JAFFE subject	111
5.7	Signal of emotion anger, IMF1 to IMF3 and residue of a JAFFE subject	112
5.8	Scatter plot of feature distributions of IMF1 features due to PCA + LDA on (a) JAFFE dataset and (b) CK dataset	114
5.9	Scatter plot of feature distributions of IMF1 features due to PCA + LFDA on (a) JAFFE dataset and (b) CK dataset	116
5.10	Scatter plot of feature distributions of IMF1 features due to KLFDA on (a) JAFFE dataset and (b) CK dataset	117
5.11	The box plot of six features IMF1 + PCA + LDA features for seven distinct emotions using the Kruskal-Wallis test	121
5.12	The box plot of six IMF1 + PCA + LFDA features for seven distinct emotions using the Kruskal-Wallis test	123
5.13	The box plot of six IMF1 + KLFDA features for seven distinct emotions using the Kruskal-Wallis test	124
5.14	The box plot of six IMF1 + PCA + LDA features for seven emotions using independent-samples Kruskal-Wallis test on CK database	128
5.15	The box plot of six IMF1 + PCA + LFDA features for seven emotions using independent-samples Kruskal-Wallis test on CK database	129
5.16	The box plot of six IMF1 + KLFDA features for seven emotions using independent-samples Kruskal-Wallis test on CK database	130

5.17	Example of noisy images: (a) the free image, (b) noisy image by variance $\nu = 0.01$, (c) noisy image by variance $\nu = 0.02$ and (d) noisy image by variance $\nu = 0.03$	133
5.18(a)	Noise free image and IMFs and residue	134
5.18(b)	Noise-corrupted image by Gaussian white noise with zero mean and variance $\nu = 0.01$ and its corresponding IMF and residue	134
5.19	Performance of the IMF1 + LDA features under different occlusion modes	138
5.20	Performance of the IMF1 + LFDA features under different occlusion modes	139
5.21	Performance of the IMF1 + KLFDA features under different occlusion modes	139

©This item is protected by original copyright

LIST OF ABBREVIATIONS

A	Sum of Amplitude
ANOVA	Analysis of Variance
CK	Cohn-Kanade
D	Diagonal
DWT	Discrete Wavelet Transform
ELM	Extreme Learning Machines
EMD	Empirical Mode Decomposition
ERM	Empirical Risk Minimization
FACS	Facial Action Coding System
FLD	Fisher's Linear Discriminant
H	Horizontal
HCI	Human-Computer Interaction
HLAC	High Order Local Autocorrelation Coefficient
HOS	Higher Order Spectra
IMFs	Intrinsic Mode Functions
JAFFE	Japanese Female Facial Expression
KLFDA	Kernel Local Fisher Discriminant Analysis
k-NN	k-Nearest Neighbour
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LFDA	Local Fisher Discriminant Analysis
LGBP	Local Gabor Binary Patterns
LPP	Locality-Preserving Projection

MANFIS	Multiple Adaptive Neuro-Fuzzy Inference System
M_{ave}	Mean Magnitude
MFNs	Multilayer Feedforward Networks
NBCE	Normalized Bispectral Cubic Entropy
NBE	Normalized Bispectral Entropy
NBSE	Normalized Bispectral Squared Entropy
PCA	Principal Component Analysis
RBF	Radial Basis Function
SLFNs	Single Hidden Layer Networks
SRM	Structural Risk Minimization
SVM	Support Vector Machines
V	Vertical

©This item is protected by original copyright

LIST OF SYMBOLS

$f(x, y)$	2D image
$\psi(k, x)$	2D Gabor function
σ	Sigma
c_1^x	1 st order cumulant
$c_2^x(\tau_1)$	2 nd order cumulant
$c_3^x(\tau_1, \tau_2)$	3 rd order cumulant
c_n^x	n th order cumulant
m_1^x	1 st order moment
$m_2^x(\tau_1)$	2 nd order moment
$m_3^x(\tau_1, \tau_2)$	3 rd order moment
m_n^x	n th order moment
γ_2	Variance
γ_3	Skewness
γ_4	Kurtosis
$B(f_1, f_2)$	Bispectrum
$E[.]$	Expectation operator
$g(s, \theta)$	Radon transform
s	Small distance to the origin
θ	Angle
$E_u(t)$	Upper envelope

$E_l(t)$	Lower envelope
μ	Mean
S	Covariance
λ	Eigenvalue
v	Eigenvector
S_W	Within-class scatter
S_B	Between-class scatter
\tilde{S}_W	Local within-class scatter
\tilde{S}_B	Local between-class scatter
$A_{i,j}$	Similarity matrix
$\tilde{A}_{i,j}^W$	Weight matrices of the local within-class adjacency graph
$\tilde{A}_{i,j}^B$	Weight matrices of the local between-class adjacency graph
$K_{i,j}$	Kernel matrix
$\langle \cdot, \cdot \rangle$	Inner product
Δ	Non-redundant region
ξ	Slack variable
α_i	Lagrangian multiplier
H	Hidden layer output matrix
H^\dagger	Moore-Penrose generalized inverse of H

Kajian Terhadap Teknik Pengekstrakan Tak Linear bagi Pengenalan Emosi

Wajah

ABSTRAK

Sejak beberapa dekad yang lalu, pengenalan emosi wajah telah menerima minat yang ketara di kalangan penyelidik dalam bidang penglihatan komputer, pengenalan corak dan bidang yang berkaitan. Peningkatan aplikasi terhadap pengenalan emosi muka telah menunjukkan kesan yang besar dalam beberapa bidang termasuklah daripada psikologi kepada interaksi manusia-komputer (HCI). Walaupun pengenalan emosi wajah telah mencapai tahap kejayaan tertentu, namun prestasinya masih jauh dari persepsi manusia. Banyak pendekatan telah dicadangkan dalam kesusasteraan. Malah, keupayaan pengenal emosi muka untuk beroperasi secara automatik sepenuhnya dengan ketepatan yang tinggi masih mencabar disebabkan masalah-masalah seperti perbezaan antara kelas, persamaan antara kelas dan perubahan halus ciri-ciri wajah. Masalah ini turut dikhuatiri apabila *physiognomies* muka berkaitan dengan umur, etnik dan jantina berbeza dari individu yang lain, seterusnya meningkatkan kesukaran mengiktiraf emosi muka. Untuk menyelesaikan masalah ini, tesis ini bertujuan untuk membangunkan teknik pengekstrakan tak linear dengan menggunakan Perintah spektral Tinggi (HOS) dan Empirikal Mod Penguraian (EMD) secara berasingan dalam mengenal tujuh emosi muka (marah, jijik, takut, gembira, neutral, sedih dan terkejut) berdasarkan imej- imej statik. Langkah awal pra-pemprosesan adalah untuk mengasingkan kawasan muka daripada imej wajah asal dengan menggunakan pengecaman muka. Imej wajah 2-D kemudian diunjurkan ke dalam 1-D isyarat muka dengan unjuran berturut-turut melalui pengubah Radon. Pengubah Radon adalah translasi dan putaran tak berubah, oleh itu ia mengekalkan variasi dalam keamatan piksel. Isyarat muka yang menggambarkan emosi diekstrak menggunakan HOS dan EMD untuk mendapatkan satu set ciri-ciri yang ketara. Dalam rangka kerja HOS, statistik tertib ketiga atau *bispectrum* yang menangkap kontur (bentuk) dan maklumat tekstur telah digunakan pada isyarat muka. Dalam kajian ini, satu set ciri-ciri bispectral digunakan untuk menghuraikan ciri-ciri ketujuh-tujuh kelas emosi. Sementara itu, dalam rangka kerja EMD, isyarat muka telah diurai menggunakan EMD untuk menghasilkan satu set kecil fungsi mod intrinsik (IMFs) melalui proses saringan. Ciri-ciri IMF yang mempamerkan corak yang unik telah digunakan untuk membezakan emosi-emosi wajah. Dalam usaha untuk mengurangkan dimensi tinggi ciri-ciri HOS dan EMD, tiga teknik pengurangan dimensi telah digunakan: Analisis Pembezaan Linear (LDA), Analisis Pembezaan Fisher Tempatan (LFDA) dan *Kernel* LFDA (KLFDA). Ciri-ciri yang diperolehi kemudian dimasukkan kepada pengelas mesin pembelajaran yang berbeza seperti k-jiran terdekat (k-NN), Mesin Sokongan Vektor (SVM) dan Extreme Mesin Pembelajaran (ELM) untuk mengklasifikasikan tujuh emosi wajah. Untuk menilai keberkesanan kaedah yang dicadangkan, dua pangkalan data penanda aras telah digunakan iaitu Ekspresi Muka Perempuan Jepun (Jaffe) dan Ekspresi Muka Cohn-Kanade. Keputusan eksperimen menunjukkan bahawa kaedah yang dicadangkan bukan sahaja menunjukkan keputusan yang lebih hebat berbanding dengan beberapa algoritma yang sedia ada tetapi juga dapat menangani imej separa tertutup serta imej bising.

INVESTIGATION OF NONLINEAR FEATURE EXTRACTION TECHNIQUES FOR FACIAL EMOTION RECOGNITION

ABSTRACT

Over the last decades, facial emotion recognition has received a significant interest among researchers in areas of computer vision, pattern recognition and its related field. The increasing applications of facial emotion recognition have shown a sizeable impact in many areas ranging from psychology to human-computer interaction (HCI). Although facial emotion recognition has achieved a certain level of success, however its performance is far from human perception. Many approaches have been constantly proposed in the literature. In fact, the ability of facial emotion recognition to operate in fully automated with high accuracy remains challenging due to various problems such as intra-class variations, inter-class similarities and subtle changes of facial features. The adhered problem is further hampered as physiognomies of faces with respect to age, ethnicity and gender, thus increase the difficulties of recognizing the facial emotion. In order to resolve this problem, this thesis aims to develop nonlinear features extraction techniques of using Higher Order Spectra (HOS) and Empirical Mode Decomposition (EMD) separately in recognizing the seven facial emotions (anger, disgust, fear, happiness, neutral, sadness and surprise) from static images. A pre-processing step of isolating face region from different background was first employed by means of face detection. The 2-D facial image was then projected into 1-D facial signal by successive projection via Radon transform. Radon transform is translation and rotation invariant, hence preserves the variations in pixel intensities. The facial signal that describes the expression was extracted using HOS and EMD to obtain a set of significant features. In HOS framework, the third order statistic or *bispectrum* that captures contour (shape) and texture information was applied on facial signal. In this work, a new set of bispectral features was used to characterize the distinctive features of seven classes of emotion. While, in EMD framework, the facial signal was decomposed using EMD to produce a small set of intrinsic mode functions (IMFs) via sifting process. The IMF features which exhibit the unique pattern were used to differentiate the facial emotions. In order to reduce high dimensionality of HOS- and EMD features, three dimensionality reduction techniques were adopted: Linear Discriminant Analysis (LDA), Local Fisher Discriminant Analysis (LFDA) and Kernel LFDA (KLFDA). The obtained features were then fed to different machine learning classifiers such as k-nearest neighbor (k-NN), Support Vector Machines (SVM) and Extreme Learning Machines (ELM-RBF) for classifying the seven facial emotions. To evaluate the effectiveness of the proposed method, two benchmark databases are used namely, Japanese Female Facial Expression (JAFFE) and Cohn-Kanade Facial Expression Database (CK). Experimental results show that the recognition rate of HOS + KLFDA + ELM-RBF and IMF1 + KLFDA + ELM-RBF have achieved the accuracy of 99.26% and 99.75%, respectively. Therefore, the proposed method not only demonstrates the superior results compared with some existing algorithms but also satisfactorily deal with partially occluded images as well as noisy images.

CHAPTER 1

INTRODUCTION

1.1 Research Background

Facial emotion or synonymously facial expression refers to the explicit transformation of human face due to the automatic response of the emotional states. Facial expressions in fact are induced by the activations of facial muscles, which result in temporally deformed permanent facial features (eyebrows, eyelids, mouth and nostril) and transient facial features (frown and bulges). Temporal dynamics of muscular activities are typically brief, ranging from 5 seconds or less 250 milliseconds (Fasel & Luetttin, 2003). A study conducted by psychologist (Mehrabian, 1968) on communications pertaining to feelings and attitudes shows that 7% of meaning resides in the spoken word, 38% in vocal utterance (the way that the word are said) and the other 55% in facial expression. This implies that facial expression forms a major part of human communication. Hence, facial expression plays an important role in human face-to-face interactions in delivering messages or intentions.

© This item is protected by original copyright
Cross-cultural research in facial expression (Ekman, 1972, 1992; Ekman & Oster, 1979; Keltner & Ekman, 2000) has shown that the six facial expression of emotions such as *anger*, *disgust*, *fear*, *happiness*, *sadness* and *surprise* are universal across human ethnicities and cultures. Ekman & Friesen (1978) have introduced Facial Action Coding System (FACS) as a tool for measuring and describing facial muscular activity.

The FACS refers to the description of an individual or combine facial muscles and tongue generated from analysis of facial anatomy. They measure the facial appearance changes in terms of FACS with 44 different action units (muscle actions) that produce them. Therefore, FACS provides a tool for behavioral science research, cognitive process and also becomes a strong basis for the development towards human-computer interaction systems.

Recent advances in image analysis, pattern recognition and computer vision open up a window towards automatic detection and classification of facial emotion. The automatic facial emotion analysis could bring facial emotion into human-machine interactions as a new modality and makes the interaction tighter and more efficient. Therefore, analysis of facial expression would be highly beneficial for fields as diverse as in behavioral sciences, medicines, physiology, security and computer science (Pantic & Rothkrantz, 2000a).

Nowadays, with advanced technology, facial emotion recognition (FER) is directed into the development of robust human-computer interface (HCI). This HCI has focused on the invention of social welfare robot that can assist the physically disabled who are bedridden or wheelchair bound to gain mobility. As the robot becomes part of our living space, it is important for a robot to understand human's mood and emotion so that a better understanding and interaction between humans and machine can be achieved. For instance, the invention of nurse robot could tend to a patient in hospital via estimating his/her emotion and create an appropriate response to the particular emotions displayed.

Dai et al. (2001) proposed a new method for monitoring patients in bed by utilizing the analysis of FER. (Breazeal, 2003) developed an expressive humanoid robot called *Kismet* that is able to perceive natural social cues either from visual or auditory

channel and delivers the social signals to the man through facial expression. (Anderson & McOwan, 2004) developed *EmotiChat* application that capture the expression (e.g. *happy*) and automatically insertion emotion icon “*happy*” into the text in the *EmotiChat* application without typing.

In behavioral sciences and medicine for instance, facial expression has been used for pain monitoring system in the work of (Hammal & Kunz, 2012). Their system utilizes facial expression to help patients who are unable to convey pain in words (e.g. newborn baby or persons with serious cognitive impairments like autism). Recently, facial expression was used to generate the differences between genuine and simulated pain. (Littlewort, Bartlett, & Lee, 2009) introduced a system that was able to differentiate between real and faking with better accuracy by tracking patterns of the subtle muscle movement from subject faces.

Due to the sizable impact of FER in daily life, various approaches have been proposed in the literature. Although FER has reached a certain level of success (Deng, Jin, Zhen, & Huang, 2005; Donato, Bartlett, Hager, Ekman, & Sejnowski, 1999a; Feng, Pietikäinen, & Hadid, 2005; Gu, Xiang, Venkatesh, Huang, & Lin, 2012; Owusu, Zhan, & Mao, 2014; Pantic & Rothkrantz, 2000b; Shan, Gong, & McOwan, 2009; Shih, Chuang, & Wang, 2008; Zhao & Zhang, 2011) development of a robust FER is still ongoing and challenging as there are still many unsolved aspects due to various unpredictable facial variations and complicated exterior environment conditions. The problems such as intra-class variations which exist in facial expression images of the same type, inter-class similarities in facial expression images and subtle change of nonlinear facial features make it difficult to pre-locate facial regions and perform robust and accurate feature extraction.

The problem is further compounded by physiognomies of faces that vary from individual to another, thus making the recognition harder. Many efforts (Pantic & Rothkrantz, 2000a) have been made to deal with these variations in FER. Ideally, facial features should be robust to intra-class variations such as small amounts of translation, rotation, spatial scale and additive noise. Higher order spectra (HOS) features and empirical mode decomposition (EMD) features are suitable from this perspective. Therefore, the main objective of this research is to propose nonlinear feature extraction techniques (using HOS and EMD) to improve robustness and performances in FER system.

HOS offer some advantages in identifying non-linear coupling, Gaussianity deviation and features obtained from it can be invariants to rotation, translation and scaling (Chandran & Elgar, 1993). These features can be used for various practical applications. Thus, it motivates this research to further investigate HOS techniques on facial emotion recognition.

EMD is a multi-resolution technique suitable for nonlinear data. It decomposes complicated signals into frequency components so-called intrinsic mode functions (IMFs) (Huang et al., 1998). The EMD offers the benefits in which the basis functions can be directly derived from the signal itself based on the local characteristic time scale of the signal which provides full data-driven approach (Nunes et al., 2003) and often brings not only high decomposition efficiency but also sharp frequency and time localizations (Qing et al., 2010).