

**Emotion Recognition Using Facial Expression  
and Electroencephalography Features  
with Support Vector Machine Classifier**

**Student**

**Henry Candra**

**Supervisor**

**Associate Professor Steven Su**

**Professor Hung. T. Nguyen**

A thesis submitted in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy



Centre for Health Technologies

School of Electrical, Mechanical and Mechatronic Systems

Faculty of Engineering and Information Technology

May, 2017

## **Certificate of Authorship/Originality**

I, Henry Candra, certify that the work in this thesis has not previously been submitted for a degree, nor has it been submitted as part of the requirements of a degree, except as fully acknowledged within the text.

I also certify that this thesis has been written by me. Any help that I have received in my research work and in the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature:

Date: 8 May 2017

*“As long as we are mindful and aware, no one practice is better than another.”*

*DJKR*

## Acknowledgements

First and foremost I would like to express my gratitude to my principal supervisor, Associate Professor Steven Su for the opportunity to carry out this research and for his invaluable guidance, support, encouragement, and life teachings. He has been with me since the beginning of my PhD and has been through my ups and downs in my PhD life. Without his help and patience I will not be able to complete my PhD study.

I am also deeply grateful to my co-supervisor, Professor Hung Tan Nguyen for his continuous support, constant guidance, valuable time, inspiration and valuable advice.

My special thanks to all the members of the Centre for Health Technologies (CHT) and other school who have been a source of collaboration, friendships as well as good advice and sharing. My grateful especially to Dr. Mitchell Yuwono for his continuous support, guidance and encouragement; to Dr. Rifai Chai for his sharing, discussion and inspiration; to (Fr.) Dr. Ardi Handojoseno (RIP) for his prayer, advice, and motivation; to Adrian Johannes for the listening and heartening; to Daniel Roxby and Zhichao Sheng for our inspiring lunch discussion; to Ganesh Naik, Steve Ling, Simon Ting for the advice and support; to Kai Cao, Lin Ye, Tao Zhang, and many others whose name has not been mentioned here, through all of you I learn to raise my young spirit once more.

Finally and most importantly, my deep appreciation goes to my Mum, my wife, my son, and all my family. Their encouragement, moral support, understanding, constant love, and prayer become my inspiration, and this thesis is my gift to all of you.

# Contents

<b>List of Figures</b>	<b>ii</b>
<b>List of Tables</b>	<b>v</b>
<b>Nomenclature</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Emotion Recognition . . . . .	3
1.3 Application . . . . .	3
1.3.1 Facial Expression Recognition System . . . . .	5
1.3.2 EEG Emotion Recognition System . . . . .	6
1.4 Challenges . . . . .	7
1.5 Contribution of This Thesis . . . . .	8
1.6 Outline of The Thesis . . . . .	10
1.7 List of Publications . . . . .	11

<b>2</b>	<b>Literature Review</b>	<b>13</b>
2.1	Emotion Measurement . . . . .	13
2.1.1	Measurement of Motor Expression . . . . .	14
2.1.2	Measurement of Physiological Arousal . . . . .	15
2.1.3	Measurement of Subjective Feeling . . . . .	16
2.2	Face Detection and Recognition . . . . .	18
2.2.1	Face Detection . . . . .	18
2.2.2	Face Recognition . . . . .	21
2.2.3	Applications of Face Recognition . . . . .	25
2.3	Facial Expression Recognition System . . . . .	26
2.3.1	Geometric Based Method . . . . .	28
2.3.2	Appearance Based Method . . . . .	29
2.4	Database for Facial Expression Recognition . . . . .	32
2.5	Summary of Latest Researches Methodology on Facial Expression Recognition System . . . . .	34
2.6	EEG Measurement and Its Applications . . . . .	35
2.6.1	EEG Signal Processing and Signal Conditioning . . . . .	36
2.6.2	EEG Electrodes and Electrodes Positioning . . . . .	37
2.6.3	Brain Rhythms . . . . .	39
2.6.4	Applications of EEG Signals . . . . .	40
2.7	EEG emotion recognition system . . . . .	41
2.7.1	Features, Channel Selection and Number of Electrodes for EEG Emotion Recognition . . . . .	42

2.7.2	Wavelet Transform for EEG Feature Extraction . . . . .	43
2.7.3	Window Segmentation in EEG Emotion Classification . . . . .	46
2.7.4	Database for EEG Emotion Classification . . . . .	47
2.8	Summary of The Latest Researches Methodology in EEG Emotion Recognition . . . . .	49
2.9	Support Vector Machine (SVM) Classifier for Emotion Recognition System	50
2.9.1	Optimizing the SVM Classifier . . . . .	51
2.10	Summary . . . . .	53
<b>3</b>	<b>Facial Expression Recognition using E-HOG and RED E-HOG Features</b>	<b>56</b>
3.1	Method Overview . . . . .	57
3.2	Experimental Setup . . . . .	59
3.2.1	Dataset Preparation . . . . .	59
3.2.2	Experimental Outline . . . . .	59
3.3	Experimental Details . . . . .	60
3.3.1	Viola-Jones Algorithm for Facial Landmarks Detection . . . . .	61
3.3.2	E-HOG Feature Extraction Strategy . . . . .	62
3.3.3	Improving E-HOG Feature with Dimensional Reduction Technique: Constructing RED E-HOG . . . . .	64
3.3.4	Classifying Facial Expression using Optimized Multi-class SVM .	66
3.4	Discussion . . . . .	67
3.4.1	Classification Results with E-HOG Feature and Comparison to Original HOG . . . . .	67
3.4.2	Classification Results of RED E-HOG Feature . . . . .	69

3.4.3	Comparison of Classification Processing Time Between RED E-HOG vs. ORI E-HOG . . . . .	70
3.4.4	Analysis of Classification Results for Each Emotion with Confusion Matrix of RED E-HOG Face . . . . .	71
3.4.5	Analysis of The Dimensional Reduction Process with Scatter Plots Between Feature Vectors . . . . .	72
3.4.6	Comparison of ORI E-HOG and RED E-HOG to Other Methods	72
3.5	Summary . . . . .	74
<b>4</b>	<b>Discrete EEG-Emotions Recognition Using Wavelet Features</b>	<b>75</b>
4.1	Method Overview . . . . .	76
4.2	Experimental Setup . . . . .	78
4.2.1	Dataset Resource and Dataset Preprocessing . . . . .	78
4.2.2	Subject Grouping from The Dataset . . . . .	79
4.2.3	Dataset Mapping . . . . .	79
4.2.4	Experimental Outline . . . . .	80
4.3	Experimental Detail . . . . .	81
4.3.1	Subject Grouping . . . . .	81
4.3.2	EEG Feature Extraction with Discrete Wavelet Transform (DWT)	83
4.3.3	EEG subband selection . . . . .	84
4.3.4	EEG channel selection . . . . .	84
4.3.5	Optimizing the Multi-class SVM Classifier . . . . .	85
4.3.6	Receiver operating characteristic (ROC) for Classification Analysis	86
4.3.7	Normalized Mutual Information (NMI) for EEG Channel Analysis	88



4.4	Discussion . . . . .	89
4.4.1	Classification Results of The Discrete EEG-Emotion Recognition	89
4.4.2	Confusion Matrix of The Best Trained Classifier . . . . .	91
4.4.3	ROC Analysis of The Classification Results . . . . .	92
4.4.4	NMI Analysis of The EEG Channel Selection . . . . .	92
4.5	Summary . . . . .	94
<b>5</b>	<b>Improving EEG Emotion Recognition Accuracy Using OWS Method</b>	<b>95</b>
5.1	Method Overview . . . . .	96
5.2	Experimental Setup . . . . .	99
5.2.1	Dataset Resource and Dataset Preparation . . . . .	99
5.2.2	Subject Grouping . . . . .	100
5.2.3	Experimental Overview . . . . .	100
5.3	Experimental Detail . . . . .	102
5.3.1	Subject Grouping . . . . .	103
5.3.2	EEG Emotion Signals Segmented into 12 Window Sizes . . . . .	103
5.3.3	Wavelet Feature Extraction: Wavelet Energy and Wavelet Entropy	104
5.3.4	Creating an array of 32 channels wavelet PSD / entropy in 5 or 3 bands formation . . . . .	105
5.3.5	Statistical Features: <i>MTD</i> and <i>MAE</i> . . . . .	105
5.3.6	Fast Fourier Transform Power Spectral Density ( <i>FFT PSD</i> ) . . . . .	106
5.3.7	Recognizing the EEG Emotion with SVM Classifier . . . . .	107
5.4	Discussion . . . . .	107

5.4.1	Analysis of Segmented EEG Signals with Wavelet PSD Spectrum: Benefit of Window Segmentation . . . . .	108
5.4.2	Comparative Summary of The Classification Results for 7 Features	109
5.4.3	Anaysis for The Wavelet Features . . . . .	110
5.4.4	Comparison of MAE to other Features . . . . .	111
5.4.5	Analysis of Training and Testing Time: Finding Window with Less Processing Time . . . . .	112
5.4.6	Suggested optimal window selection (OWS) in EEG emotion recognition using 7 features for all emotions . . . . .	113
5.5	Summary . . . . .	115
<b>6</b>	<b>Conclusions and Future Direction</b>	<b>116</b>
6.1	Discussion . . . . .	116
6.2	Conclusion . . . . .	118
6.3	Limitation of The Research . . . . .	121
6.4	Future Directions . . . . .	122
	<b>Appendix</b>	<b>123</b>
	<b>Bibliography</b>	<b>165</b>

# List of Figures

1.1	Block diagram an emotion recognition system . . . . .	4
1.2	Russell’s circumplex model of affect Russell (1980). Horizontal axis represents valence (pleasure); vertical axis represents arousal. Artwork is as seen in Valenza et. al (Valenza <i>et al.</i> , 2014). . . . .	6
2.1	Block diagram of a consensual componential model of emotion. (Mauss & Robinson, 2009) . . . . .	13
2.2	Self Assessment Manikin SAM (Bradley & Lang, 1994) . . . . .	18
2.3	The Haar feature (Viola & Jones, 2001) . . . . .	20
2.4	Sources of facial expressions (Fasel & Luetttin, 2003) . . . . .	27
2.5	Facial expressions analysis (Tian <i>et al.</i> , 2011) . . . . .	28
2.6	Block diagram of HOG computation (Dalal & Triggs, 2005) . . . . .	30
2.7	EEG electrodes positioning with 10/20 international system (Roman-Gonzalez, 2012) . . . . .	39
2.8	Brainwave Rhythms in various band frequencies . . . . .	40
2.9	Segmentation of EEG signal with different window size. The EEG signal is segmented into different window size to obtain most effective length of EEG signal to be used in EEG emotion classification (Candra <i>et al.</i> , 2015a). . . . .	47

3.1	The block diagram of the proposed emotion recognition algorithm. . . . .	60
3.2	The facial landmarks extraction using Viola-Jones' algorithm. . . . .	62
3.3	HOG vs E-HOG features of the eye and mouth. Sparse <b>E</b> matrix in the E-HOG computation contributes to a substantial amount of cells having zero values, yielding simpler features and a leaner extraction process. . .	63
3.4	Classification results with various E-HOG features after applying PCA. .	69
3.5	Classification result with various E-HOG features after applying a combination of PCA and LDA. . . . .	70
3.6	Scatter plot of ORI E-HOG features before and after dimensionality reduction using PCA and a combination of PCA and LDA. . . . .	73
4.1	Mapping of 4 discrete emotions to Russell's circumplex model of affect (Russell, 1980). Horizontal axis represents valence (pleasure); vertical axis represents arousal. . . . .	77
4.2	Block diagram of the Discete EEG-emotion recognition. . . . .	80
4.3	Confusion matrix of binary classification with 4 possible outcomes. . . .	87
4.4	Receiver Operating Characteristic (ROC) of an average case scenario. . .	92
4.5	The Normalized Mutual Information between each channel and combination of emotions. . . . .	93
5.1	Recommendation of 3 to 12s for EEG emotion classification of arousal and valence emotional states (Candra <i>et al.</i> , 2015a). . . . .	98
5.2	Block diagram of Optimal Window Selection (OWS) method. . . . .	100
5.3	Result of subject grouping to overcome mistranslation effect of SAM rating method. The Euclidean distance reveals how closely a participant is related with another. Participants with similar response patterns will have relatively closer distance. . . . .	103

5.4	The signal segmentation scheme with 12 window sizes. . . . .	104
5.5	An array of 32 channels wavelet PSD / entropy in 5 or 3 bands formation as the representation of one segment EEG-emotion signal. . . . .	105
5.6	Visualization of 1 channel EEG signals in their related 5 stacked bands wavelet PSD spectrum of 60s window compared to 15 consecutive 4s window (marked by red dotted oval) showing the repetitional pattern between both windows. This is the benefit of the segmentation process. . . .	108
5.7	Comparative summary of classification results for 4 emotions with 7 features represented in line graphs. . . . .	109
5.8	Simplified graph of 7 features using weighted average accuracy of 4 emotions. . . . .	110
5.9	Graphical representation of optimal training and testing time in EEG-emotion recognition for 7 features using weighted average of 4 emotions. . . . .	113
5.10	Suggested optimal window selection (OWS) color graph. The numbers in the graph reflect the weighted average accuracy of 4 emotions for each 7 features. The color heatmap expresses various levels of accuracy yield by window size between 1s and 60s for each of the 7 features, where darker colors represent higher accuracy. Referring to the graph, the optimal window can be identified. . . . .	114

# List of Tables

2.1	Summary of Databases for Facial Expression Recognition . . . . .	33
2.2	Comparison Between Facial Expression Recognition Methods . . . . .	34
2.3	Summary of Databases for EEG Emotion Recognition . . . . .	48
2.4	Comparison Between EEG Emotion Recognition Methods . . . . .	49
3.1	Comparison of average classification accuracy between ORI HOG and E-HOG (30 repetitions) . . . . .	68
3.2	Comparison of average classification processing time between ORI HOG and E-HOG (30 repetitions) . . . . .	68
3.3	Confusion Matrix of the best trained classifier using Face E-HOG features	69
3.4	Average training time of ORI E-HOG vs. 10 Dimensions RED E-HOG trained with multi-class SVM in 30 repetitions . . . . .	71
3.5	Average testing time of ORI E-HOG vs. 10 Dimensions RED E-HOG trained with multi-class SVM in 30 repetitions . . . . .	71
3.6	Confusion Matrix of the best trained SVM classifier using 10 dimensions RED E-HOG Face . . . . .	71
3.7	Comparison of ORI E-HOG, RED E-HOG Faces to Other Facial Expre- sion Recognition Methods . . . . .	72

4.1	Classification Results of The Discrete EEG-Emotion Recognition . . . . .	91
4.2	Confusion matrix of the best trained classifier using wavelet entropy under $CH_{18}$ , $SB_3$ setting. . . . .	91
5.1	t-test of 3 vs. 5 bands wavelet entropy features for valence emotion with hypothesis that, "The difference of accuracy between 3 and 5 bands entropy features is not significant" ( $p > 0.05$ ) . . . . .	111
5.2	t-test of $MAE$ vs. $FFT PSD$ feature for liking emotion with hypothesis that, "The difference of accuracy between $MAE$ and $FFT PSD$ features is significant" ( $p < 0.05$ ) . . . . .	112

# Nomenclature

AAM	:	Active Appearance Models
ADCs	:	Analogue-to-Digital Converters
AI	:	Artificial Intelligent
AU	:	Action Units
BCI	:	Brain-Computer interface
CK+	:	Extended Cohn–Kanade database
CoC	:	Coefficient of Correlation
CT	:	Computerized Tomography
CWT	:	Continuous Wavelet Transform
DEAP	:	Database for Emotion Analysis Using Physiological Signals
DWT	:	Discrete Wavelet Transforms
ECG	:	Electrocardiography
EEG	:	Electroencephalography
EGG	:	Electrogastrography
E-HOG	:	Edge-Histogram of Oriented Gradients
EMG	:	Electromyography
EOG	:	electroocclugraphy
ERCE	:	Ensemble Rapid Centroid Estimation
ERD	:	Event-Related Desynchronization
ERP	:	event- related potentials
ERS	:	event-related synchronization
FACS	:	Facial Action Coding System
FD	:	Fractal Dimension
FFT	:	Fast Fourier Transform
fMRI	:	Functional Magnetic Resonance Imaging
GB	:	Giga Bytes
GLCM	:	Gray-Level Co-occurrence Matrix
GMM	:	Gaussian Mixture Model
GSR	:	Galvanic Skin Response
HCI	:	Human-Computer Interaction
HOG	:	Histogram of Oriented Gradients
HOS	:	Higher Order Spectra



IADS	:	International Affective Digitized Sound System
IAPS	:	International Affective Picture System
ICA	:	Independent Component Analysis
ID	:	Identification
JS	:	Jensen-Shannon
KL	:	Kullback-Leibler
KPCA	:	Kernel Principal Component Analysis
LBP	:	Local Binary Patterns
LDA	:	Linear Discriminant Analysis
MEG	:	Magnetoencephalography
MSCE	:	Magnitude Square Coherence Estimation
MSR	:	Manifold based Sparse Representation
NSI	:	Non Stationary Index
ORI E-HOG	:	Original Edge-Histogram of Oriented Gradients
OWS	:	Optimal Window Selection
PAD	:	Pleasure, Arousal and Dominance Emotion Scales
PCA	:	Principal Components Analysis
PET	:	Positron Emission Tomography
PSD	:	Power Spectral Density
RBF	:	Radial Basis Function
RCE	:	Rapid Centroid Estimation
RED E-HOG	:	Reduced Edge-Histogram of Oriented Gradients
RHMM	:	Regional Hidden Markov Model
SAM	:	Self-Assessment Manikin
SIFT	:	Scale Invariant Feature Transform
SMO	:	Sequential Minimal Optimization
SPET	:	Single Photon Emission Tomography
SVM	:	Support Vector Machine
WE	:	Wavelet Energy
WT	:	Wavelet Transform

# Abstract

Recognizing emotions from facial expression and electroencephalography (EEG) emotion signals are complicated tasks that require substantial issues to be solved in order to achieve higher performance of the classifications, i.e. facial expression has to deal with features, features dimensionality, and classification processing time, while EEG emotion recognition has the concerned with features, number of channels and sub band frequency, and also non-stationary behaviour of EEG signals. This thesis addresses the aforementioned challenges.

First, a feature for facial expression recognition using a combination of Viola-Jones algorithm and improved Histogram of Oriented Gradients (HOG) descriptor termed Edge-HOG or E-HOG is proposed which has the advantage of insensitivity to lighting conditions. The issue of dimensionality and classification processing time was resolved using a combination of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) which has successfully reduced both the dimension and the classification processing time resulting in a new low dimension of feature called Reduced E-HOG (RED E-HOG).

In the case of EEG emotion recognition, a method to recognize 4 discrete emotions from arousal-valence dimensional plane using wavelet energy and entropy features was developed. The effects of EEG channel and subband selection were also addressed, which managed to reduce the channels from 32 to 18 channels and the subband from 5 to 3 bands.

To deal with the non-stationary behaviour of EEG signals, an Optimal Window Selection (OWS) method as feature-agnostic pre-processing was proposed. The main objective

of OWS is window segmentation with varying window which was applied to 7 various features to improve the classification results of 4 dimensional plane emotions, namely arousal, valence, dominance, and liking, to distinguish between the high or low state of the aforementioned emotions. The improvement of accuracy makes the OWS method a potential solution to dealing with the non-stationary behaviour of EEG signals in emotion recognition. The implementation of OWS provides the information that the EEG emotions may be appropriately localized at 4–12 seconds time segments.

In addition, a feature concatenating of both Wavelet Entropy and average Wavelet Approximation Coefficients was developed for EEG emotion recognition. The SVM classifier trained using this feature provides a higher classification result consistently compared to various different features such as: simple average, Fast Fourier Transform (FFT), and Wavelet Energy.

In all the experiments, the classification was conducted using optimized SVM with a Radial Basis Function (RBF) kernel. The RBF kernel parameters were properly optimized using a particle swarm ensemble clustering algorithm called Ensemble Rapid Centroid Estimation (ERCE). The algorithm estimates the number of clusters directly from the data using swarm intelligence and ensemble aggregation. The SVM is then trained using the optimized RBF kernel parameters and Sequential Minimal Optimization (SMO) algorithm.