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

UNIVERSITY OF TECHNOLOGY SYDNEY

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

List of Figures

List of Tables

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	:	electrooclugraphy
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.