RECOGNITION OF FACIAL ACTION UNIT BASED ON SPATIAL-TEMPORAL BAYESIAN PROBABILISTIC TECHNIQUE

LOH YEE KOON

UNIVERSITI SAINS MALAYSIA

2016

RECOGNITION OF FACIAL ACTION UNIT BASED ON SPATIAL-TEMPORAL BAYESIAN PROBABILISTIC TECHNIQUE

by

LOH YEE KOON

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

September 2016

ACKNOWLEDGEMENTS

With the completion of this thesis, I would like to show my deepest gratitude to many people who had assisted and guided me in generating research ideas and giving me mental support in completing this thesis.

First of all, I would like to thank Universiti Sains Malaysia for providing the research grants that support this project. This project was supported in part by the Universiti Sains Malaysia's Postgraduate Research Grant Scheme (1001/PELECT/8046014), and Universiti Sains Malaysia's Research University Grant (1001/PELECT/814208).

I would like to express my greatest appreciation to my supervisor, Assoc. Prof. Dr. Shahrel Azmin Suandi for providing invaluable guidance, advice, support and criticism throughout the years. Assoc. Prof. Dr. Shahrel has always been motivating, encouraging and inspiring. Thanks to Assoc. Prof. Dr. Shahrel for introducing me to the fundamental concepts of Artificial Intelligence and Machine Learning. He has provided me with a large scale of freedom in exploring the opportunities to excel in my research. Besides that, he has encouraged me to attend conferences to deepen and widen my foresight by meeting with people from academia and industry. His knowledge and instructions helped me to conquer a lot of challenges throughout the journey.

Furthermore, I would like to express my appreciation to my comrades Jing Huey, Boon Tatt and Alex. They have shared knowledge and experience and being supportive and cheery. I truly appreciate their companion, sharing of ideas as well as their concern.

My warmest regards to my beloved family for their caring encouragement and moral support throughout this journey. They are always being understanding and believing in me for these years. Last but not least, I would like to thank all whoever directly or indirectly help in making this thesis successful.

Loh Yee Koon 2016

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	2 Dimensional
2TBN	2-slice Temporal Bayes Net
AAM	Active Appearance Model
AdaBoost	Adaptive Boosting
AURS	Action Unit Recognition System
AUs	Action Units
BIC	Bayesian information criterion
BN	Bayesian Network
CK	Cohn-Kanade
CPD	Conditional Probability Distribution
CPT	Conditional Probability Table
CR	Classification Rate
DAG	Directed Acyclic Graph
DBN	Dynamic Bayesian Network
DNMF	Discriminant Non-negative Matrix Factorization
EM	Expectation-Maximization
FACS	Facial Action Coding System
FAR	False Alarm Rate
FERA	Facial Expression Recognition and Analysis
GERG	Geneva Emotion Research Group
HMM	Hidden Markov Model
I-AURS	Improved-Action Unit Recognition System
KFM	Kalman Filter Model
LBPs	Local Binary Patterns
LGBPs	Local Gabor Binary Patterns
LDA	Linear Discriminant Analysis
LPQ	Local Phase Quantization

MDL	Minimum Description Length
MML	Minimum Message Length
ML	Maximum likelihood
NN	Neural Network
PCA	Principal Component Analysis
PGM	Probabilistic Graphical Model
PHOG	Pyramid of Histogram of Gradient
PI	Person Independent
PS	Person Specific
ROI	Region Of Interest
SSPNET	Social Signal Processing Network
SSS	Stochastic Structure Search
SVM	Support Vector Machine
TPR	True Positive Rate
VC	Vapnik-Chervonenkis

LIST OF SYMBOLS

θ	Parameters for a node
$d-$	Shortest distance from the separating hyperplane to the closest negative example
$d+$	Shortest distance from the separating hyperplane to the closest positive example
ε	Error
$f(x)$	Function of x
$h(x)$	Weak classifier
$H(x)$	Strong classifier
i	Integer
I	Image
J	Junction tree
$K(x,x')$	Kernel function of x
$L(\theta)$	Log likelihood of parameters
M	Marginal value
n	Positive integer
$Pa(Z(i,t))$	Parent nodes of $Z(i,t)$
$P(Z_t\|Z_{t-1})$	Conditional probability of Z at time slice t given the probability of Z at previous time slice $t-1$
$P(X)$	Joint probability distribution for X
$P(X_i\|Pa_i)$	Conditional probability of X_i given the probability of the parents of node X_i
s	Seconds
S	Network structure
t	Time
T	Total time slice
$w_{t,i}$	Weights
$\mathbf{w}$	The normal to hyperplane

x, y	The spatial coordinates on the image
X	Set of variables in a network structure
y_T	Threshold marginal value
$Z(i,t)$	A node at time slice t

PENGECAMAN UNIT TINDAKAN WAJAH BERDASARKAN TEKNIK SPATIAL-TEMPORAL KEBARANGKALIAN BAYESIAN

ABSTRAK

Pengecaman Unit Tindakan muka sering digunakan sebagai kerja-kerja asas untuk mengkaji ekspresi wajah atau aplikasi pergerakan manusia seperti pemantauan video dan pengenalan muka. Unit Tindakan (AUs) bekerja sebagai unit asas dalam Sistem Kod Tindakan Muka (FACS) untuk taksonomi pergerakan muka, yang mengaitkan setiap AU dengan pengaktifan satu atau lebih otot muka khusus. Pembinaan sistem pengecaman Unit Tindakan yang stabil tetap menjadi cabaran bagi penyelidik disebabkan aksi, pencahayaan dan gabungan rumit ekspresi wajah. Sehubungan dengan itu, kerja penyelidikan ini mencadangkan satu pendekatan kebarangkalian yang model hubungan statik dan tempoh antara AUs dari urutan imej menggunakan Dynamic Bayesian Network (DBN). DBN yang menggabungkan imej ukuran kepada model DBN direka untuk menjadi satu struktur umum untuk hubungan AU. Support Vector Machine (SVM) digunakan untuk mendapatkan ukuran AU dari pangkalan data dengan mengklasifikasikan setiap AU daripada ciri imej. Ukuran AU tersebut kemudian digunakan sebagai bukti kepada DBN untuk membuat kesimpulan kewujudan pelbagai AU. Kemuncak penyelidikan ini adalah bahawa parameter AU dalam model DBN dipelajari daripada kaedah data tidak lengkap, dengan nod AU pembolehubah tersembunyi dan disimpulkan daripada ukuran imej secara langsung dan dimodelkan dengan cara kebarangkalian yang dinamik. Kerja penyelidikan ini mencadangkan bahawa setiap AU mempunyai keputusan ambang berbeza kerana sambungan yang berbeza daripada AU

dalam model dengan mencari ambang yang terbaik bagi setiap AU. Keputusan eksperimen menunjukkan bahawa dengan membuat kesimpulan AU dari ukuran imej sebagai model terdahulu, model yang dicadangkan mencapai keputusan yang setanding dengan model yang dipelajari sepenuhnya daripada pangkalan data tertentu. Sistem ini mencapai kadar pengiktirafan purata sebanyak 94.78% dengan kadar positif benar sebanyak 70.54% dan kadar penggera palsu 2.31% menggunakan pangkalan data Cohn-Kanade (CK). Pendekatan kebarangkalian yang dicadangkan itu juga telah digunakan untuk cabaran Pengiktirafan Ekspresi Wajah dan Analisis (FERA) yang dianjurkan oleh Pemprosesan Isyarat Rangkaian Sosial (SSPNET) pada tahun 2011. Cabaran ini bertujuan untuk membolehkan perbandingan yang adil di antara sistem dengan mempunyai keperluan untuk prosedur penilaian yang seragam. Cabaran ini digunakan sebagai penanda aras sistem ekspresi wajah di seluruh dunia. Pendekatan kebarangkalian yang dicadangkan itu telah direka bentuk semula dengan mengikut arahan yang diberikan oleh cabaran dan model baru dibina dan dilatih untuk cabaran FERA. Sistem yang dicadangkan mencapai prestasi lebih baik daripada kaedah asas dalam cabaran dan ia telah menunjukkan hasil yang setanding dengan keadaan-keadaan lain dan peserta dalam cabaran tersebut. Metrik prestasi yang digunakan di FERA ialah ukuran F1 dan keputusan keseluruhan mencapai ukuran F1 pada 0.494 mengatasi kerja-kerja lain termasuk satu-satunya pasukan yang menggunakan pendekatan kebarangkalian dalam kerja mereka. Oleh itu, sistem yang dicadangkan telah memenuhi objektif kajian dengan pembelajaran parameter dari kaedah data tidak lengkap, umum kepada pangkalan data yang berbeza serta keadaan yang berbeza untuk bersaing dengan kerja-kerja lain di dunia.

RECOGNITION OF FACIAL ACTION UNIT BASED ON SPATIAL-TEMPORAL BAYESIAN PROBABILISTIC TECHNIQUE

ABSTRACT

Facial Action Unit recognition is often used as elementary works for facial expressions analysis or human motions applications such as video surveillance and face identification. Action Units (AUs) are employed as basic unit in Facial Action Coding System (FACS) to taxonomize facial movements; by associating each AUs with the activation of one or more specific facial muscles. A stable Action Unit recognition system still remains a challenge for researchers due to pose, illuminations and complicated combination of facial expression. With this regards, this research work proposes a probabilistic approach which models spatial and temporal relationships of AUs from image sequence using Dynamic Bayesian Network (DBN). The state-of-the-art DBN, which incorporates AU measurements from images to a DBN model is designed to be a generic structure for AU relationships. Support Vector Machine (SVM) is used to obtain AU measurements from database by classifying each AUs from image features. Such AU measurements are then applied as evidence to the DBN for inferring existence of various AUs. The highlight of this work is that AU parameters in DBN model are learned from incomplete data method, where the AU nodes are hidden variables and directly inferred from image measurements and modeled in dynamic and probabilistic way. This research work proposed that each AUs has different decision threshold due to different connections of AUs in the model by searching the best threshold for each AUs. Experimental results show that by inferring AU from im-

age measurements, the proposed model achieves comparable results to the model that learned completely from specific database. This system achieves average recognition rate at 94.78% with a true positive rate of 70.54% and false alarm rate of 2.31% using Cohn-Kanade (CK) database. The proposed probabilistic approach has also been applied to the Facial Expression Recognition and Analysis (FERA) challenge which was hosted by the Social Signal Processing Network (SSPNET) in 2011. The challenge aims to allow a fair comparison between systems, by having a need for standardized evaluation procedures. This challenge is used as the benchmark of facial expression system around the world. The proposed probabilistic approach has been redesigned to follow the instructions given by the challenge and a new model is built and trained for FERA challenge. The proposed probalisitic approach is proven to be applicable and generalized to different conditions. The proposed system is compared against the baseline system for the challenge provided by the FERA organizers. The proposed system achieved better performance than the baseline system and achieved comparable results with other state-of-the-art and participants in the challenge. The performance metric used in FERA is F1-measure and the overall result achieves 0.494 for F1-measure, outperforming other works including the one and only team which use probabilistic approach in their work. Hence, the proposed system has met the objectives of research by learning parameters from incomplete data method, generalized to different database as well as different conditions to compete with other works in the world.

CHAPTER ONE

INTRODUCTION

1.1 Background

A facial expression is indeed the clues for human behavior. A person's internal emotional states and intention can be figured out through his facial changes. Facial behavior analysis has gained extensive attentions from researchers across many areas of studies such as biometrics, computer vision and psychological studies (Matsumoto and Hwang, 2011). Facial expression recognition is one of very active topics in research in recent years.

Face is one's identity and unique for everyone (Novotney, 2011), and it changes with age. It is a complex study due to different expressions for different people and the expressions differ at different age. A facial expression is generated by activation or relaxation of facial muscle focusing around eyes and mouth (Ekman, 2009). Around the world, more than 500 people including neurologists, psychiatrists and psychologists have learned Dr. Ekman's research tool called FACS, or Facial Action Coding System, for deciphering which of the 43 muscles in the face are working at any given moment, even when an emotion is so fleeting that the person experiencing it may not be conscious of it (Foreman, 2003).

Facial expression is important for market researchers and product developers (nViso, 2011b). The emotional response of consumers are measured to involve customers earlier in the development cycle so improving likelihood of success in launching new