Real–time Facial Expression Analysis

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Ph.D.)

in

Computer Science

at Massey University, Auckland,

New Zealand.

Chao Fan

2008

Abstract

As computers have become more and more advanced, with even the most basic computer capable of tasks almost unimaginable only a decade ago, researchers and developers are focusing on improving the way that computers interact with people in their everyday lives. A core goal, therefore, is to develop a computer system which can understand and react appropriately to natural human behavior.

A key requirement for such a system is the ability to automatically, and in real time, recognises human facial expressions. In addition, this must be successfully achieved regardless of the inherent differences in human faces or variations in lighting and other external conditions.

The focus of this research was to develop such a system by evaluating and then utilizing the most appropriate of the many image processing techniques currently available, and, where appropriate, developing new methodologies and algorithms.

The first key step in the system is to recognise a human face with acceptable levels of misses and false positives. This research analysed and evaluated a number of different face detection techniques, before developing a novel algorithm which combined phase congruency and template matching techniques. This novel algorithm provides key advantages over existing techniques because it can detect faces rotated to any angle, and it works in real time. Existing techniques could only recognise faces which were rotated less than 10 degrees (in either direction) and most could not work in real time due to excessive computational power requirements.

The next step for the system is to enhance and extract the facial features. To successfully achieve the stated goal, the enhancement and extraction of the facial features must reduce the number of facial dimensions to ensure the system can operate in real time, as well as providing sufficient clear and detailed features to allow the facial expressions to be accurately recognised. This part of the system was successfully completed by developing a novel algorithm based on the existing Contrast Limited Adaptive Histogram Equalization technique which quickly and accurately represents facial features, and developing another novel algorithm which reduces the number of feature dimensions by combining radon transformation and fast Fourier transformation techniques, ensuring real time operation is possible.

The final step for the system is to use the information provided by the first two steps to accurately recognise facial expressions. This is achieved using an SVM trained using a database including both real and computer generated facial images with various facial expressions.

The system developed during this research can be utilised in a number of ways, and, most significantly, has the potential to revolutionise future interactions between

humans and computers by assisting these reactions to become natural and intuitive. In addition, individual components of the system also have significant potential, with, for example, the algorithms which allow the recognition of an object regardless of its rotation under consideration as part of a project aiming to achieve non-invasive detection of early stage cancer cells.

Acknowledgement

In 1999, I left my beautiful motherland China to move to a strange country called New Zealand to improve my English and further my studies. I was an international student fully supported by my parents (Prof. Dr. ZhanGuo Fan, QiuLing Gong). It was impossible for me to continue my study without their help and it has been their continuing support and encouragement which has made my study in New Zealand possible, special thank you Mum and Dad.

In 2001, I met my main supervisor Dr. Abdolhossein Sarrafzadeh, who changed my life. He helps me thoughtfully and he gave valuable advice for my study and in 2003, he encouraged me to undertake my PHD study. His good suggestions and comments always kept my study and my research on the right track. I learned not only knowledge, but also spirit from him. This spirit is never give up no matter what the problem and this spirit will stay with me forever.

I would like to thank for Dr Hamid Gholamhosseini and Dr Martin Johnson, my co-supervisors for their kind help and advice with my research, and the way they generously shared their knowledge during the project.

I would like to thank for Zhuo Li, my wife, for her continuous support and help at all times.

I would like to thank for Ben Powles, who kindly and carefully read and corrected my thesis. Without his help this thesis would not read so smoothly.

I would like to thank Sam Alexander, who always helps me correct my English, as well as Farhad Dadgostar and Andre Barczak, my classmates and my friends. We helped each other, and we had wonderful time in IIMS, Massey University.

Finally, thank you for all New Zealanders who helped and supported me, your kind help made me consider New Zealand as another home.

Table of contents

Abstract II
Acknowledgement IV
Table of contents V
List of Figures VII
List of Tables IX
Chapter 1. Introduction 1
1.1. Background 2
1.2. Summary of research goals 5
1.3. Statement of the problem 6
1.4. Motivation and scope of this research 8
1.5. Contribution of this research 10
1.6. Thesis Overview 12
1.6.1. Thesis outline 12
Chapter 2. Literature review 13
2.1. Face detection 14
2.1.1. Face detection based on Skin Color 14
2.1.2. Template matching 15
2.1.3. Frequency domain 16
2.1.4. Geometrical Moment and Zernike Moment 17
2.1.5. Viola and Jones's method 19
2.1.6. ANN for object detection 19
2.1.7. Statistical K-mean 20
2.1.8. Support Vector Machines 21
2.2. Image processing for facial feature extraction 22
2.2.1. Image noise removal 22
2.2.2. Lighting correction and facial feature enhancement 26
2.3. Facial expression analysis 29
2.3.1. Basic facial expressions 30
Chapter 3. Real-time facial expression analysis system overview 35
3.1. Background 36
3.2. Equipment and environment 36
3.3. System structure 38
3.3.1. Face detection 38
3.3.2. Facial feature extraction 40
3.3.3. Facial expression analysis 42
Chapter 4. Scale and Rotation Invariant Face Detection using Template Matching and SVM 43
4.1. Introduction 43
4.2. Optimized Candidates using Similarity Measurement (OCSM) 47
4.2.1. Image feature extraction by phase congruency algorithm 50
4.2.2. Template matching for facial feature extraction 54

4.2.3. Template for face matching 60
4.2.4. Fast Fourier Transform 61
4.2.5. Scale invariance and rotation invariance 64
4.2.6. Face Classifier training using an SVM 68
4.2.7. Experimental result 71
Chapter 5. Facial image enhancement 75
5.1. Introduction 75
5.2. Eliminating noise 76
5.2.1. Sigma Filtering – modified for real time applications 81
5.3. Lighting correction 83
5.4. Facial features contrast enhancement 84
5.4.1. A novel facial feature enhancement method 87
5.5. Edge preserving smoothing process 90
5.6. Summary of experimental results 92
5.6.1. Image size 93
5.7. Eye detection and face normalization 94
5.8. Summary 97
Chapter 6. Real time Facial expression analysis 98
6.1. Basic facial expressions 102
6.2. Facial feature extraction for facial expression analysis 107
6.3. Classification of facial expressions - ANNs and SVMs 119
6.3.1. Summary 119
6.3.2. ANNs Overview 119
6.4. Training of the SVM 127
6.4.1. “One against All” and “One against One” classification: 128
6.4.2. Cross validation 129
6.5. Testing of kernel functions 132
6.6. Facial expression classification system - summary of findings 134
6.7. Facial expression recognition – testing and results 135
6.7.1. Still images and video sequence analysis 135
6.7.2. Screen shots of real time facial expression analysis system 138
6.7.3. Conclusion 140
Chapter 7. Discussions and Conclusions 142
7.1. Real-time face detection 142
7.2. Facial feature enhancement and extraction 144
7.3. Facial expression detection and analysis 146
7.4. Suggestions for future research 147
Bibliography 150
Glossary 159
Index 161

List of Figures

Figure 1.1 Structure of NGITS 9
Figure 3.1 The set up for the normal environment including a computer & web camera .. 36
Figure 3.2 Logitech QuickCam 4000 used for the experiments in the research 37
Figure 3.3 Real time facial expression analysis system structure 38
Figure 4.1 Difference of Gaussian (DOG) pyramid 45
Figure 4.2 Key-point descriptor 45
Figure 4.3 Matched image for the same person - 51 key points found 46
Figure 4.4 Matched image for different people - 33 key points found. 46
Figure 4.5 Face detection using an averaged face pyramid. 49
Figure 4.6 Flowchart of the scale and rotation invariant face detection 50
Figure 4.7 The relationship between phase congruency, Local energy and the sum of the Fourier amplitudes 53
Figure 4.8 Comparison of gradient based edge detection and phase congruency edge detection 54
Figure 4.9 Selected images from the Face database. 57
Figure 4.10 (a) Averaged face – color and (b) averaged face - grey level 57
Figure 4.11 Training samples for face detection 60
Figure 4.12 Rotated Face templates 66
Figure 4.13 Fourier Transformation and Correlation coefficients 66
Figure 4.14 Image template matching by Correlation coefficients 67
Figure 4.15 Training samples for face detection 69
Figure 4.16 Artificial Neural Networks for Face Detection. 70
Figure 4.17 Comparison result with openCV 71
Figure 4.18 Result of rotate face detection 72
Figure 4.19 Comparison result with openCV 73
Figure 4.20 Performance of OCSM 74
Figure 4.21 Accuracy of OCSM 74
Figure 5.1 Steps required for facial image enhancement. 76
Figure 5.2 Performance comparsion on different size of images by different noise removal algorithms 78
Figure 5.3 Comparison of noise removal algorithms as applied to face images without a moustache 79
Figure 5.4 Comparison of noise removal algorithms as applied to face images with a moustache 80
Figure 5.5 Pseudo code of sigma filtering 82
Figure 5.6 Sigma Filtering (Sigma = 10) 82
Figure 5.7 Result of Lighting Correction Using Morphological Operation 84
Figure 5.8 Different feature enhancement algorithms as applied to a face image. 85
Figure 5.9 Skin color histogram analysis 85
Figure 5.10 A novel method for facial feature enhancement 88

Figure 5.11 The results of applying the algorithms on faces 89
Figure 5.12 More test results of facial feature enhancement 89
Figure 5.13 Kuwahara filtering of an image with 5x5 pixels. 91
Figure 5.14 The performance comparison between K-means clustering and Kuwahara filtering 92
Figure 5.15 More results for facial features enhancement on different people using the novel algorithm developed during this research. 93
Figure 5.16 Time complexity of the Facial Feture Enhancement Algoritnm 94
Figure 5.17 Integral image feature calculation 95
Figure 5.18 Property of eye region 95
Figure 5.19 Result of eye detection and face detection 96
Figure 5.20 Positions of the eyes in the normalized image 96
Figure 6.1 Examples of happy expressions 102
Figure 6.2 Examples expressions of Sadness 103
Figure 6.3 Examples of anger expressions 103
Figure 6.4 Examples of fear expressions 103
Figure 6.5 Examples of disgust expressions 104
Figure 6.6 Examples of surprise expressions 104
Figure 6.7 Comparison of an image's histogram before and after rotation 108
Figure 6.8 Multi-Classifier for Facial Feature Expression Analysis 109
Figure 6.9 Outline extractions for Chinese characters 110
Figure 6.10 Eight-directions connections and an example 111
Figure 6.11 The radon projection for 45 and 135 degrees 113
Figure 6.12 Feature Extraction and Radon Transformation 114
Figure 6.13 Images of extracted feature and their discrete cosine transform 115
Figure 6.14 Feature extracted image and its discrete cosine transform 116
Figure 6.15 Features dimensions reduced using the feature extraction novel algorithm 117
Figure 6.16 Comparison of the novel algorithm (b) with AAM (a) 118
Figure 6.17 Structure of an ANNs 120
Figure 6.18 The Back Propagation Algorithm 121
Figure 6.19 the structure of ANN for training facial features 121
Figure 6.20 Tansig transfer function 122
Figure 6.21 performance and training state for ANN 122
Figure 6.22 Gradient for training ANN 122
Figure 6.23 Regression on ANN (after 178073 iterations) 123
Figure 6.24 Separating two classes by hyper-planes 125
Figure 6.25 Artificial faces generated by FaceGen (natural expression) 131
Figure 6.26 SVM parameter searching 134
Figure 6.27 Video sequences analysis. 136
Figure 6.28 Frame difference on different facial expressions 137
Figure 6.29 Natural expression detected 138
Figure 6.30 Angry expression detected 139
Figure 6.31 Sad expression detected 139
Figure 6.32 Smile expression detected 140

Figure 6.33 Surprised expression detected 140

List of Tables

Table 2.1 Comparison of different noise removal algorithms..... 26
Table 5.1 Methods for image noise remove 77
Table 6.1 Facial Action Coding System 105
Table 6.2 Accuracy of different kernels. 132
Table 6.3 Comparison of Datcu & Rothkrantz's method with the proposed method. 133
Table 6.4 Exhausted SVM parameters searching 133