DEVELOPMENT OF FACIAL EXPRESSION RECOGNITION SYSTEM

BY

EL MEHDI BOUHABBA

A dissertation submitted in fulfilment of the requirement
for the degree of Master of Science
(Mechatronic Engineering)

Kulliyyah of Engineering
International Islamic University
Malaysia

JUNE 2011

ABSTRACT

Enabling computer systems to recognize facial expressions and infer emotions from them in real time presents a challenging research topic. The recognition of emotional information is a key step towards giving computers the ability to interact more naturally and intelligently with people. One of the potential applications of face detection and facial expression recognition is in human computer interfaces. The system will be used for the interaction between human and humanoid robot head, where the detected expression will be mimicked by the robot head. The problem of facial recognition can be divided into two major areas: detection of the face region and identification of the detected region. Detecting human face in computer vision proves to be very challenging due to the fact that human faces can have different forms and colors, adverse lighting conditions, varying angles or view points, scaling differences and different backgrounds. Attempting recognition on an inaccurate detected face region is hopeless. This thesis describes a face detection framework that is capable of processing input images swiftly while achieving high detection rates. The presented face detection system is developed using the response of Haar-Like features and AdaBoost algorithm. A set of experiments in the domain of face detection is presented in this research. The developed system yields face detection performance comparable to the best existing systems, where its accuracy is up to 98%. The face and facial features detected in the video stream are used as input to a Support Vector Machine classifier, which is used for facial expression recognition. The method was evaluated in terms of recognition accuracy for a variety of interaction and classification scenario, and it was proven that the system is able to detect the four expressions successfully. The person-dependent and person-independent experiments demonstrate the effectiveness of a support vector machine to fully automatic and unobtrusive expression recognition in real time.

خلاصة البحث

بإمكان أنظمة الحاسوب التعرف على التعابير الوجهية للإنسان العادي و استنتاج عواطفه وذلك من خلال التواصل المباشرعن طريق آلة التصوير الموصلة بالحاسوب. و نظام هـــذا التواصل هو الذي يزيد من صعوبة و تحدي هذا البحث. إن المفتاح الرئيسي الذي يمكن اعتباره قاعدة لهذا النظام هو التعرف على تعابير الوجه لإنسان ما حيث يمكن إدخال هذه النظم الى الحاسوب بعد ذلك سيتمكن هذا الأخير من التعرف تلقائيـــا علـــى التعـــابير الوجهية لكل إنسان يوجد في وضعية مقابلة مع آلة التصوير .و هذا النظـــام ســـيوظف للتعامل بين الإنسان الآلي و الإنسان العادي حيث التعابير التي تلتقطها آلة التصوير تمرر إلى الحاسوب الذي بدوره يعطي أوامر إلى الإنسان الآلي ليقوم بنفس التعابير التي قـــام بهـــا الإنسان العادي. للحصول على تعابير الوجه واجهتنا عدة مشاكل من بينها تعدد ألـــوان الوجه و كذلك الإضاءة الخارجية و اختلاف زوايا الرؤية للإنسان العادي. بعد إنجاز هذا البحث توصلنا الى نظام يمكن من كشف وجه الإنسان بسرعة جيدة ودقة عالية 98 %.
و هذا النظام كشف الوجه المتقدم و المتطور تم فيه اعتماد طريقة الهارليك Haar-Like و الخوارزمية أذابوست AdaBoost و من خلال عـــدة تحســـينات في الخوارزميـــة تم الحصول على نظام قادر على إلتقاط الوجه بجميع مواصفاته التي تميز كل إنسان عن الآخر، بعد الحصول على هذا النظام القادر على التعرف على وجه الإنسان تم تطبيق هذا النظـــام مع طريقة تثبيت المتجهة SVM للحصول على العواطف التي يشعر بها الإنسان و ذلـــك بالتركيز على حركات كل من العين و الفم للحصول على أربعة تعابير إنسانية و هـــي : الحزن و الفرح و الغضب و العادية و بعد التعرف على التعابير إنطلاقا من آلة التصـــوير الموصلة بالحاسوب سيقوم الحاسوب بإعطاء الأوامر الى الإنســـان الآلي ليقـــوم بـــنفس الحركات الوجهية (أي العواطف) إنطلاقا من عواطف الإنسان العادي.

APPROVAL PAGE

I certify that I have supervised and read this study and that in my opinion; it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Mechatronics Engineering

...……

Amir Akramin Shafie
Supervisor

……
Rini Akmeliawati
Co-Supervisor

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master Mechatronics Engineering

……
Abiodum Musa Aibinu
Internal Examiner

……
Roselina Arelhi
External Examiner

This dissertation was submitted to the Department of Mechatronics Engineering and is accepted as a fulfilment of the requirement for the degree of Master of Mechatronics Engineering.

...……

Asan Gani Abdul Muthalif
Head, Department of Mechatronics Engineering

This dissertation was submitted to the Kulliyyah of Engineering and is accepted as a fulfilment of the requirement for the degree of Master of Mechatronics Engineering.

……

Amir Akramin Shafie
Dean, Kulliyyah of Engineering

DECLARATION

I hereby declare that this dissertation is the result of my own investigation, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degree at IIUM or other institutions.

El Mehdi Bouhabba

Signature:…… Date:……

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

DEVELOPMENT OF FACIAL EXPRESSION RECOGNITION SYSTEM

I hereby affirm that International Islamic University Malaysia (IIUM) holds all rights in the copyright of this work and henceforth any reproduction or use in any form or by means whatsoever is prohibited without the written consent of IIUM. No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by means, electronic, mechanical, photocopy, recording or otherwise without prior written permission of the copyright holder.

Affirmed by El Mehdi Bouhabba

…… ……

Signature Date

ACKNOWLEDGEMENTS

In the name of Allah the most gracious and most merciful, All praise is due to Allah (S.W.T) where without whose help, this thesis would not have reached this stage. I would like to express my most sincere appreciation and gratitude to Associate Prof. Dr. Amir Akramin Shafie and Associate Prof. Dr. Rini Akmeliawati and Prof. Momoh Jimoh E. Salami for their advice, guidance, suggestions, critical comments and supervision through the course of my study and for their support in the preparation and completion of this thesis. Special thanks also go to the staff of the department of Mechatronics Engineering in particular and the Faculty of Engineering in general.

To my lovely wife Sharifah Norezila Bt Syed Yussof, I thank you for your infinite support, love and patience. For their endless encouragement, my deepest gratitude goes to my family, specially my father Muhammed Bouhabba and my mother Naima Benedaou, also to my brothers, Najib Bouhabba and Ali Bouhabba, and dearest sister Sara Bouhabba.

My sincerest thanks to the Ambassador of Morocco, His Excellency Mr Ahmed Amaziane and his deputy, Mr Said Mimoun Amenzou and all the embassy's staffs, whose help and assistance have paved way for the success of this research.

Finally, I am grateful to all my friends, Mr. Mohammed Yamlahi Alami, Mr Elhoussaine Kiai, Mr Mohamed Zekriwi, Amin Yamlahi Alami, Jamil Radih, Mohd Farid Md Alias, Moustapha Adnan, Sara Bilal, Aseef Iqbal, Adil Khambashi and to all those whose names are not mentioned, but have in one way or the other contributed to the success of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAM	Active Appearance Model
ACK	Acknowledgement
ASM	Active Shape Model
EMFACS	Emotional Facial Action System
FACS	Facial Action Coding System
FER	Facial Expression Recognition
FFT	Fourier Function Transform
HCI	Human Computer Interaction
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
OpenCV	Open Computer Vision
OS	Operating System
PCA	Principal Component Analysis
RBF	Radial Basis Function
RGB	Red Green Blue
SVM	Support Vector Machine
XML	Extensible Mark-up Language

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW

In today's society, automatic face recognition has become increasingly relevant in many commercial and law enforcement applications. The recent interest in face recognition can be attributed to the increase in the commercial interest and the development of feasible technologies to support the development of face recognition (Chellappa and Wilson, 1992; Samal and Iyengar, 1992; Valentin and Abdi ,1994). Mugshot matching, user verification, user access control, crowd surveillance, enhanced human computer interaction; all become possible when an effective face recognition system is implemented.

However, facial recognition is not possible if the face is not isolated from the background; thus the need for face detection. In this project, face detection can be simply viewed as a pre-processing step, for obtaining the "object". Face detection is a beneficial technology that has application over a wide area of fields. Even though face detection is one of the tasks which human vision can do effortlessly; detecting a human face in computer vision proves to be very challenging due to the fact that human faces can have different forms and colors, adverse lighting conditions, varying angles or view points, scaling differences (a face being near or far away) and different backgrounds. Current systems have advanced to be fairly accurate in recognition under constrained scenarios, but extrinsic imaging parameters such as pose, illumination, and facial expression still cause much difficulty in correct

recognition. The problem of facial recognition can be divided into two major areas: detection of the face region and identification of the detected region. Attempting recognition on an inaccurate detected face region is hopeless. Human Computer Interaction (HCI) could greatly be improved by using emotion, pose, and gesture recognition, all of which require face and facial feature detection and tracking (Bradski, 1998).

Although different algorithms exist to perform face detection, each has its own weaknesses and strengths. Some use skin tones, some use contours, and other are even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive (Lienhart and Maydt, 2002). An image is only a collection of colour and/or light intensity values. Analyzing these pixels for face detection is time consuming and difficult to accomplish because of the wide variations of shape and pigmentation within a human face. Pixels often require reanalysis for scaling and precision.

Viola and Jones devised an algorithm, called Haar Classifiers, to rapidly detect any object, including human faces, using AdaBoost classifier cascades that are based on Haar-like features and not pixels (Lienhart and Maydt, 2002; Menezes and Barreto, 2004; Cristinacce and Cootes, 2003). This is used for face detection and recognition in this project. Its advantages for object detection include:

i. Face detector already implemented
ii. Face detection (for videos) at 15 frames per second for 384*288 pixel images; as this thesis detects the face in real time
iii. 90% objects are detected

Another way to isolate the face is through face alignment. Face alignment aim to achieve a more accurate localization of the face by matching a 2D face shape or

appearance with a facial image. The difference between face detection and face alignment is that detection considers the images in terms of areas whereas face alignment has a precision of pixels.

Face alignment or fitting is essentially an image registration problem, where a face model needs to be deformed to match the image of a face, so that the natural facial features are aligned with the model. The dramatic variations of facial appearance due to shape, pose, illumination, expression, occlusions, and image resolution make this a challenging problem. Due to its importance in a wide range of applications, there is a sizable literature on face alignment.

Active Shape Model (ASM) and Active Appearance Model (AAM) are two most representative face alignment models. ASM (Menezes and Barreto, 2004) is one of the early approaches that attempt to fit the data with a model that can deform in ways consistent with a training set whereas AAM (Samal and Iyengar, 1992; Cristinacce and Cootes, 2003) is a popular extension of the ASM.

Facial recognition can be applied once the face is detected in the image. There are a number of different approaches to performing face recognition, which have varying levels of success. Some of the better-known algorithms utilize eigenfaces (Turk and Pentland, 1991) or active appearance models (Edwards and Taylor, 1998) to identify a face. However, eigenface approaches suffer from requiring extremely constrained frontal face images and potentially large amounts of training data to deal with high variability. Active appearance models are more promising for a noisy environment but require computationally expensive models.

In geometric feature-based methods (Goldstein and Harmon, 1971; Brunelli and Poggio, 1993) facial features such as eyes, nose, mouth, and chin are detected. Properties and relations such as areas, distances, and angles, between the features are

used as the descriptors of faces. This thesis attempt at facial recognition follows the popular use of machine learning to determine the differences between features. Support Vector Machines (SVMs) have been recently proposed by Corinna and Vapnik (1995) as an effective method for general purpose pattern recognition. By utilizing Support Vector Machines (SVM) to create models, it enables the creation of a complicated description of what features characteristics determine an expression.

1.2 PROBLEM STATEMENT AND ITS SIGNIFICANCE

Automatic face detection and recognition are two challenging problems in the domain of image processing and computer graphics that are yet to be perfected.

Manual recognition is a very complicated task when it is vital to pay attention to primary components like: face configuration, orientation, location where the face is set (relative to the body), and movement (i.e. traces out a trajectory in space).

It is more complicated to perform detection in real time. Dealing with real time capturing from a camera device, fast image processing would be needed. Haar features by Viola and Jones (2001) is the first real time frontal-view face detector. Hence, this project proposed to use this method.

An additional challenge is the need to carry out face detection and recognition in variable real world environments. This will involve the use of machine learning techniques with a large amount of training data for reliable and robust performance. To achieve this, goal of the project focuses on the camera capturing the image, system detects the face in the image and four expressions are recognised with high accuracy in various environments.

1.3 RESEARCH OBJECTIVES

The objectives of this research project are:

a) To develop real-time face and feature detection system from webcam acquired images.

b) To recognize four expressions which are Neutral, Joy, Sad and Surprise, by detecting the changes in the features.

c) To implement the Facial Expression Recognition system in the Interacting Humanoid Robot Head project done by other team, where the robot head will imitate the expressions detected from the FER.

1.4 RESEARCH METHODOLOGY

The research methodology includes:

1. Theoretical Background and Literature review of Haar-Like Features and AdaBoost for face detection and SVM using for recognizing facial expression.
2. Development of the proposed algorithm for face detection system.
3. Develop system for face detection using Matlab, C++ or C#.
4. Develop facial expression recognition system using SVM algorithm.
5. Testing the program for emotion Robot Head following Human emotion.
6. Implementation of the system to the Robot Head project, by another team.
7. Comparison and performance analysis.

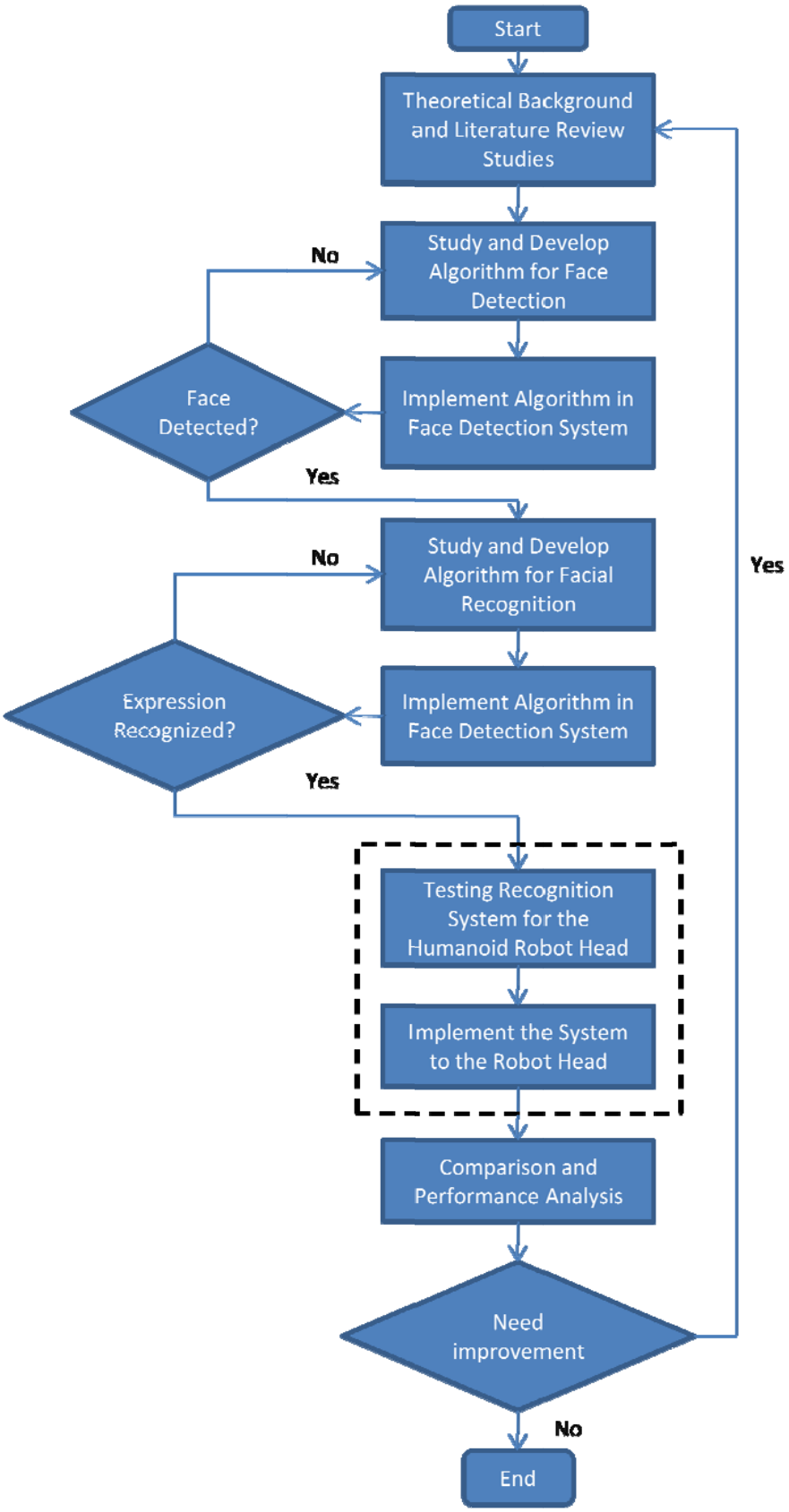

Figure 1.1: Flow Chart of Methodology

1.5 SCOPE OF RESEARCH

Reviewing the literature, it is quite difficult to state that there exists a complete system which solves face detection problem with all variations included. In this thesis, for decreasing variation in the training examples, pose variations are constrained by limiting the movement of the subject in front of the camera. The face will only be detected when it is in an upfront position and looking straight into the camera. The Haar-Classifier together with AdaBoost-Classifier will be implemented to detect and recognize the face in real time basis.

Once the face detection is done, the facial recognition of the face will be done by Support Vector Machine application. For the purpose of this project, four basic expressions will be detected by the system; NEUTRAL, HAPPY, SAD and SURPRISE. This limitation is presented in this thesis for the reason of minimizing the mechanical work in implementing these expressions to the robot head. If the robot head is successful in mimicking these four expressions, adding a few more expression is just a matter of altering and varying the algorithm.

1.6 DISSERTATION ORGANIZATION

The outline of this thesis is organized as follows:

Chapter 2 introduces the literature for the face detection research, including the research done on the Haar-Features and AdaBoost Classifier. The chapter also gives an introduction on face alignment and facial recognition. Chapter 3 is based on the face detection and facial recognition itself. This chapter shows how a fully automated face detection system is derived using Haar-Features and AdaBoost Classifier. This chapter also describes the component-based approach to face recognition and face emotion using SVM. Also outlined is the software implementation of the system

where it will present the applications of the developed Facial Expression Recognition software. Chapter 4 explains the systems tests and evaluation of the proposed systems which are implemented in this thesis. It will also contain the present experimental results. The chapter gives all the information of the system built and elaborates on how it works. Finally, conclusions and suggestions for future work that could be done to extend the capabilities are presented in Chapter 5.

CHAPTER TWO

BACKGROUND AND LITERATURE REVIEW

2.1 INTRODUCTION

The importance of facial expression in social interaction and social intelligence is widely recognized. For instance, in medicine, it could be used to continuously monitor a patient's pain level or anxiety, in gaming, a virtual avatar could be directed to mimic the user's facial expressions and in security the analysis of facial expressions could be used to assert a person's credibility. Facial expression analysis started to be an active research topic since the 19th century, when Suwa, Sugie and Fujimora (1978) introduced the first automatic facial expression recognition system. This system tracks the motion of 20 identified spots on an image sequence in an attempt to analyze facial expressions. Since then, lots of works have been done in this domain and later, various computer systems have been made to help us understand and use this natural form of human communication.

This chapter reviews what has been done in processing and understanding of facial expression. When building an FER system, these main are the issues that must be considered: face detection and alignment, image normalization, feature extraction, and classification. Implementing these steps sequentially and independently are the methods that most of the current works in FER are based on. This chapter will briefly describe the problems faced in facial expression analysis, before it review on methods existed for facial expression analysis.