FACIAL FEATURE EXTRACTION

MUHAMMAD MARZUQ BIN MOHD SHARIP

A project report submitted in partial fulfilment of the
requirements for the award of the Bachelor Degree of Electrical Engineering
(Control and Instrumentations)

Faculty of Electrical & Electronics Engineering
Universiti Malaysia Pahang

NOVEMBER 2009

DECLARATION

I declare that this project report "*Facial Feature Extraction*" is the result of my own research except for works that have been cited in the reference. The project report has not been accepted any degree and not concurrently submitted in candidature of any other degree.

Signature	:	....
Name of Author	:	MUHAMMAD MARZUQ BIN MOHD SHARIP
Date	:	12th NOVEMBER 2009

ACKNOWLEDGEMENT

Alhamdulillah, I am greatly indebted to Allah SWT on His mercy and blessing for making this project successful.

I would like to express my deepest gratitude and thanks to Mdm. Nurul Wahidah Binti Arshad, my honourable supervisor, for his continuous guidance, committed support and invaluable advice throughout my study.

I sincerely thank to my uncle and auntie, and my family for their support and bless. Not to mention, to all my friends, thank you for sharing useful idea, information and moral support during the course of study.

ABSTRACT

This project presents the facial feature extraction system and face recognition system. The test image that used for this project contain various type. There are ten different images of each of 40 disntinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions; open or closed eyes, smiling or not smiling, and facial details; glasses or no glasses. All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position with tolerance for some side movement. For the facial feature extraction system, it is more focus on eye extraction. The eye will extracted from the face by finding the centroid of the eye region using threshold technique. For the recognition system, Principle Component Analysis (PCA) is used to match the test image with the database image. The system will find which database image has a maximum percentage based on similarity of the pattern of the image.

ABSTRAK

Projek ini menengahkan tentang sistem pengekstrak ciri muka dan sistem pengecaman muka. Pelbagai jenis imej ujian di gunakan untuk projek ini. Terdapat sepuluh imej untuk setiap 40 orang subjek. Untuk sesetengah subjek, imej diambil pada waktu yang berlainan, pengcahayaan yang berlainan, ekspresi muka, buka atau tutup mata, senyum atau tidak senyum, dan ciri tambahan; berkacamata atau tidak. Kesemua gambar ini diambil mengunakan latarbelakang yang gelap yang sekata dengan subjek memandang ke hadapan dengan sedikit kepelbagaian dalam palingan ke sisi. Untuk sistem pengekstrak cirri muka, ia akan menumpukan pada pengekstrakan mata sahaja. Ciri mata akan diekstrak dari muka dengan mencari titik tengah kepada kawasan mata dengan mengunakan teknik ambang (threshold). Untuk sistem pengecaman muka, Analisis Komponen Prinsip (Principle Component Analysis-PCA) digunakan untuk mencari padanan bagi imej ujian dengan imej di dalam pangkalan data. Sistem ini akan mencari imej di dalam pangkalan data yang menpunyai peratusan tertiggi berdasarkan persamaan pola bagi imej itu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

d_{eye}	Distance of the eye
Cl, Cr	The ground truth positions
$\widetilde{Cl}, \widetilde{Cr}$	The results of automatic localization
M	The number of faces in the set
Ψ	The average face in face space
Γn	The vector of converted face image
N	The length of vector of converted face image
Φ	Face's difference from the average face in face space
C	Covariance matrix for dataset
p, p_i	Pixel of face n
A	Matrix of each face's difference from the average face in face space
L	Value of eigenvectors of M x M matrix
u_i	eigenvector of L

LIST OF ABBREVIATIONS

PCA	Principal Component Analysis
2DPCA	Two-Dimensional Principal Component Analysis
GUI	Graphical User Interface
AT & T	American Telephone & Telegraph Corporation
LED	Light Emitting Diode
IBM	International Business Machines Corporation
HMM	Hidden Markov Model

CHAPTER 1

INTRODUCTION

1.1 Introduction of Facial Feature Extraction and Recognition

Facial feature extraction consists in localizing the most characteristic face components such as eyes, nose, and mouth within images that depict human faces. This step is essential for the initialization of many face processing techniques like face tracking, facial expression recognition or face recognition. Among these, face recognition is a lively research area where it has been made a great effort in the last years to design and compare different techniques.

In this chapter, an automatic method for facial feature extraction that used for the initialization of our face recognition technique presented. Basiclly, to extract the facial components equals to locate certain characteristic points. For example, locate the center and the corners of the eyes, the nose tip, and others. Particular emphasis will be given to the localization of the most representative facial features, namely the eyes, and the locations of the other features will be derived from them.

A general statement of the automatic face recognition problem can be formulated as follows: given a stored database of face representations, one has to identify subjects represented in input probes. This definition can then be specialized to describe either the identification or the verification problem. The former requires as input a face image, and the system determines the subject identity on the basis of the database of known individuals; in the latter situation the system has to confirm or reject the identity claimed by the subject.

As noted by Zhao et al.[1], whatever the problem formulation, its solution requires the accomplishment of three subsequent subtasks: face detection, feature extraction and face recognition as shown in Figure 1.1.

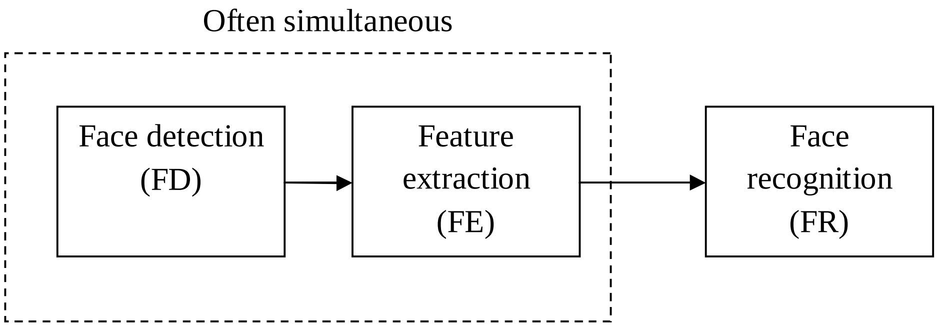

Figure 1.1: The subtasks of the face recognition problem

1.2 Eigenface Method

The typical approach based on the idea of template pattern matching is the "eigenface" method. To generate a set of eigenfaces, a large set of digitized images of human faces, taken under the same lighting conditions, are normalized to line up the eyes and mouths. They are then all resampled at the same pixel resolution. Eigenfaces can be extracted out of the image data by means of a mathematical tool called principal component analysis (PCA).

The eigenfaces that are created will appear as light and dark areas that are arranged in a specific pattern. This pattern is how different features of a face are singled out to be evaluated and scored. There will be a pattern to evaluate symmetry, if there is any style of facial hair, where the hairline is, or evaluate the size of the nose or mouth. Other eigenfaces have patterns that are less simple to identify, and the image of the eigenface may look very little like a face.

1.3 Background of the Problems

Though people are good at face identification, recognizing human face automatically by computer is very difficult. Face recognition has been widely applied in security system, credit-card verification, and criminal identifications, teleconference and so on. Face recognition is influenced by many complications, such as the differences of facial expression, the light directions of imaging, and the variety of posture, size and angle. Even to the same people, the images taken in different surroundings may be unlike. The problem is so complicated that the achievement in the field of automatic face recognition by computer is not as satisfied as the fingerprints. Facial feature extraction has become an important issue in automatic recognition of human faces. Detecting the basic feature as eyes, nose and mouth exactly is necessary for most face recognition methods.

In fact, given an input image depicting one or more human subjects, the problem of evaluating their identity boils down to detecting their faces, extracting the relevant information needed for their description, and finally devising a matching algorithm to compare different descriptions.

On one hand, the modularity of the original problem is a beneficial property as it allows to decompose it and to concentrate on the specific difficulties of each task in order to achieve a more effective solution. On the other hand, care must be taken in recomposing the separate modules: a common approach is to devise techniques that face only a task at once without considering the problems that can arise at the "interfaces" between them.

An important aspect of any localization algorithm is its precision. The face recognition techniques presented in literature only occasionally face the issue and rarely state the assumptions they made on their initialization, many simply skip the feature extraction step, and assume perfect localization by relying upon manual annotations of the facial feature positions. However, it has been demonstrated that face recognition heavily suffers from an imprecise localization of the face components.

This is the reason why it is fundamental to achieve an automatic, robust and precise extraction of the desired features prior to any further processing. In this respect, we investigate the behaviour of the face recognition technique when initialized on the real output of the extraction method.

1.3.1 How to Select the Best Facial Feature Point

In order to perceive and recognize face, what to do is extract the prominent characteristics on the faces. Usually those features like eyes, nose and mouth together with their geometry distribution and the shape of face is applied.

Applying human visual property in the recognition of faces, people can identify face from very far distance, even the details are vague. That means the symmetry characteristic is enough to be recognized. Human face is made up of eyes, nose, mouth and chin etc. There are differences in shape, size and structure of those organs, so the faces are differ in thousands ways, and can describe them with the shape and structure of the organs so as to recognize them. One common method is to extract the shape of the eyes, nose, mouth and chin, and then distinguish the faces by distance and scale of those organs as shown in Figure 1.2. The other method is to use deformable model to describe the shape of the organs on face finely [2].

Figure 1.2: Relationship of face features.

Characteristics of the organs easily to determine by locating the feature points from a face image. If the characteristics is normalized which have the properties of scale, translation and rotation invariance, the faces in the database also can be normalized through pre-treatment, so as to extend the range of database, reduce the storage and recognize the faces more effectively.

Additionally, the selection of face feature points is crucial to the face recognition. The feature points should be picking up which represent the most important characteristics on the face and can be extracted easily. The number of the feature points should take enough information and not be too many. If the database has different postures of each people to be recognized, the property of angle invariance of the geometry characteristic is very important. Most extraction technique has presented a method to locate the vital feature points of face, which select 9 feature points that have the property of angle invariance, including 2 eyeballs, 4 near and far corners of eyes, the midpoint of nostrils and 2 mouth corners, as shown in Figure 1.3. According to these, we can get other feature points extended by them and the characteristics of face organs which are related and useful to face recognition.

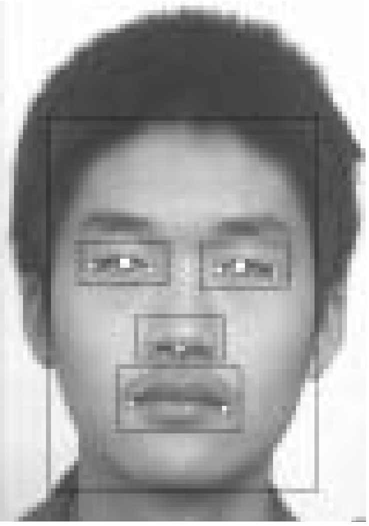

Figure 1.3: The 9 vital feature points on face.

1.4 Objective of the Project

Objective of this project are:

i. To build a system that can recognize the eyes on human face.

- The knowledge of the eye positions allows to roughly identify the face scale (the distance is relatively constant from subject to subject) and it is in plane rotation.

ii. To compare the extracted feature with facial images database.

- From the eyes extracted, that feature required for the initialization of any face recognition technique.

iii. To declare a matches of the same facial features.

- The accurate eye localization permits to identify all the other facial features of interest.

1.5 Scope of the Project

Scopes of this project are:

i. The recognition is based on the location of the pupil of the eyes;

- Eyes are the only facial features required for the initialization of any face recognition system, actually this is the only information needed by those that operate an alignment of the face region. However some techniques may require more features than just the eyes.
- The system will be find which database image has a maximum percentage of similarity of the location and the distance of the eyes.

ii. Facial database from American Telephone & Telegraph (AT&T) laboratories, Cambridge;

- This database contains 400 images consist of ten different images of each 40 distinct subjects.
- For some subject, the images were taken at different times, varying the lighting, facial expressions, open or close eyes, smiling or not smiling, glasses or no glasses.
- All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position with tolerance for some side movement.

iii. Using Graphical User Interface (GUI) in MATLAB.

- The suitable software to build a Graphical User Interface (GUI) that is user-friendly to everyone with a simple button to operate.

CHAPTER 2

LITERATURE REVIEW

2.1 Face Detection

The image acquisition technique is critical in analyzing facial expressions. The image resolution, camera type, digitizer used, size of image, image illumination all influence the facial expression analysis. Pantic and Rothkrantz [3] first used two cameras mounted on a headphone like, one capturing the frontal portion of the face and one on the right hand side capturing the profile image of the face. Image resolution of 69 x 93 is optimum for any automated face processing related techniques. At a resolution of 48 x 64, corners of the mouth and eye are hard to detect. Facial expression cannot be recognized at an image resolution of 24 x 32 [4]. Face acquisition deals with detecting the exact location of face, position of different features on the face and orientation of the face either from an arbitrary image or from a facial image. Usually the facial image in the frontal or near frontal view is used to recognize facial expressions. Thus in order to get a properly registered facial image, two approaches are used. In the holistic approach, the face is identified as a whole unit. In the analytic approach, certain features of the face are identified individually and then recombined to get the overall face boundary. Huang and Huang [5] used a point distribution model and canny edge detector based approach to locate the frontal view of the face in an image. Pantic and Rothkrantz [3] used image histogram thresholding methods to detect the horizontal and vertical boundaries of the face in dual views. Kobayashi and Hara [6] used brightness distribution techniques on monochrome facial images to extract the position of the irises, and then eventually the entire location of the face. Yoneyama [7] used an analytic approach to detect the outer corner of the eye, height of the eye and mouth automatically. PersonSpotter [8] system performs real time tracking of the head in arbitrary images. Pentland et al., [9] used Principal Component Analysis based eigenspace method to track the head in real time. Heisele et al., [10] developed component based, trainable system for detecting frontal and near frontal views of faces in still gray images. Rowley et al., [11] developed a neural network based system to detect frontal views of face.

2.2 Technique of the Facial Extraction

Recently, techniques achieved in the researches for detection of facial feature points can be broadly classified as: (i) approaches based on luminance, chrominance, facial geometry and symmetry [12, 13], (ii) template matching based approaches [13, 14], (iii) PCAbased approaches[12, 15, 16], and the combination of the above approaches along with curvature analysis of the intensity surface of the face images[17]. Also other facial feature detection approaches exist. Feris et al. present a technique for facial feature localization using a two-level hierarchical wavelet network [18].

Regarding feature extraction, there is a general agreement that eyes are the most important facial features, thus a great research effort has been devoted to their detection and localization [19, 20, 21, 22, 23, 24, 25, 26]. This is due to several reasons, among which:

- Eyes are a crucial sourse of information about the state of human beings.
- The eye appearance is less variant to certain typical face changes. For instance they are unaffected by the presence of facial hair (like beard or mustaches), and are little altered by small in-depth rotations and by transparent spectacles.
- The knowledge of the eye positions allows to roughly identify the face scale (the inter-ocular distance is relatively constant from subject to subject) and its in plane rotation.

- The accurate eye localization permits to identify all the other facial features of interest.

2.3 Facial Feature Data

Two types of features can be extracted from the facial image: geometric features and appearance features. Geometric feature measures the variations in shape, location, distance of facial components like mouth, eyes, eyebrows, and nose in different expressions. The appearance features present the appearance (skin texture) variations of the face, such as wrinkles and furrows. The appearance feature can be extracted on either the whole face or specific regions in a facial image. To recognize facial expressions, an automatic facial expression analysis system can use geometrical features only [3, 27], appearance features only [28, 29, 30], or hybrid features (both geometric and appearance based) [31, 32, 33]. Cohn, Kanade and Tian [32, 33] have obtained normalized feature measurements by using a reference neutral face image to remove the variations in face style, motion, and other factors.

2.4 Pupil Detection

The pupil detection system detects the pupils using the redeye effect. The system's robustness to occlusions and head motions makes it ideal to be used for automatic facial action analysis. As the pupil positions can be recovered very efficiently and robustly, it eliminates the need of manual labelling or pre-processing of the images, a required step that plagues a number of previous approaches.

Although the red-eye effect has been known for quite sometime, it is in recent years that it has grabbed a lot of attention for vision applications. Morimoto et al [34] have described a system to detect and track pupils using the redeye effect. Haro et al [35] have extended this system to detect and track the pupils using a Kalman filter and probabilistic PCA. Using an infrared camera equipped with infrared LEDs, which is used to highlight and track pupils and is an in-house built version of the IBM Blue Eyes camera. The whole unit is placed under the monitor pointing towards the users face. The system has an infrared sensitive camera coupled with two concentric rings of infrared LEDs. One set of LEDs is on the optical axis and produces the red-eye effect. The other set of LEDs, which are off axis, keeps the scene at about the same illumination. The two sets of LEDs are synchronized with the camera and are switched on and off to generate two interlaced images for a single frame. The image where the on-axis LEDs are on has white pupils whereas the image where the off-axis LEDs are on has black pupils. These two images are subtracted to get a difference image, which is used to track the pupils.

The pupils are detected and tracked using the difference image, which is noisy due to the interlacing and motion artifacts. Also, objects like glasses and earrings can show up as bright spots in the difference image due to their specularity. To remove this noise, threshold the difference image using an adaptive thresholding algorithm [35]. First, the algorithm computes the histogram and then thresholds the image keeping only 0.1% of the brightest pixels. All the non-zero pixels in the resulting image are set to 255 (maximum value). The thresholded image is used to detect and to track the pupils.

2.5 The Importance of Precise Eye Localization

Given the true positions of the eye centers (by manual annotation), the eye localization accuracy is expressed as a statistics of the error distribution made over each eye (usually the mean or the maximum), measured as the Euclidean pixel distance. In order to make these statistics meaningful, so that they can be used to compare the results obtained on any dataset, it is necessary to standardize the error by normalizing it over the face scale.

One popular error measure has been introduced by Jesorsky et al. [36], and it has been already adopted by many research works on eye localization. The measure, which can be considered a worst case analysis, is defined as