UNIVERSITI TEKNOLOGI MARA

FACE EXPRESSION RECOGNITION USING ARTIFICIAL NEURAL NETWORK (ANN)

A Project paper submitted to the
MARA University of Technology
In partial fulfillment of the requirement for the
BACHELOR OF SCIENCE (Hons) INTELLIGENT SYSTEMS

By

MAZURAINI BINTI GHANI
2002326490
BSc (Hons) INTELLIGENT SYSTEMS
FACULTY OF INFORMATION TECHNOLOGY
AND QUANTITATIVE SCIENCES

APRIL 2005

DECLARATION

I hereby declare that this research report together with all of its contents is no other than those of my own work, except for some information taken and extracted from other sources that have been quoted respectively.

.....

11th MARCH 2004

MAZURAINI BINTI GHANI

(2002326490)

ACKNOWLEDGEMENT

In the name of ALLAH, the Most Gracious and the Most Merciful…

Alhamdulillah, praise Allah for His Almighty and Graciousness. With his utmost blessings, I was able to finish this written report within the time duration given.

Firstly, my appreciate goes to Assoc. Prof. Puan Zaidah Ibrahim who has effortlessly enriched me with ideas and opinion in order for me to complete this study successfully. I really appreciate her unstoppable suggestions, guidance and insight throughout this research. The most important things are her ability to encourage and motivate me to do the research within a very limited time. Special thanks to Associate Prof. Dr.Azlinah for your guideline, motivate and support in conducting this research. With her expertise, I was able to conduct this study within the domain required.

I would like to express my gratitude for who was giving unlimited support, understanding, and love me constantly; they are my parent and siblings. I really appreciate that and it will remain in my heart for the rest of my life.

I would also convey my gratitude to all my friends who contribute in my research as a face model, gave motivation and an idea. They are my course mate CS223 and my housemate. From the bottom in my heart I would like to say thank you very much to all of you. May Allah bless you.

ABSTRACT

Over a few decades, many computer vision systems haven been developed. One of the applications related to computer vision is face recognition and was being interested by many researches. This project is all about implementing the back-propagation neural network algorithm in classification of face expression. This project has 3 objectives. The first objective is to collect and digitized images with different expressions which is neutral, happy and angry. The second is to design and develop a prototype for classifying human emotions by face expression recognition of a given image, using back propagation neural network. The last is to study and experiment the suitable edge detection techniques for binary facial image. There are two important phases that were focused in developing this project. The phases are pre-processing phase and neural network design phase. In pre-processing phase, s detail studies and intensive experiments were conducted to obtain a suitable method of segmentation. Meanwhile, in the neural network design and implementation phase, intensive experiments have been conducted to obtained appropriate design and parameter value of neural network. In this project, the suitable method of segmentation is local adaptive threshold. However, the performances of neural network in learning and classification task should be enhanced by redesigning and conducting experiment on other learning algorithm than back-propagation.

TABLE CONTENTS

CHAPTER 3: METHODHOLOGY

CHAPTER 4: ANALYSIS AND RESULT

CHAPTER 5: CONCLUSION

APPENIDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

BACKGROUD OF THE PROJECT

1.1 Introduction

Human beings possess and express emotions in everyday interactions with others. Emotions are often reflected on the face, in hand and body gestures, and in the voice, to express our feeling. Facial expressions and vocal emotions are commonly use in everyday human-to-human communication, as one smile to show greeting, end frowns when confused. People do great deal of inference from perceived facial expression.

In recent years, an exploration in research in pattern recognition system using Artificial Neural Network method has been observe because it is a successful method in pattern recognition [11]. Examples of pattern recognition are handwritten recognition, fingerprints recognition, speaker recognition and voice recognition. Face expression classification is also one of the pattern recognition.

Face expression classification is a specialized pattern recognition task with several applications. There are many types of face expression. Some examples of face expression are neutral, sad, angry, happy, nervous, shock, laugh and scared. Because of the various types of face expression many researcher has develop system on face expression classification.

The recognition of face expression has been focus by a large number of researchers over the past several decades. There are many approaches successfully used for face recognition system. Every approach has own special purpose depending on what the system is used for and how the image is captured.

1.2 Problem Description

Emotion detection using face expression recognition is a complex task. There are several problems occurred during an image processing stage and face expression recognition process [6]:

1. Scale variance

   Each of facial images is captured at different scales or sizes due to the distant between the face and the image acquisition device such as camera.

2. Pose variance

   Pose variance also called shift variance can be described as the different orientation of each facial image input to system.

3. Illumination variance

   Illumination can be described as the amount of source light occurrence on the picture being viewed and the amount of light reflected by the objects in the picture. The major cause of illumination variance is the lighting condition when the picture is taken. Lighting condition will determine the brightness and clearness of an image.

4. Detail variance

   Facial features on the images may be covered by some details such as headscarf, breads, hairs and moustaches.

5. Facial expression

   Difficult to capture real face expression image, because face expression will clearly shown based on the feeling at that time.

Problems that stated at the first and second point can be overcome manually using the image editing software such as Adobe Photoshop 7.0, and the process called image normalization.

1.3 Project Objectives

The main objectives to the development of this project are:

1. To collect and digitized images with different expressions which is neutral, happy and angry.

2. To design and develop a prototype for classifying human emotions by face expression recognition of a given image, using back propagation neural network.

3. To study and experiment the suitable edge detection techniques for binary facial image.

1.4 Project Scope

The following is scope of the project:

1. Classification is based on the three main emotions (neutral, happy, angry) only.
2. Images are captured from young male and female, the age of individuals is around 20 years old to 25 years old.
3. Each an individual's face captured with three different expressions image (neutral, happy and angry).
4. The image of an individual's face captured does not were any type of glasses for both male and female.

5. The facial images captured and to be processed are in frontal view only.
6. The images to be processed are in grayscale mode.
7. Implement Neural Network by using back-propagation algorithm for training and learning task.
8. The proposed project uses six facial point feature:
    a) Eyebrow inner startpoint (2)
    b) Eyebrow inner endpoint (2)
    c) Corner of the mouth (2)

1.5 Project Significance

Many computer vision systems have been developed, such as finger print recognition, speech recognition and other pattern recognition. Face recognition is one of the applications related to computer vision. All these computer visions have own benefits and significances. Therefore the benefits and significance of this proposed project are:

1. Provide a better understanding on the process involved in typical computer vision system especially on the image processing stage and classification stage.

2. Defined suitable edge detection techniques for binary facial image.

3. Identify the appropriate values of parameter in neural network to solve the problem of this project.

4. To improve the perspective of people about the importance of computational intelligence approaches in solving the complex problem such as face recognition.

1.6 Summary

This chapter provided the background of the project to be developed. Introductory of this project raptly described here. The problems and objectives of project have been identified. The scope of the project have been determined and stated based on the limitation and the time constraint. The significance of the project has also stated in this chapter based on an objectives, to ensure this project does not waste time and effort.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter discusses the principals and framework commonly used in developing computer vision application. Computer vision applications usually involves pattern recognition covering from simple object recognition to more complex object such as speech recognition and face recognition.

The beginning of this chapter discusses basic principals in development of computer vision systems. After that, the discussion focuses on the digital image processing steps especially in face recognition problem. Then, the pre-processing steps of digital image especially used in facial image processing are discussed. The final section in this chapter will discuss Artificial Neural Network approaches in recognition and classification tasks.

All the technical and important terms will be explained clearly to help in understanding insight on the subject.

2.2 Fundamental steps in digital image processing

In this section, the fundamental steps required in digital image processing are discussed. The basic steps involve digital image processing is depicted Figure 2.1

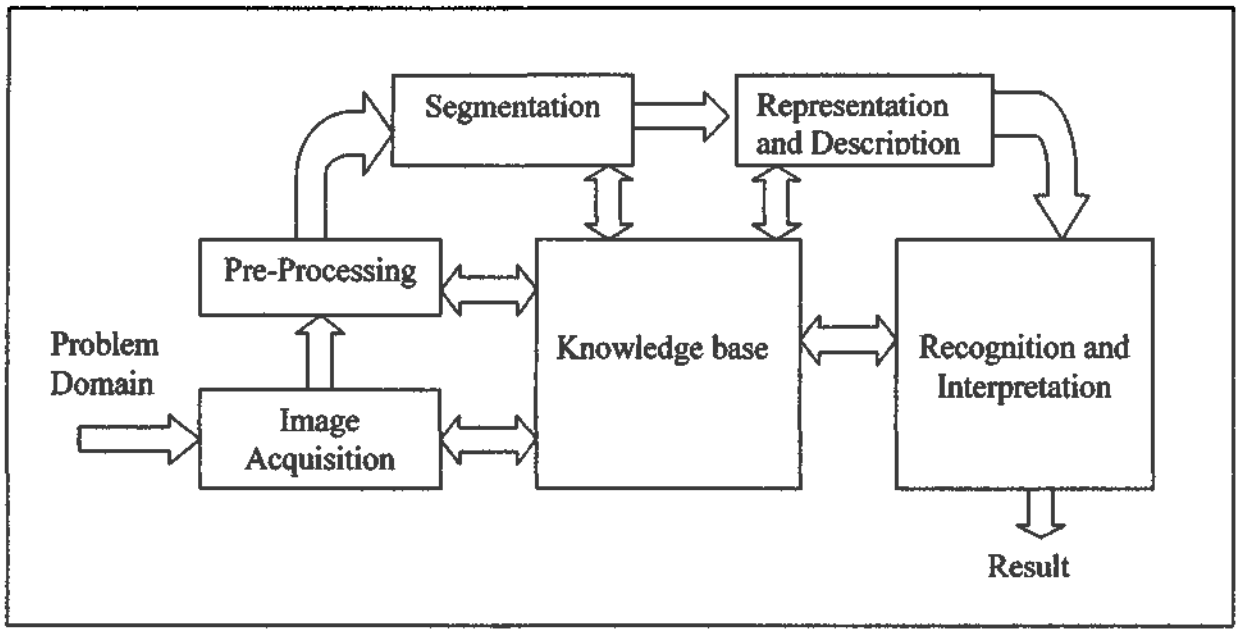

Figure 2.1 Fundamental steps in digital image processing (Rafael C. G, Richard E.W, 1992).

2.2.1 Image Acquisition

Image acquisition is the first step in digital image processing. In a simple word, this is stage where a digital image is acquired. In other words, the image to be processed is obtained and presented to the computer vision system [7]. The data image may be any image files that come from several different sources such as [12] monochrome or color TV camera, frame grabber and scanner.

2.2.2 Pre-processing

The purpose of image pre-processing is to improve the image quality hence ensures that the image can be process at the other steps or stage more accurately [12]. For example, pre-processing normally relates with techniques for contrast enhancements, noise filtering and edge detection.

There are several important tasks in image pre-processing to be implemented in computer vision system. The tasks involved are image normalization, histogram equalization, media filtering, high-pass filtering, background removal, translational and rational normalizations, and illumination normalization [7]. The detail of each task is discussed as follows:

1. Image size normalization

The acquired image is usually altered and edited to be in the form of default format. Each of the images must be in the same size, same type of file, same mode either grayscale mode or color mode, and same resolution on which the computer vision system runs.

2. Histogram equalization

To enhance the quality of image and improving the face recognition performance it is necessary to modify the dynamic range (contrast range) of the image. Consequently, the desired facial features will be more apparent.

3. Median filtering

Median filtering one of the method to reduce the noise information that exist in an image the method is more effective when it deals with strong noise patterns and to preserve the sharpness of edge.

4. High-pass filtering

High-pass filtering stresses and deals with desired features of and image such as the contour of eye s and mouth. High pass filtering can improve edge detection performance significantly. This is beneficial when the system is emphasized on certain important features on the image.

5. Background removal

This stage remove unnecessary background in the image to ensure only the facial component will be used.

6. Translational and rational normalizations

In certain situation, it is possible to deal on face image in which the head is somehow shifted or rotated. It is necessary in a face recognition system to normalize the orientation of the head position of facial image. In face recognition, especially the systems based on the frontal views, the head is the key factor in the determination of facial features.

7. Illumination normalization

The image capture under different lighting condition (illumination) can affect the classification performance especially for face recognition systems that use the whole face information for recognition. For example, a fair human face can be dark in captured image.

2.2.3 Segmentation

After the image has been normalized, enhanced and filtered image are ready for further step in digital image processing called segmentation. In general, segmentation divides and input image into its elementary parts or objects [12]. In the problem of face recognition, segmentation is a process of extraction the individual facial features such as eyes, nostril and mouth. One the most important approach in segmentation is thresholding [12]. Thresholding is the process of conversion grayscale image into binary image.

Thresholding produce only two value either foreground value or background value by setting all pixels whose intensity values above a threshold value to a foreground

and all the remaining pixels to background value. The foreground value can be 1's and background value can be 0,s or vice versa.

The foreground value referred to objects and background values referred to background.

$$g(x, y) = \begin{cases} 1, \text{ if } f(x, y) > T \\ 0, \text{ if } f(x, y) \leq T \end{cases}$$

Where T is a threshold value, f(x, y) is pixel value at point (x, y) and g(x, y) is a binary image as a result of thresholding process.

Basically, thresholding can be categorized into three types. The types that are usually applied are [12]:

1. Global Thresholding

Global threshold is when the threshold value, T is manipulated only on single pixel value *f(x, y)* only.

2. Local Thresholding

Local thresholding compute threshold value, T based on some local properties called *p(x, y)* of *f(x, y)*.

3. Local Adaptive Thresholding

Local Adaptive Threshold is when threshold value depends on both local properties *p(x, y)* and spatial coordinates x and y.

2.2.4 Representation and Description

Segmentation process produces raw data in the structure of pixels that represents a boundary and pixels contain in the region. Representation is process where conversion pf raw data into the proper form for next process which is recognition and interpretation.

The data yield from segmentation process can be represented as a boundary or as whole region. Boundary representation is suitable when emphasizes on external shape, while regional representation is emphasized on internal properties such as texture and intensity [12].

There are several representations scheme that are traditionally used the approaches to represent the region of interest are chain codes, polygonal approximations, signatures, boundary segments and skeleton of a region [12].

Description is the process of extracting important features and differentiating the region of interest in an image from other region. In face recognition problem, the region of interest can be the eyes or the mouth. These two regions must be differentiated among each other. There are two types of descriptors called boundary descriptor and regional descriptors.

Some simple descriptors in boundary descriptors are the length contour of particular region of interest, shape number which depends on the chain code representation of particular curve, Fourier descriptors and moments [12].

While for regional descriptor descriptors the approaches typically used are the area of a region, topological descriptor and texture. Area can be described as the number of pixels consists inside the particular boundary. Meanwhile, topological descriptors mean a region of interest or objects described by its topological properties. Example of the topological descriptor is shown in Figure 2.2

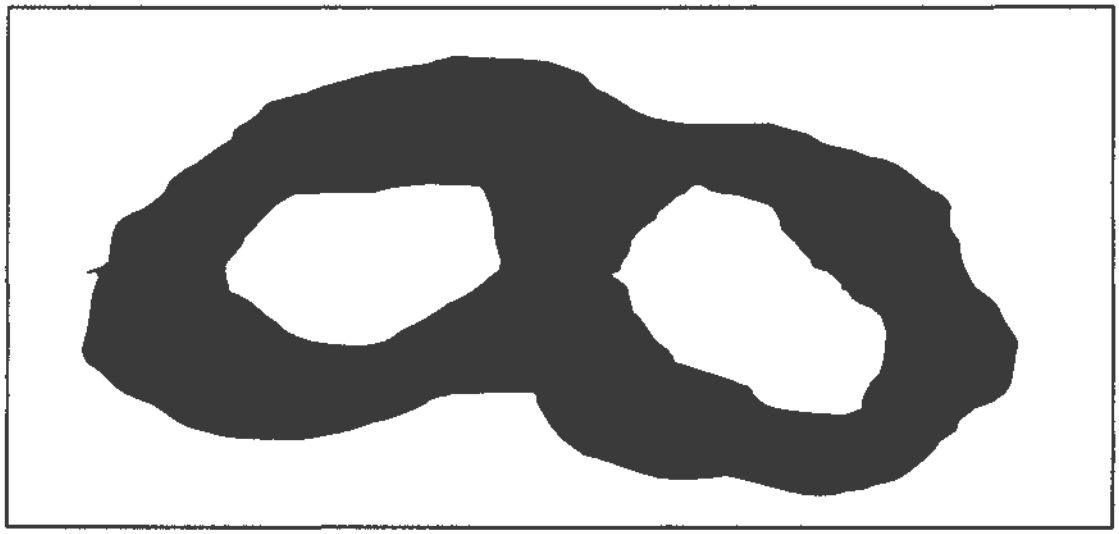

Figure 2.2 Example of topological descriptor (Negnevitsky, M., 2002)

Figure 2.2 show the region has two holes and topological descriptor is described as the number of holes consists in the region. In other way, topology can be defined as the analysis of properties of figure that are unchanged by any deformation such as stretching or change in orientation.

Another regional descriptor is based on texture content. Texture descriptor can be defined as the measure of properties such as smoothness, coarseness and regularity of the image being analyzed.

2.3 Artificial Neural Network

This section will discuss how ANN is inspired by the characteristics and function of human brain, the learning approaches in ANN and how the neural network learn by examples of training data.

2.3.1 Recognition and Interpretation

This is the stage where the testing objects are classified into a similar group and labeling the objects based on the information provide by the descriptors. At this stage, the system being developed must have the following characteristics:

1. The ability to extract relevant information from necessary background and details.
2. Make generalization from the knowledge generated by the learning from training data and example.
3. The capability to make classification from imperfect of noisy information.

The major factor in recognition or classification of the image into particular group is the similarity of features possessed by each object of a group. There are several methods of recognition and classification including Nearest Neighbor, K-Nearest Neighbor, Maximum-Likelihood and Neural Network [6].

Many researchers and studies categorized the recognition can be divided into to types which are supervised recognition and unsupervised recognition [7][9][11].

2.3.1.1 Supervised Recognition

In supervised recognition, the real output is known. The actual output is obtained based on a priori knowledge about the patterns of input data [11]. Beside that, the other factor is the number of output class is available. Supervised recognition consists of two phases which are learning or training phase and testing or validation phase.

1. Training or Learning Phase

   In this stage, while a set of training data are being input into the classifier, the internal properties of the classifier are also being adjusted. For example in neural network, the desired output is compared with actual output produces by the network. The error gradient is calculated to adjust the weight of each neuron in the network. [9].

2. Testing and Validation Phase

   Testing is the stage where, the trained network used with the final internal properties of the classifier that was contribute after a several repetitive process of adjustment, to test mutually exclusive testing data. The performance of the network and classifier where measured by number of correctness of classifier classified the data.

2.3.1.2 Unsupervised Recognition

Differentiation between supervised recognition and unsupervised recognition is target output classes. The real target output is unknown, so that unsupervised recognition more difficult compared to supervised recognition. This recognition is suitable when a set of training patterns of unknown class is available. This type of recognition will try to learn to categorize the input pattern into finite and unknown numbers of output class. Most of the unsupervised recognition is based on clustering algorithm and operations. The tow main categories of cluster algorithm are [11]:

1. Agglomerative algorithm

   Star taking into account every point in feature space in feature space as a single cluster and group them using some similarity function.

2. Divisive algorithm

   This algorithm assumes the set of features as a single cluster and will be divided until a predefined condition is achieved.

Beside cluster algorithm, other popular algorithm in unsupervised learning is Hebb Rule as discussed in [9].

2.3.2 Characteristic of brain function

Neural networks are named after the cells in the human brain that perform intelligent operations. The brain is made up of billions of neuron cells. Neural networks are formed from hundreds or thousands of simulated neurons connected together in much the same way as the brain's neurons.

Neural network is a system that modeled the human brain. It is a statistical model built by tuning a set of parameter known weights. The weights form a mapping from a set of input value to the associated of output values [5]. Figure 2.3 below illustrate the typical component of a human brain.