USER VERIFICATION FOR SECURITY APPLICATION

Muhammad Akmal Bin Hanzah

May 2009

"I hereby declare that I have read through this report and found that it has comply the partial fulfillment for awarding the degree of Bachelor of Electrical Engineering (Control, Instrumentation and Automation)

Signature :……………………………………….
Supervisor"s Name : Encik Ahmad Idil Bin Abdul Rahman
Date : 11 MAY 2009

USER VERIFICATION FOR SECURITY APPLICATION

MUHAMMAD AKMAL BIN HANZAH

This report is submitted in partial fulfillment of requirement for the degree of Bachelor in Electrical Engineering
(Control, Instrumentation and Automation)

Faculty of Electrical Engineering
Universiti Teknikal Malaysia Melaka (UTeM)

11 MAY 2009

“I hereby declare that this report is a result of my own work except for the excerpts that have been cited clearly in the references.”

Signature:…………………………………
Name : Muhammad Akmal Bin Hanzah
Date : 11 MAY 2009

ACKNOWLEDGEMENT

First and foremost, I thank Allah the Almighty for blessing me to complete my Project Sarjana Muda 2. I want to take this opportunity to record my utmost and sincere gratitude to my supervisor, En. Ahmad Idil Bin Abdul Rahman. Without him, I can never start work on my project and to proceed until this point of development. He has shown me guidance, important advice, and inspiration throughout my project. He has also given me knowledge essential in doing this project.

Besides, I would like to show my appreciation to my lectures, who have taught me over the years in UTeM. The have taught me the basic of Electrical Engineering, and this invaluable knowledge has provided ma a firm foundation for doing this project. Most importantly, the knowledge I required from them has prepared me for my career in the future.

Furthermore, I would like to thank my friends and fellow classmates for sharing and discussing, knowledge with me. Their support, opinion, and advised will be not forgotten.

To my beloved family, I would like to forward my obliged to them for their continuous support during my study period, their patience and benevolence. Lastly, I would like to thank everyone who has contributed during my Projek Sarjana Muda 2. your kindness and cooperation of my paperwork is much appreciated.

ABSTRACT

In daily life, human needs to communicate and change information among each other. Speech recognition has been an important subject for research, and its development has come to a stage where it has been actively and successfully applied in many industrial and consumer applications. The method used for speech recognition have since been developed and improved, with increasing accuracy and efficiency leading towards a better human-machine interface. This project is about to design and to develop a User Identification for Security Application via MATLAB programming language. User verification for security application is a project to recognize user or speaker by using their voice to apply on security system. The system importance for security purposes such as access rooms, phone banking, robots, military etc. Personal identity is usually claimed by presenting a unique personal possession such as a key, a badge, or a password.

ABSTRAK

Dalam kehidupan seharian manusia perlu berhubung dan bertukar maklumat sesama sendiri. Pengesanan suara telah menjadi subjek utama untuk kajian dan kajian tersebut telah sampai ke tahap ia telah digunakan di dalam industri dan juga aplikasi pengguna. Kaedah yang digunakan untuk pengesanan suara telah dibangunkan dan dinaiktarafkan sejajar dengan ketepatan dan kecekapan dalam memajukan antaramuka mesin-manusia. Projek ini bertujuan untuk merekabentuk dan membangunkan sebuah Aplikasi Pengenalpastian/Pengecaman Pengguna untuk Keselamatan menggunakan MATLAB. Aplikasi ini bertujuan untuk mengenalpasti pengguna melalui suara pengguna dalam sesebuah system keselamatan. Sistem amat berguna untuk tujuan keselamatan seperti akses kepada sesebuah bilik, urusan bank melalui talian, robot, ketenteraan dan sebagainya. Pengenalan diri selalunya diakui dengan menyerahkan atau menunjukkan hakmilik peribadi yang unik contohnya kunci, lencana atau kata laluan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

LIST OF TABLE

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The human ear is a marvelous organ. Beyond our uniquely human ability to receive and decode spoken language, the ear supplies us with the ability to perform many diverse functions. These include, for example, localization of objects, enjoyment of music, and the identification of people by their voices. Currently, along with the efforts to develop computer procedures that understand spoken messages, there are also considerable interests in developing procedures that identity people form their voices.

User verification for security application is a project to recognized user or speaker by using their voice to apply on security system. The system importance for security purpose such as access rooms, phone banking, robots, military etc. Personal identity is usually claimed by presenting a unique personal possession such as a key, a budge, or a password.

Now days, there are several techniques to identify or verify person such as thumb print, retina, DeoxyriboNucleic Acid (DNA) and voice. As the technology moves forward, the technique to identify user also comes with alternative by using voice to verify user identity.

Among these methods, identity verification based on a person‟s voice has special advantages for practical deployment. Speech our most natural means of communication and therefore, user acceptance of the system would be very high. Advances in digital signal processors and speech technology have made possible the design of fast, cost effective, high performance speaker verification systems. These system can be integrated into providing access control for banking transactions by telephone, automatic telephone transactions by mail and credit card verification, and access to computer via modems on dial-up lines.

One of the advantages of using voice is that it leads to faster time for solution of problems. In Malaysia, speaker verification is still in its infancy stage. Speaker verification is a process where an unknown word is given as an input to a system, and the system would recognize the word and perform an output action accordingly. Thus the comparison of different algorithm has been carried out for years in various R&D laboratories in the world. Throughout the development of voice recognition different methods have been produced.

1.2 OBJECTIVE OF THE PROJECT

- To design a user / speaker verification system using speech recognition system.
- To learn about the technique use for the automated speech recognition system such as Fast Fourier Transform (FFT) and Dynamic Time Warping (DTW).
- To implement the algorithm of the systems and develop the interface by using MATLAB programming language.

1.3 SCOPE OF PROJECT

The scope of this project involve with two different module, training module and testing module. For the training module, the system will be train by the known voice to get the data and store into the reference template. For the testing module, the system will be test with unknown voice and known voice.

1.3.1 Training Method

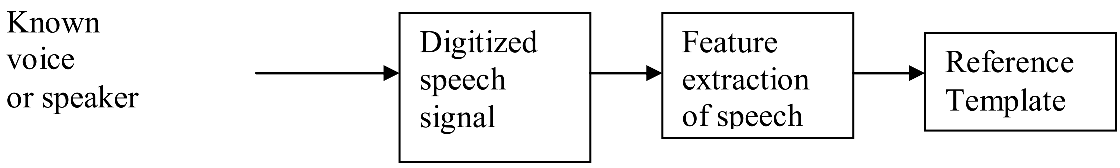

Figure 1.1: Training method block diagram

On training module, speech input, or known voice or speaker, is converted to its digital form. This can be accomplished using a microphone which records the human voice, and using an analog-digital-converter (ADC) to convert the signal to its digital form. Next, the digitized speech samples are then processed, using digital signal processing methods, to produce speech features. Features extraction is a process of jettisoning as much as irrelevant information as possible and representing relevant data in compact and meaningful form. The method that use is Fast Fourier Transform (FFT). Then the signal will be stored in reference template or database.

1.3.2 Testing Method

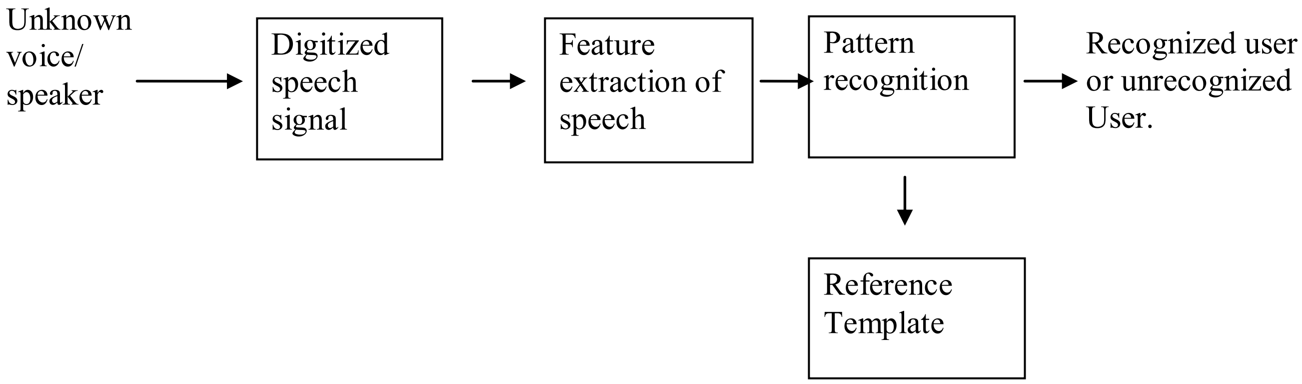

Figure 1.2: Testing method block diagram

On testing module, speech input, known or unknown voice or speaker is converted to its digital form. This can be accomplished using a microphone which records the human voice, and using an analog-digital-converter (ADC) to convert the signal to its digital form. Next, the digitized speech samples are then processed, using digital signal processing methods, to produce speech features. Features extraction is a process of jettisoning as much as irrelevant information as possible and representing relevant data in compact and meaningful form. The method that use is Fast Fourier Transform (FFT). Then the pattern matching stage, where the unknown words feature vectors are compared with the previously stored templates to find the best match. The method use is called Dynamic Time Warping (DTW). Then the output will be the recognized user or unrecognized user determine.

CHAPTER 2

LITERATURE REVIEW

2.1 OVERVIEW

This chapter will discuss about source or article that related to the project. There have many sources or researches done before and from there, details about this project are known and can understand briefly about the software.

2.2 SPEECH RECOGNITION SYSTEM

The definition of speech recognition system is the ability of a computer to understand spoken words for the purpose of receiving commands and data input from the speaker. Speech recognition system also can be defined as the ability for a software application to understand spoken human commands and act on them. The challenge with voice recognition technology lies in the ability for the program to distinguish contrasting patterns in the way people speak [1].

Speech recognition (also known as automatic speech recognition or computer speech recognition) converts spoken words to machine-readable input (for example, to keypresses, using the binary code for a string of character codes). The term "voice recognition" may also be used to refer to speech recognition, but can more precisely refer to speaker recognition, which attempts to identify the person speaking, as opposed to what is being said.

Speech recognition applications include voice dialing (e.g., "Call home"), call routing (e.g., "I would like to make a collect call"), domotic appliance control and content-based spoken audio search (e.g., find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g., a radiology report), speech-to-text processing (e.g., word processors or emails), and in aircraft [2].

In order to understand how speech recognition system works, figure 2.1 shows the conventional block diagram for speech recognition system.

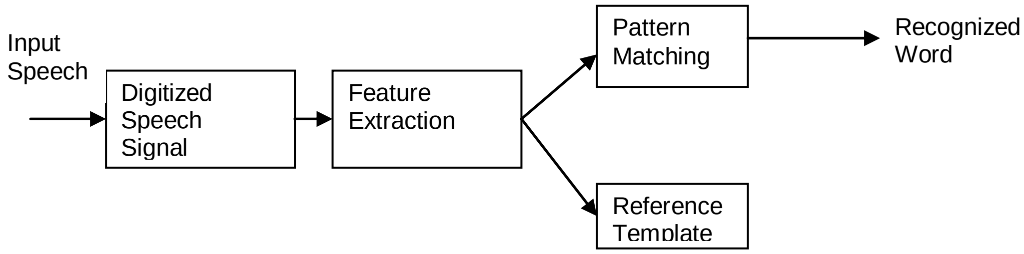

Figure 2.1: Conventional Speech Recognizer

Speech input, or human voice, is converted to its digital form. This can be accomplished using a microphone which records the human voice, and then using an analog-digital-converter (ADC) to convert the analog signal to its digital form. Thus, we will have digital data representing each level of signal at every time step.

Next, the digitized speech samples are then processed, using digital signal processing methods, to produce speech features. Feature extraction is a process of jettisoning as much irrelevant information as possible and representing relevant data in compact and meaningful form. A method of feature extraction includes Fast Fourier Transform, LPC analysis, Cepstrum, etc. These methods normally produce a sequence of feature vectors useful for the next stage, the pattern comparison stage.

Speech recognition system normally needs to be trained, where repeated utterance of known speech samples will be processed and their speech features stored in reference template. If unknown word needs to be recognized, the word is also processed but after the feature extraction stage, where the unknown word‟s feature vectors are compared with the previously stored templates to find the best match. The conventional method of pattern matching is called the Dynamic Time Warping (DTW) method.

2.3 SPEECH PATTERN MATHCING

One way to recognize a word is to compare it with a set of known words, or templates, stored in a machine, and to choose that word which it matches best.

2.3.1 Dynamic Time Warping (DTW)

The Dynamic Time Warping is a state of the art distance measure widely used in sequential pattern matching and it outperforms Euclidean distance in most cases because its matching is elastic and robust. The DTW is a method which two patterns of different lengths. The DTW method will map the time axis of the test patterns onto the time axis of the training pattern in such way that the resulting dissimilarity is minimized [3]. Suppose that there are 2 sequences patterns sets in graph 2.2 and graph 2.3

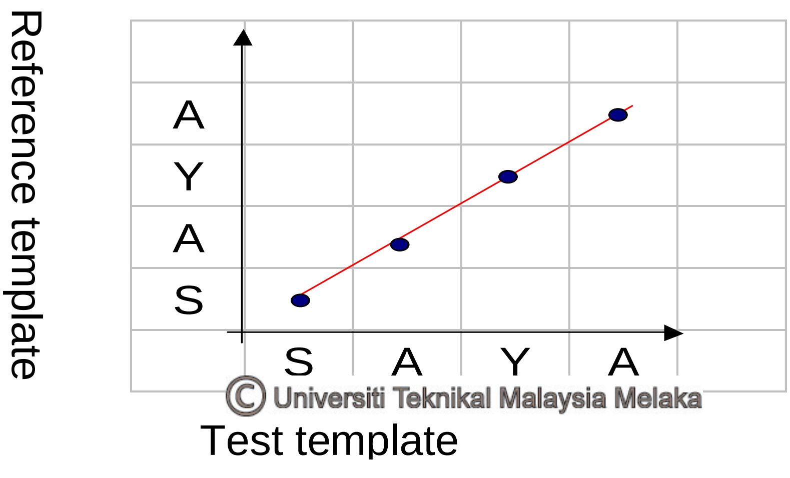

Figure 2.2: Graph for normal sequence pattern set.

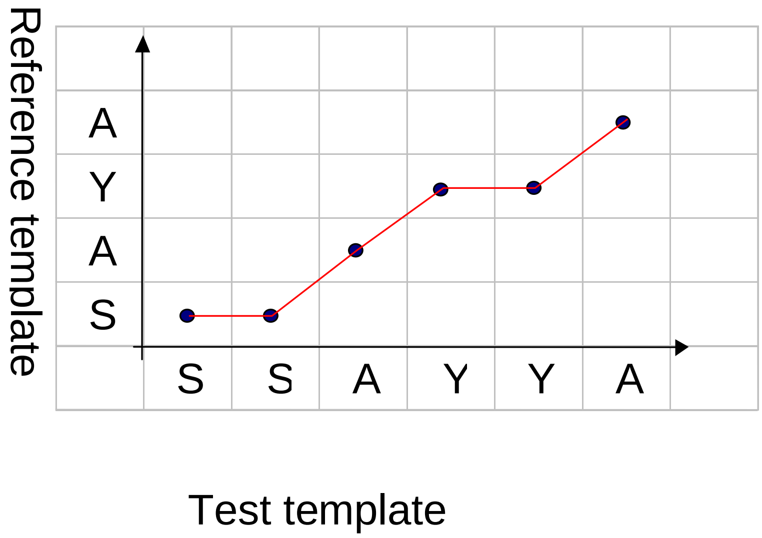

Figure 2.3: Graph for test sequence pattern set.

The line is called the warping path and each point of the line, the distance between the 2 sets. The test pattern is matched against all training patterns in the template, and the training pattern which gives the lowest accumulated distance is said to be recognized word.

2.3.2 Neural Networks

Another method of speech recognition is by using Neural Networks. A neural network is made up of computational units which mimic the biological neuron called a threshold logic unit, originally proposed by McCulloch and Pitts.

The artificial neuron is shown in figure 2.4:

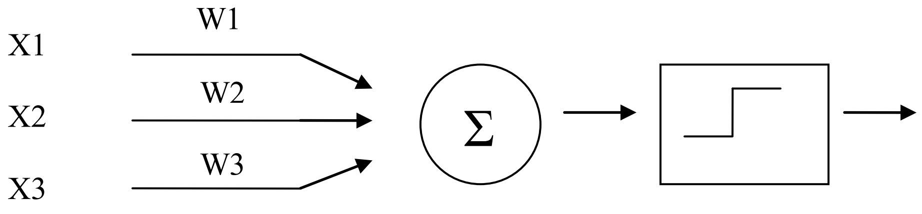

Figure 2.4: The Artificial Neuron

Here, a number of inputs (X1, X2…Xn) are weighted (W1, W2….Wn) and then summed up to produce an activation. If this activation exceeds a certain threshold, then the unit would produce a response, in the above case, either „1" or „0". The threshold function can also be a continuous function, such as the sigmoid function, where the output is then a continuous response between „1" and „0". A certain configuration of the neural units, called the multi-layer perceptron is normally used for the speech recognition [4]. Figure 2.5 show the model of multilayer perceptron for the neural networks.

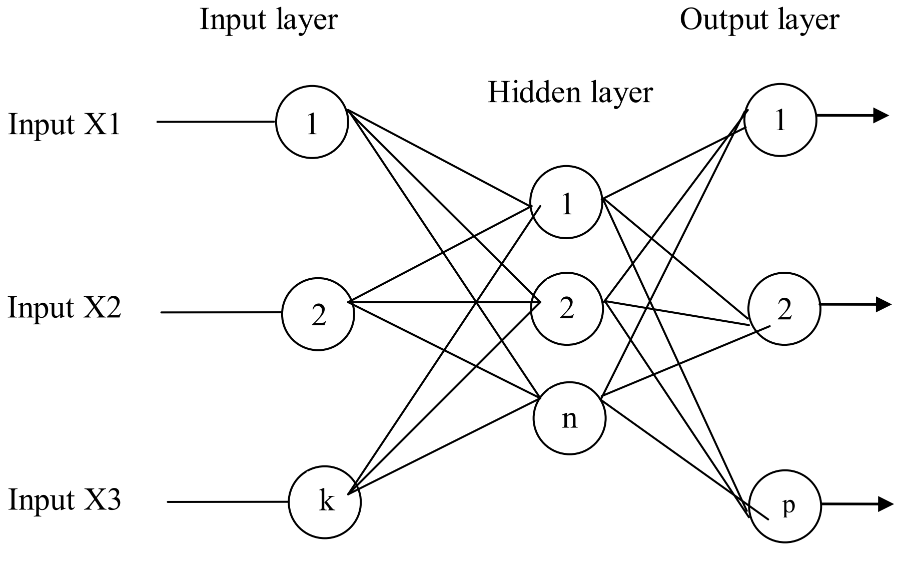

Figure 2.5: Multilayer Perceptron for the Neural Networks

This neural network is trained using speech data, using supervised learning in the context of neural network, to produce excepted outputs. Then if an unknown word is given as an input to the network, the network will then output the result accordingly.

2.3.3 Hidden Markov Model.

A hidden Markov model (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unknown parameters, and the challenge is to determine the hidden parameters from the observable parameters. The extracted model parameters can then be used to perform further analysis, for example for pattern recognition applications. An HMM can be considered as the simplest dynamic Bayesian network.

In a regular Markov model, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a *hidden* Markov model, the state is not directly visible, but variables influenced by the state are visible.

Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states.

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics [2]. Figure 2.6 shows the HMM model.