104793
A- K1

raf
TK7882.S65 .C46 2013
0000104793
Biometric voice recognition in security system / Chow Kok Lin.

BIOMETRIC VOICE RECOGNITION IN SECURITY SYSTEM

Chow Kok Lin

Bachelor of Mechatronics Engineering

May 2013

© Universiti Teknikal Malaysia Melaka

"I hereby declare that I have read through this report entitle "BIOMETRIC VOICE RECOGNITION IN SECURITY SYSTEM" and found that it has comply the partial fulfilment for awarding the degree of Bachelor of Mechatronic Engineering".

Signature : .....

Supervisor's Name : HAIROL NIZAM BIN MOHD SHAH

Date : 18/6/13

© Universiti Teknikal Malaysia Melaka

BIOMETRIC VOICE RECOGNITION IN SECURITY SYSTEM

CHOW KOK LIN

A report submitted in partial fulfillment of the requirements for the degree of Bachelor of Mechatronic Engineering

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

YEAR 2013

© Universiti Teknikal Malaysia Melaka

I declare that this report entitle "BIOMETRIC VOICE RECOGNITION IN SECURITY SYSTEM" is the result of my own research except as cited in the references. The report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : .....

Name : CHOW KOK LIN

Date : 18/6/2013

© Universiti Teknikal Malaysia Melaka

I wish to dedicate this thesis to my Mon and Dad, my Supervisor En.Hairol Nizam bin Mohd Shah.
Thank you for all of your support and helping.

© Universiti Teknikal Malaysia Melaka

Acknowledgement

First of all, I would like to thank to my supervisor, lecturers, my family and also my fellow friends. Thank to my supervisor En. Hairol Nizam bin Mohd Shah who guide me, teach me and give me some idea in order to complete this project. I am very thankful for his guidance and advice in this project.

After that, I would like to thank to all the lecturers who help me during I face the trouble in this project. All the lecturers are so kindly to help and give information that they know to help me complete this project.

Besides that, I want to give my gratitude to my friends for helping me when doing the analysis of experiment and also supporting and sharing ideas throughout the project.

Lastly, I would thank to my family for supporting me so that I can always look for them when I face problem.

© Universiti Teknikal Malaysia Melaka

Abstract

Biometrics is the new technology to measure the user features parameters. Voice recognition is the process of automatically recognize a certain word spoken by the admin user based on individual information included in speech waves. Signal processing for extracting the feature parameters of the user is an important stage in the voice recognition system. In this project, voice recognition system is designed to identify an administrator voice. By using MATLAB software for coding the voice recognition, the administrator voice can be authenticated. The key is to convert the speech waveform to some type of parametric representation for further analysis and processing. A wide range of possibilities exist for parametrically representing the speech signal for the voice recognition system such as Mel-Frequency Cepstrum Coefficients (MFCC). The input voice signal is recorded and computer will compare the signal with the signal that stored in the database by using MFCC method. The voice based biometric system is based on single word recognition. An administrator utters the password once in the training session so as to train and stored. In testing session the users can utter the password again in order to achieve recognition if there is a match. By using MATLAB simulation, the output can obtain either the user is being recognized or rejected. From the result of testing the system, it successfully recognizes the specific user's voice and rejected other users' voice. In conclusion, the accuracy of the whole system is successfully recognizing the user's voice. It is a medium range of the security level system.

© Universiti Teknikal Malaysia Melaka

Abstrak

Biometrik ialah teknologi yang canggih untuk mengukur ciri-ciri yang tersendiri oleh pengguna sendiri. Pengecaman suara adalah proses yang automatic mengiktiraf perkataan-perkataan tertentu yang disuarakan oleh pengguna tertentu berdasarkan maklumat tersendiri yang terdapat di dalam gelombang suara sendiri. Pemprosesan suara isyarat untuk mengeluarkan ciri pengguna tersendiri adalah penting dalam sistem pengecaman suara. Dalam project pengecaman suara, sistem pengecaman suara direka untuk mengenal pasti suara pengguna tertentu. Dengan megguanakan perisian MATLAB, suara pengguna tertentu dapat dikenal pasti. Yang penting adalah menukar gelombang suara kepada beberapa kemungkinan yang boleh digunakan untuk membuat analisis dan pemprosesan. Dengan mengguanakan kaegah MFCC, suara pengguna tertentu akan direkod dan computer akan menganalisis dan berbanding dengan data yang telah disimpan di dalam database. Suara pengguna tertentu akan direkod dan disimpan di dalam sesi latihan database. Pengguna akan menyuarakan kata laluan di sesi ujian. Kata laluan dikenal pasti, boleh mendapat luluan manakala kata laluan yang dinafikan, pengguna kena cuba sekali lagi. Daripada experiment yang dijalankan, suara pengguna tertentu dapat dikenal pastikan dan tidak mengenal suara pengguna lain. Akhirnya, the ketepatan pengecaman suara system adalah mengecam suara pengguna tertentu.

© Universiti Teknikal Malaysia Melaka

Table of Contents

Acknowledgement IV
Abstract VI
Abstrak VII
Chapter 1 1
Introduction 1
1.1 Project Background 1
1.2Problem Statement 3
1.3 Project Objectives 3
1.4 Project Scope 4
1.5 Thesis outlines 4
Chapter 2 6
Literature Reviews 6
2.1 Introduction 6
2.2 Speech Signal 6
2.2 Characteristics of Human Voice 8
2.3 Case Study 8
2.5 Conclusion 17
Chapter 3 18
Methodology 18
3.1 Chapter overview 18
3.2 Block Diagram of the Voice Recognition System Overview 19
3.3 K-Chart 20
3.4 Gantt Chart 22
3.5 Project implementation flow chart 23
..... 23

© Universiti Teknikal Malaysia Melaka

3.6 Voice Recognition Flow Chart 25
3.7 Control System Flow Chart 27
..... 27
3.8 Component used 28
3.8.1 ARDUINO UNO 28
3.8.2 LCD screen 29
3.8.3 Relay 30
3.8.4 Transistor 31
3.8.5 Magnetic Door Lock 31
3.9 Software 32
3.9.1 MATLAB 32
3.9.2 Arduino IDE 35
3.10 Performance Analysis 36
3.10.1 Experiment set up 36
3.10.2 Experiment I 36
Experiment II 37
3.11 Methodology Summary 38
Chapter 4 39
Results 39
4.1Chapter Overview 39
4.2Breadboard diagram of the Arduino Uno 39
4.3 Schematic diagram of Arduino Uno ..... **Error! Bookmark not defined.**
4.4 MATLAB Result 41
4.5 The Project System 43
4.5.1 Analysis of Project System 43
4.6 Performance Result 44
4.6.1 Experimental result for the accuracy of the admin's voice verification process (to compare the energy within each window, mean square error, and the average pitch) 44
4.6.2 Experimental result for verify the admin's voice between 10 different people. 47

© Universiti Teknikal Malaysia Melaka

4.6.3 Experimental Result for calculates the energy pitch with different tone of uttered speech from the admin..... 49
4.6.4 Experimental Result for different word of uttered speech from the admin..... 51
4.7 Summary of the results..... 53
Chapter 5..... 54
Conclusion and Recommendation..... 54
5.1 Chapter Overview..... 54
5.2 Conclusion..... 54
5.3 Recommendation..... 55
References..... 56
Appendix..... 57

© Universiti Teknikal Malaysia Melaka

LIST OF TABLES

TABLE	TITLE	PAGE
Table 2.1	Pair Wise Comparison Table of the Method's Criteria	15
Table 2.2	Weight Objective Table for the Method	15
Table 2.3	Point Explanation	15
Table 2.4	Comparison between MFCC and the VQ	16
Table 4.1	Result for accuracy for voice recognition system	45
Table 4.2	Result for recognize admin user among 10 users.	48
Table 4.3	Accuracy that word "HELLO" with different condition verified	50
Table 4.4	Different words other than "HELLO" that can be verified	52

© Universiti Teknikal Malaysia Melaka

LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure1.1	Utterance of "HELLO"	2
Figure2.1	Schematic Diagram of the Speech Production	7
Figure2.2	Speaker Verification	9
Figure2.3	Fusion classifiers system	10
Figure 2.4	Dynamic Time warping (DTW)	12
Figure 2.5	Pitch track with the autocorrelation method.	14
Figure 2.6	Vector quantization of codebook information.	14
Figure3.1	Block Diagram of the Voice Recognition System	19
Figure3.2	K-Chart	21
Figure3.3	Gantt chart	22
Figure3.4	Project implementation flow chart	23
Figure3.5	MATLAB Software Development	25
Figure3.6	ARDUINO Software Control System	27
Figure3.7	ARDUINO UNO Board	28

© Universiti Teknikal Malaysia Melaka

FIGURE	TITLE	PAGE
Figure3.8	LCD screen	29
Figure3.9	Relay	30
Figure3.10	Transistor	31
Figure3.11	Magnet Door Lock	31
Figure3.12	MATLAB	32
Figure3.13	Uttered speech of "Hello"	33
Figure3.14	MFCC Approach	34
Figure3.15	Arduino platform	35
Figure4.1	Breadboard diagram of Arduino Uno	39
Figure4.2	Schematic diagram of Arduino Uno	40
Figure4.3	the word "HELLO" as voice input signal	41
Figure4.4	the word "HELLO" after silence detection	42
Figure4.5	the word "HELLO" after using Hamming Window	42
Figure4.6	the word "HELLO" after FFT	43
Figure4.7	the word "HELLO" after Mel-warping	43

© Universiti Teknikal Malaysia Melaka

LIST OF ABBREVIATIONS

VRS	Voice Recognition System
DTW	Dynamic Time Warping
MFCC	Mel Frequency Cepstrum Coefficients
GMM	Gaussian Mixture Model
HMM	Hidden Markov Model
VQ	Vector Quantization
FFT	Fast Fourier Transform
MFN	Multilayer Feedforward Network
SVM	Support Vector Machines

© Universiti Teknikal Malaysia Melaka

LIST OF APPENDICS

Appendix A	Training phase code
Appendix B	Testing phase code
Appendix C	Arduino Uno R3 on board code
Appendix D	Gantt chart

© Universiti Teknikal Malaysia Melaka

Chapter 1

Introduction

1.1 Project Background

Nowadays, a lot of residential area and the companies are using all kinds of security system to make sure their property is secured such as using password and User ID/Pin for protection. Unfortunately, all these security system is not secured at all because the pin code can be hacked, the ID card can be stolen and duplicated. Based on the reasons, a whole new technology of security system must bring out to increase back the confidential of the civilian about the security system. [8]

A biometric technology is the one which use the user features parameter as the password. The feature parameters of everyone is unique, even the users are twins. Therefore, the voice recognition system is safe for the administrator user. Voice is the most natural way to communicate for humans. In this thesis, the issue of voice recognition is studied and a voice recognition system is developed for certain word being spoken. [8]

Voice biometric technology for authentication user is more convenient and accurate. This is because the biometrics characteristic if an individual are unique and belongs to the personal until the user dead. It is convenient for the user because nothing

© Universiti Teknikal Malaysia Melaka

to be carried or remembered and would not scare the ID card being stolen or password being hacked.

From a technological perspective it is possible to distinguish between two broad types of ASR: direct voice input (DVI) and large vocabulary continuous speech recognition (LVCSR). These systems will analyze users' specific voice and use it to fine tune the recognition of that user's speech, resulting in more accurate transcription.

The voice recognition system contains two main modules which are feature extraction and feature matching. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent each user while feature matching involves the step to identify the unknown user by comparing extracted features from admin voice input with the ones from a set of known user.

The speech signal and its characteristics can be represented in two different domains which are time and frequency domain. An utterance is the vocalization of a word or words that represent a single meaning to computer. From the figure 1.1, a sample utterance of "HELLO" is shown.

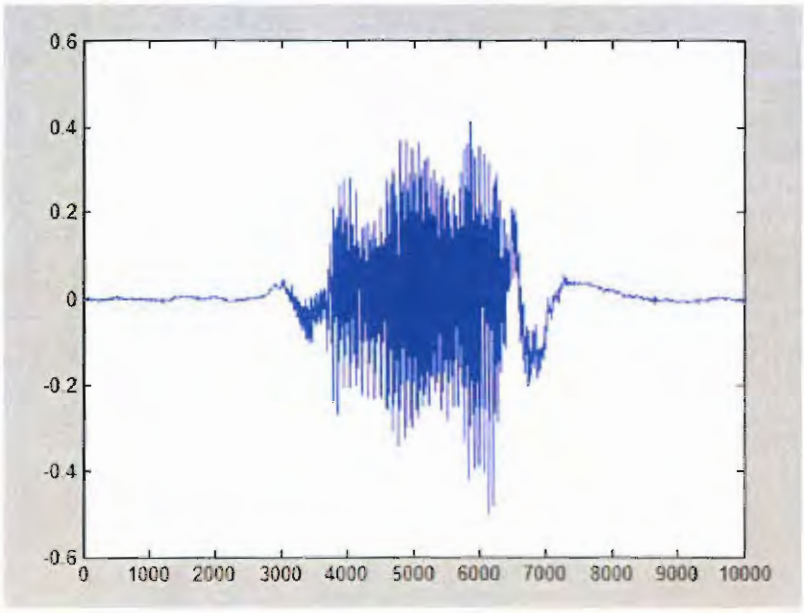

Figure1.1. Utterance of "HELLO"

© Universiti Teknikal Malaysia Melaka

Voice biometric is potential to growth worldwide because it requires the simple hardware. However, the imperfection of the voice recognition due to the background noise and some unwanted noise can increase the false rejection rate. Therefore, voice recognition software still can be improved.

1.2Problem Statement

In information technology, biometric technology is very secure and authenticable tool. It refers to a technology that is sensitive to measure and analyze human unique physical or behavior characteristics. Goal of the user is to create a voice recognition system with a high level of security. Previously, the common methods for the security system is using the password, pin code or user ID to access the security system. This sounds that the password, pin code and the user ID is very private, just knowing by the user themselves, but actually it is not. The previous security system will prone to hack; password will be cracked by the hacker and access into the security system and make the security system down.

Besides that, the user ID needs to bring along the user always. This increase the probability of the user ID being stolen or forget to bring. But with the voice recognition system, this problem will be solved easily.

1.3 Project Objectives

1.3.1 To develop voice recognition algorithm.

1.3.2 To authenticate a user's identity that able to declare as the admin voice.

1.3.3 To determine the suitable communication interface for the VRS

© Universiti Teknikal Malaysia Melaka

1.4 Project Scope

The main scope for this project is to create a voice recognition system by using MATLAB to perform the voice signal analysis for voice recognition purpose. By using the method of Mel Frequency Cepstral Coefficients (MFCC) to has a great success in speaker recognition in order to determine user voice as the administrator voice. The voice recognition system will only accept the admin voice and reject the impostor voice. Indoor will be more suitable choice for the voice recognition system because at the indoor situation will produce less irrelevant noise. At last, the communication interfaces between PC and device which are using the serial Port.

1.5 Thesis outlines

The voice recognition system is a combination of five chapters that contains and elaborates specific topics such as Introduction, Literature Review, Software Design, Result, Discussion, Conclusion and Further Development that can be applied in this project.

Chapter 1 basically is about an introduction of the whole project. The main idea about the introduction of voice recognition system and objectives of the project will be discussed and elaborated in this chapter.

Chapter 2 is about the literature review and the methodologies for the development of the voice recognition system.

Chapter 3 is discussed about the design of the voice recognition system by using MATLAB for the voice recognition and ARDUINO coding for the communication system. In this chapter, it will discuss about the feature extraction method and algorithm used in this thesis. It wills shows and explain the flow chart that been used to write the coding, developing the process using MATLAB and ARDUINO.

© Universiti Teknikal Malaysia Melaka

Chapter 4 is discussed about the results obtained and the limitation of the project. The entire voice signal figure and the experimental table will be placed on this chapter.

Chapter 5 is discussed about the analysis and discussion of the project. The discussion is concentrate on the result of the experiment and accuracy of the voice recognition system.

Chapter 6 is discussed about the conclusion and recommendation of the project.

© Universiti Teknikal Malaysia Melaka

Chapter 2

Literature Reviews

2.1 Introduction

In order to complete this project, a lot of literature reviews and research have been done. These studied including the noise filtering and feature extraction that is main important in this project. To succeed this voice recognition system, feature extraction is the main part to extract the feature parameters of the user's voice and set that user's voice as the only administrator. There are a lot of the general theories to understand the feature extraction behind each technology have been studied well. Articles, journals, reference books, notes, and internet have provided the sources of the literature reviews of the voice recognition. For more details of the voice recognition are describe below.

2.2 Speech Signal

In this chapter, the basic knowledge and fundamental concept of voice recognition will be referred and discussed. Below is the schematic diagram about the speech production.

© Universiti Teknikal Malaysia Melaka

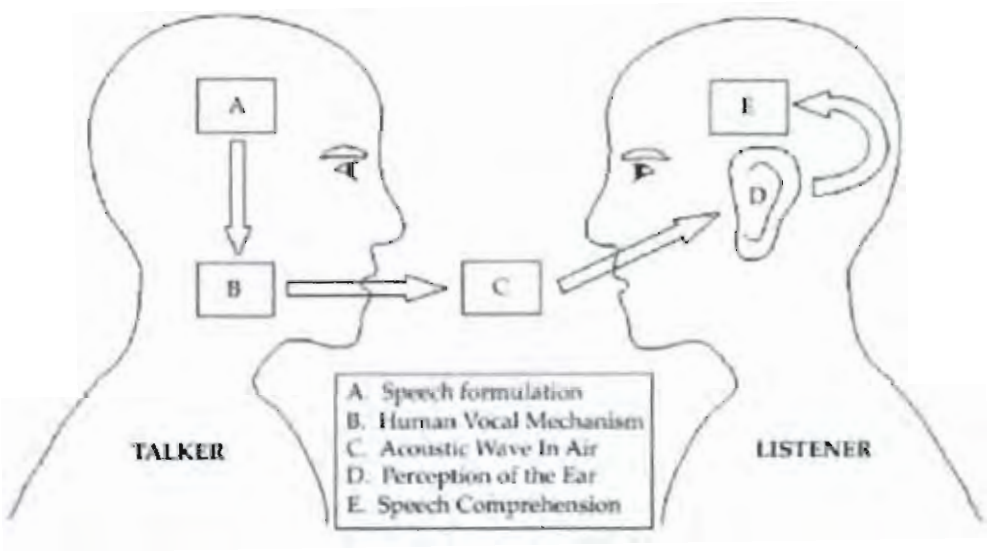

Figure2.1. Schematic Diagram of the Speech Production

Five different elements from the figure 2.1:

A. Speech formulation
B. Human Vocal Mechanism
C. Acoustic Wave in Air
D. Perception of the Ear
E. Speech Comprehension

The speech formulation is associated with the formulation of the speech signal in the talker's mind and used by the human vocal mechanism to produce the actual speech waveform. The waveform is transferred via the air to the listener. During the period of the voice signal transfer to listener, the wave can be affected by external sources such as background noise. The listener percepts the waveform when the voice signal reach to listener's ear. Then the listener's mind starts processing this waveform to comprehend its content so the listener understands what the talker is trying to talk.

Speech is:

- Time-varying signal
- Well-structured communication process
- Is different from every speaker

© Universiti Teknikal Malaysia Melaka

- Categorize to fast, slow, or varying in speed
- Categorize to high pitch, low pitch or whispered

2.2 Characteristics of Human Voice

Each human has a unique tone, rhythm, frequency and pitch to express including where they stop in phrases and how quickly they speak depending on where they are in a phrase [5]. Obviously the average male has a lower voice than the average female but the average range of each person's voice is unique. Humans have the interesting characteristic of different accents when they speak. Even across one certain word there are several variances on the way words and in turn sound is produced. The highest value of the frequency that a human can produce is about 10 kHz while the lowest value is about 70 Hz. The frequency interval changes for every person.

Speaker characteristics,

- Differences in vocal tract length, male, female, and children's speech
- Regional accents are differences in resonant frequencies, durations, and pitch
- Individuals have resonant frequency patterns and duration patterns that are unique (to identify speaker as admin)

2.3 Case Study

Voice recognition is the process by which a computer identifies spoken words. It can split into two types which are text dependent and text independent. Text dependent is about the keywords or the phrases for the voice recognition while text independent is not specific on the text being said and is more flexible. [12]

© Universiti Teknikal Malaysia Melaka