



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Simulation Study of Motor Controller via Voice Command

Thesis submitted in accordance with the requirements of the
Universiti Teknikal Malaysia Melaka for the Degree of
Bachelor of Engineering Manufacturing (Robotic and Automation)

By

AHMAD FAUZI BIN AHMAD KAMAR

Faculty of Manufacturing Engineering

April 2007



UNIVERSITI TEKNIKAL KEBANGSAAN MALAYSIA

BORANG PENGESAHAN STATUS TESIS*

JUDUL: SIMULATION STUDY OF MOTOR CONTROLLER VIA VOICE COMMAND

SESI PENGAJIAN: 2/2006-2007

Saya AHMAD FAUZI BIN AHMAD KAMAR

mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hak milik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (√)

 SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia yang termaktub di dalam AKTA RAHSIA RASMI 1972)

 TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

 TIDAK TERHAD

Disahkan oleh:

(TANDATANGAN PENULIS)

(TANDATANGAN PENYELIA)

SAMSI BIN MD. SAID

Pensyarah

Cop Rasmi:

Fakulti Kejuruteraan Pembuatan
Universiti Teknikal Malaysia Melaka
Karung Berkunci 1200, Ayer Keroh
75450 Melaka

Alamat Tetap:
No. 40, Jalan Tenaga 12,
Taman Tenaga,
43000 Kajang, Selangor Darul Ehsan

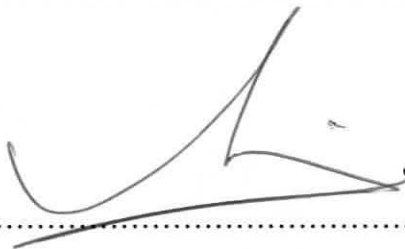
Tarikh: 16/05/2007

Tarikh: _____

* Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan, atau disertasi bagi pengajian secara kerja kursus dan penyelidikan, atau Laporan Projek Sarjana Muda (PSM).
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT atau TERHAD.

APPROVAL

This thesis submitted to the senate of UTeM and has been accepted as partial fulfillment of the requirements for the degree of Bachelor of Manufacturing Engineering (Robotic and Automation). The members of the supervisory committee are as follow:



Main Supervisor
SAMSUL HUDA SAID
Pensyarah
Fakulti Kejuruteraan Pembuatan
Universiti Teknikal Malaysia Melaka
Karung Berkunci 1200, Ayer Keroh
75450 Melaka

.....
Co- Supervisor

DECLARATION

I hereby, declare this thesis entitled “Simulation Study on Motor Controller via Voice Command” is the results of my own research except as cited in the reference.

Signature : 

Author's Name : AHMAD FAUZI B AHMAD KAMAR
Date : 08 MAY 2007

ABSTRACT

Every human in this world need a safer home to make sure their lives not in the danger. That's why more application is been applying to the house to make sure a better home. Some of the application that add to make sure the house safer is automated window that will shut when raining is coming. For the door, voice recognition can be added to make sure nobody than the speaker can enter the room.

Voice recognition is a method to make a system recognize the speaker identity. Such systems extract features from speech, model them and use them to recognize the person from his/her voice. Voice recognition is system that recognizing who is speaking and speech recognition is a system to recognizing what is being said. Generally these two terms are frequently confused and voice recognition is used as a synonym for speech recognition instead. The main contribute of this project is to recognize the speaker voice and activate the motor that represent any application that can be apply. Matlab (Mathematics Laboratory) will be use to recognize and analyze the people's voice and if the voice wave match the sample in the template, the control circuit, PIC microcontroller circuit, will activate the motor. This project has great potential since everybody need a comfortable house.

This project is start by researching the simple voice recognition that available in open source. After getting the system that suitable with the project, learning session will be start before building and testing the system. In the same time, the PIC motor controller also will be build. The PIC system is not a very difficult to build as the information is easy to get in internet, book and journal. But for the for the voice recognition it is a little bit hard to create the system although there is some application or past studies about it. The objective of the project is to show the simulation the motor movement via voice command. But in the end, only the PIC motor controller and voice recognition is been

complete. The serial connection between PIC and voice recognition system is been unable to comply because of some problem that cannot be avoided.

This system is start form transferring a human voice which in an analog into the computer to change it into a digital. The wave is been receive by the microphone and than it will transmit it to the Matlab. The wave will be go through some signal processing where the wave will be change from analog signal it into a digital signal. After the signal is been process, the wave that been recorded will be train in Matlab by using Neural Network. This step is to create a template that will be use for the testing the system afterwards. The PIC motor controller is being use to represent any application that can be used for this system.

This project indeed is a very interesting and challenging project in many aspects based on the stated objectives and presented information. Although facing some problem, but I'm manage to finish this project. The system can make a different between the two words that been recorded, process and train inside the Matlab. The PIC motor controller is also working very well.

ACKNOWLEDGEMENT

Firstly, the author would like to thank supervisor, Mr. Samsi bin Md. Said for his trust to carry out this project as an undergraduate student. His guidance and advices are essential in helping to complete this project.

The author would also like to express the deepest thank to beloved parent and all my family members for their unconditional love, concern, and moral supports.

The author also want to express a special thank you to Miss Khairiah bt Yazid and Miss Nuraira from Malaysian Institute of Nuclear Technology (MINT) for their cooperation and willingness to tech me about the software that will be use in this project.

Finally, the author would like to thanks all buddies and member of 4 BMFA for their endless spiritual support and encouragement. Without their help, this project may not be completed

DEDICATION

Specially dedicated to My Father, Ahmad Kamar Roslan, My Mother, Subayah bt Masron, and to all my siblings who are concern, understanding and supporting.
Thanks for everything.

TABLE OF CONTENTS

Abstract	I
Acknowledgement	IV
Dedication	V
Table of Contents	VI
List of Figure	VIII
List of Table	XI

1 GENERAL INTRODUCTION

1.1	Introduction	1
1.2	Objectives	2
1.3	Problem Statement	2
1.4	Scopes	3

2 LITERATURE REVIEW

2.1.	Introduction	5
2.2.	Voice Recognition / Speech Recognition	5
2.3.	Past studies on Voice Recognition	14
2.4.	Past Studies on Speech Recognition	18
2.5.	Analog And Digital Signal	20
2.6.	PIC architecture	21
2.7.	Interfacing both Computer and PIC	34
2.8.	Discussion	36

3 METHODOLOGY

3.1	Introduction	37
3.2	Project Design	37
3.3	Project Methodology	42
3.4	Project Tool	43
3.5	Project Planning	44
4	BUILDING THE SYSTEM	
4.1.	Introduction	48
4.2.	Project procedure	48
4.3.	Building The Motor Controller System	57
5	RESULT ANALYSIS AND DISCUSSION	
5.1.	Introduction	65
5.2.	Result From The Voice Recognition System	65
5.3.	Result From The Motor Controlling System	69
5.4.	Discussion From The Result	71
6	SUGGESTION AND CONCLUSION FOR FUTURE WORK	
6.1	Introduction	74
6.2	List of suggestion	74
6.3	Conclusion	75
4	Reference	76
5	Appendix	79

LIST OF FIGURE

Figure 2.1. Example of the speech recognition	10
Figure 2.2: PIC16F84A & PIC16F877A	20
Figure 2.3 PIC16F84A architecture	21
Figure 2.4: Example of the PIC programming	24
Figure 2.5: PICStart	24
Figure 2.6: uJDM PIC programmer circuit	25
Figure 2.7 Capacitor	26
Figure 2.8 Transistor	26
Figure 2.9 Relay 12volt	27
Figure 2.10 Resistance	27
Figure 2.11 Crystal 4MHz	28
Figure 2.12 Switch Button	28
Figure 2.13 Wire	29
Figure 2.14 Strip Board	29
Figure 2.15 Soldering stuff	29
Figure 2.16 Battery 12 volts	30
Figure 2.17 DC motor	30
Figure 2.18 PIC circuit drawing for the project	31
Figure 2.19 Example of PIC circuit	32

Figure 2.20 Example of RS232 connector	35
Figure 3.1 System Flow for the motor controller system	37
Figure 3.2 Flow chart for project design	43
Figure 4.1 A single line wave graph for word Jalan1	46
Figure 4.2 Display of program writing for target	48
Figure 4.3 Display of program writing for input	49
Figure 4.4 Display of network data manager	50
Figure 4.5 Display of networks input and target attachment option	50
Figure 4.6 Display of setting the network and layers	51
Figure 4.7 Display of networks that been created	52
Figure 4.8 Log-sigmoid transfer functions	52
Figure 4.9 Display of networks parameters	53
Figure 4.10 Display of graph during the training session	53
Figure 4.11 Display of network data manager after the training session is complete	54
Figure 4.12 Display of network/data manager when saving it	54
Figure 4.14 Diagram of uJDM board	56
Figure 4.15 Diagram of RS 232 head	57
Figure 4.16 Diagram of PIC motor controller circuit	60
Figure 4.17 Diagram of PIC motor controller circuit	61
Figure 4.17 5 volts voltage regulator	61
Figure 5.1 Overall flow chart for voice recognition system	62
Figure 5.2 A digital signal of word Jalan1	63

Figure 5.3 A Digital signal after been process and have been divided into 25 blocks	63
Figure 5.4 A different Zerocrossdata graph between the word Jalan (a) and Stop (b).	64
Figure 5.2 Result from the training	65
Figure 5.2 Display of the answer	66
Figure 5.4 A different Zerocrossdata graph between the word Bahan (a) and Tahan (b).	69

LIST OF TABLE

Table 3.1 Gannt Chart for PSM 1	46
Table 3.2 Gannt Chart for PSM 2	47

CHAPTER 1

GENERAL INTRODUCTION

1.1. Introduction

Every people in this world will always need a safety home. Home need to be safe because it is the place that human needs protection from danger or bad weather. Stone age human live in the cave because they need a safety from beast and bad weather like rain. But the caves are not so safety because it doesn't have the safety aspect like a door. When human want a buy a house, they will do inspection to all safety aspect like how well the wall, is the door work correctly or even the toilet they will check.

That's the basic safety factor that human will be inspection when they want to buy a house. To make sure there are more safety in the house, more application need to add to the house for example automated light to make sure the house will be bright even tough the owner are not home at night. This application is famous nowadays and they is some automated is already been build in this world.

This application is add sensor to its application like the automated curtain. When the light is on, the curtain will be close because the sun is already set. Now to make an upgrade to the automated house application, there is need some new application. To open and close every curtain in the house, in case the house is big, it's to tiring. So there is some upgrade to it by using the remote control or timer to close and open it. Besides

the remote application, one of the other applications can be adding is voice command. By using this application, the human or the speaker can sit and make the curtain close and open by using its voice.

1.2. Problem Statement

To make sure the house is safer place to live, voice command is the one of the best way to make any system safer. Speaker verification is usually used in applications which require secure access. The systems operate with the user's knowledge and typically require their cooperation. Speaker identification systems are more likely to operate covertly without the user's knowledge. This can for example be used to route users to the correct mailbox, identify talkers in a discussion, alert speech recognition systems of speaker changes, check if a user is already enrolled in a system, etc. Now the system that uses voice command can be safer because when come to voice, every person in this world have a different voice. One solution to the crime and illegal immigration problem facing South Africa is the use of biometrics techniques and technology. Biometrics is methods for recognizing a user based on unique physiological and/or behavioral characteristics of the user [6].

1.3. Objectives

The objectives of this project must be clear and achieve to produce a successful system. The objectives are:

- I. Make the voice command system can identify and make a different between two words.

The main target of this project is to make sure this system can make a different between two words that been recorded as a template. This template will be use to identify which word that the speaker said.

II. Build a PIC motor controller

In this project, the movement of the motor will be control with the PIC microcontroller. This motor movement will represent any application that can be applied using the voice command system.

III. Connect the voice command system to control motor movement

When all the system is been build successful than it will be connect with using the serial connection.

1.4. Scopes

The scopes should be identified and planned appropriately. The scope is:

I. Research the latest technology in voice command system

Necessary data on the current voice command and PIC microcontroller system for any application was collected by conducting case studies, analyzing journal, book and past researches, visiting local research institute and searching from internet sources which are related to the voice command and PIC microcontroller. From the research survey and data collection, it helps the author to find few important facts which will determine the successful of building this system.

II. Building and Simulation.

In this project, the software used is MATLAB and MPLab. MATLAB is used to design the necessary voice command system while the MPLab will be use to programming the PIC.

III. Building the circuit.

So many researches also must be doing on this section. The electronic part also needs some concentration to build it so the circuit than will be build can work properly.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

Technology is the main attraction nowadays. The world now are been surround by technology. For example telephone, computers, automation industry and many more. If the technology can be apply to the bigger application like the automation industry, smaller application also must have some technology on it. For example automated house where all the application in the house can be control like closing and opening the door or gate.

One of the technologies in automated house is automated light. The house light will be on in curtain time and that is already been set earlier. There are also some other technology for the household like automated rice cooker, automated gate, or even automated sprinkle.

2.2. Voice Recognition / Speech Recognition

2.2.1. Definition

Voice or speech recognition is the ability of a machine or program to receive and interpret dictation, or to understand and carry out spoken commands. For use with

computers, analog audio must be converted into digital signals. This requires analog-to-digital conversion. For a computer to decipher the signal, it must have a digital database, or vocabulary, of words or syllables, and a speedy means of comparing this data with signals. The speech patterns are stored on the hard drive and loaded into memory when the program is run. A comparator checks these stored patterns against the output of the A/D converter.

Speech recognition (in many contexts also known as 'automatic speech recognition', computer speech recognition or erroneously as Voice Recognition) is the process of converting a speech signal to a sequence of words, by means of an algorithm implemented as a computer program. Speech recognition applications that have emerged over the last years include voice dialing (e.g., Call home), call routing (e.g., I would like to make a collect call), simple data entry (e.g., entering a credit card number), and preparation of structured documents (e.g., a radiology report). Voice Verification or speaker recognition is a related process that attempts to identify the person speaking, as opposed to what is being said [8].

Note that strictly speaking there is a difference between voice recognition (recognizing who is speaking) and speech recognition (recognizing what is being said). Generally these two terms are frequently confused and voice recognition is used as a synonym for speech recognition instead. [1]

2.2.2. Voice Recognition

Voice recognition technologies can roughly be categorized as follows from the viewpoint of telecommunications applications.

- Word recognition

Word recognition is the technology to recognize and identify a spoken isolated word among a vocabulary of words. The number of words in the Vocabulary may be small (e.g. up to 10) or very large (e.g. a few thousands).

- Word spotting

Word spotting is the technology to recognize and identify a target word in a continuously spoken phrase. The number of words in the vocabulary may be very small. It is necessary to ignore any words which are not registered in the vocabulary. In this sense, rejection of non-registered words is also important.

- Continuous speech recognition

Continuous speech recognition is the technology to recognize all or multiple words uttered in the form of phrases or sentences. The key technology here is how to integrate the linguistic constraints into the recognition algorithm.

- Speaker recognition

Speaker recognition is the technology to verify that the speaker uttering some words is the same as the one pre-registered (speaker verification), or to identify who among a set of pre-registered persons is the one (speaker identification). Among these technologies, word recognition and word spotting are widely used for practical application [12] Speaker verification by asking the speaker to utter a pre-registered word can be used for some applications, as well. Continuous speech recognition is emerging as a candidate for user-friendly interface; however its applications are still limited.

One solution to the crime and illegal immigration problem facing South Africa is the use of biometrics techniques and technology. Biometrics is methods for recognizing a user based on unique physiological and/or behavioral characteristics of the user [6]. There are two applications in voice recognition which is voice verification and voice identification. If the speaker claims to be of a certain identity and the voice is used to verify this claim this is called speaker verification or voice authentication. On the other hand, speaker identification is the task of determining an unknown speaker's identity. In a sense speaker verification is a 1:1 match where one speaker's voice is matched to one template (and possibly a general world template) whereas speaker identification is a 1: N match where the voice is matched to N templates.

Speaker verification is usually used in applications which require secure access. The systems operate with the user's knowledge and typically require their cooperation. Speaker identification systems are more likely to operate covertly without the user's knowledge. There are two phase to complete the voice recognition, enrollment and test. During enrollment the speaker's voice is recorded and typically a number of features are derived to form a voice print, template, or model. In the test phase (also called verification or identification phase) the speaker's voice is matched to the templates or models. Voice recognition systems employ three styles of spoken input: text-dependent, text-prompted and text-independent. This relates to the spoken text used during enrollment versus test.

If the text must be the same for enrollment and test this is called text-dependent recognition. It can be divided further into two cases: The highest accuracies can be achieved if the text to be spoken is fixed. This has the advantage that the system designer can devise a text which emphasizes speaker differences. However, since the text is always the same such systems are vulnerable to Impostors. Furthermore, it is not very user friendly if all users have to remember some complex text and in addition it makes the system language dependent.

Another type of text-dependent system uses pass phrases. The user is free to pick a phrase during enrollment but must use the same phrase during test. Most speaker verification applications use this type of text-dependent input. It has the advantage that an impostor must know the pass phrase, which adds a level of security. However, such systems are still vulnerable to tape recorder attacks. To counter this, many systems allow the specification of several pass phrases. Typically, these are answers to questions. During test the system randomly asks the user one of the questions and the user must provide the correct answer. However, the number of different questions is usually rather limited and thus a patient attacker could still attempt a tape recorder attack.

In text-prompted systems the speaker is asked to speak a prompted text. In principle this could be any kind of text. This however complicates the recognition process quite a bit. On the one hand the system knows the text that is being spoken. On the other hand the system must somehow know how this randomly selected text should sound if spoken by a particular speaker. This involves a much more elaborate model than in text-dependent systems where the text is always the same. The typical implementation makes use of digits. During enrollment the speaker is asked to speak a few digit sequences which are carefully selected such that all digits occur equally often. For every digit one speaker-specific model is trained. This has the advantage that only ten such models must be trained. During test a random digit sequence is selected and the corresponding digit models are concatenated into one model per speaker. Since speakers typically do not change languages frequently, such systems can be made language-independent quite easily.

Text-independent systems are most often used for speaker identification as they require very little if any cooperation by the speaker. In this case the text during enrollment and test is different. In fact, the enrollment may happen without the user's knowledge. Some recorded piece of speech may suffice. Since text-independent systems have no knowledge of the text being spoken only general speaker-specific properties of the

speaker's voice can be used. This does limit the accuracy of the recognition. On the other hand, this approach is also completely language independent.

2.2.3. Technology in Voice Recognition

The various technologies used to process and store voiceprints include frequency estimation, Hidden Markov models, pattern matching algorithms, neural networks, matrix representation and decision trees. Some systems also use "anti-speaker" techniques, such as cohort models, and world models [5].

Ambient noise levels can impede both collections of the initial and subsequent voice samples. Noise reduction algorithms can be employed to improve accuracies. Performance degradation can result from changes in behavioral attributes of the voice and from enrollment using one telephone and verification on another telephone. Voice changes due to aging also need to be addressed by recognition systems. Some systems adapt the speaker models after each successful verification to capture such long-term changes in the voice.

Many companies market speaker recognition engines, often as part of large voice processing, control and switching systems. Capture of the biometric is seen as non-invasive. The technology needs little additional hardware by using existing microphones and voice transmission technology allowing recognition over long distances via ordinary telephones (wired or wireless).

2.2.4. Speech Recognition

The ability to communicate with a robot through speech is the ultimate user interface. Speech is one of the human attributes that sets us apart from the rest of the animal kingdom and is considered proof of higher intelligence. When a robot obtains the ability to recognize words, it is well on its way to becoming a true humanoid. Speech