

**INDEPENDENT LEARNING OF QURAN (ILoQ)-ALPHABET
USING SPEECH RECOGNITION**

SHAFIQAH BINTI SHOLEHUDDIN

**A thesis submitted in fulfillment of the requirements
for the award of the degree of the Bachelor of Computer Science
(Computer Systems & Networking)**

**Faculty of Computer Systems & Software Engineering
University Malaysia Pahang**

MAY 2011

PERPUSTAKAAN UNIVERSITI MALAYSIA PAHANG	
No. Perolehan 068694	No. Panggilan TK 7883 565 543 2011 rs bc.
Tarikh 30 NOV 2012	

ABSTRACT

Speech recognition is the process by which a computer or other type of machine identifies spoken words that captured through a microphone. Speech recognition is still new and still under the research in Malaysia. Independent Learning of Quran (ILoQ)-Alphabet is developed to do more research about speech recognition, identify and advance the tools for independent learning and also to search the way to measure the accuracy of basic Quran pronunciation. ILoQ-Alphabet is a computer application to identify the correctness of user pronunciation by pronounces the Arabic's alphabet. This application is focusing on reading, listening and speaking as methods for independent learning. Waterfall model is use as a methodology to develop this application. This application is hope to be able to attract user especially kids in learning Quran independently and also give a lot of benefits for Muslims to learn Quran easily.

ABSTRAK

Pengecaman ucapan adalah satu proses yang melibatkan komputer atau mesin untuk mengenalpasti perkataan yang diucapkan melalui alat pembesar suara (mikrofon). Di Malaysia, pengecaman suara adalah masih baru dan masih dalam penyelidikan. Independent Learning of Quran (ILoQ)-Alphabet telah direka untuk memperbanyakkan penyelidikan mengenai teknologi pengecaman ucapan, mengenalpasti dan memperhebatkan alatan media untuk pembelajaran sendiri serta untuk mencari jalan untuk mengukur ketepatan bacaan asas Quran. ILoQ-Alphabet adalah sebuah aplikasi komputer untuk mengenalpasti ketepatan sebutan huruf-huruf di dalam Quran oleh pengguna. Aplikasi ini menumpukan kepada bacaan, pendengaran dan pengucapan sebagai cara dalam pembelajaran sendiri. Model Waterfall digunakan sebagai metodologi untuk membangunkan aplikasi ini. Aplikasi ini diharapkan dapat menarik pengguna khususnya kanak-kanak untuk mempelajari Quran sendirian dan juga memberi banyak manfaat kepada Muslim untuk belajar Quran dengan mudah.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	STUDENT'S DECLARATION	ii
	SUPERVISOR DECLARATION	iii
	DEDICATION	iv
	ACKNOWLEDGEMENT	v
	ABSTRACT	vi
	LIST OF TABLES	xi
	LIST OF FIGURES	xii
	LIST OF APPENDICES	xiv
	ABBREVIATION	xvi
1	INTRODUCTION	1
	1.1 Introduction	1
	1.2 Problem Statement	2
	1.3 Aim and Objective	4
	1.4 Scope	5
	1.5 Organization of Thesis	5
2	LITERATURE REVIEWS	7
	2.1 Introduction	7
	2.2 Studies on Related Research	8
	2.2.1 Design of Educational Software for Automatic Speech Recognition (ASR) Technique	8
	2.2.2 Adaptation Techniques for Children's Speech Recognition	11

CHAPTER	TITLE	PAGE
	2.3 Study on Current Existing Software	13
	2.3.1 Tell Me More	13
	2.3.2 Rosetta Stone	15
	2.3.3 Pimsleur	17
	2.3.4 Comparison of Existing Software	19
	2.4 Development Methodology	21
	2.4.1 Waterfall Model	21
	2.4.2 Agile Method	22
	2.4.3 Rapid Application Development (RAD)	23
	2.5 Software Development	24
	2.5.1 MATLAB 7.8.0	24
	2.6 Conclusion	24
3	METHODOLOGY	26
	3.1 Introduction	26
	3.2 Project Method	27
	3.2.1 Analysis of Requirement Stage	29
	3.2.2 Software Design Stage	31
	3.2.3 Implementation Stage	33
	3.2.4 Testing Stage	35
	3.3 Expectation Result	35
	3.4 Summary	35
4	IMPLEMENTATION	36
	4.1 Introduction	36
	4.2 MATLAB Environment Configuration	37
	4.2.1 M-File	37
	4.2.2 User Guide for Starting MATLAB 7.8.0	38
	4.3 Flow of ILoQ Application	41
	4.4 Interface Design and Coding Implement	42
	4.4.1 Main Interface	42

CHAPTER	TITLE	PAGE
	4.4.2 Listening Interface	45
	4.4.3 Arabic Alphabet Interface	46
	4.4.4 Speaking Interface	47
	4.4.5 Reading Interface	51
5	RESULT AND DISCUSSION	53
	5.1 Introduction	53
	5.2 Result	53
	5.2.1 Output	54
	5.2.2 Functionality Testing	56
	5.3 Project Limitation	57
	5.3.1 Development Constraint	57
	5.3.2 System Constraint	58
6	CONCLUSION	59
	6.1 Summary of the Project	59
	6.2 Contribution and Achievement	60
	6.3 Future Research Direction	61
	REFERENCES	62
	APPENDICES	65

LIST OF TABLES

TABLE NO	TITLE	PAGE
2.1	Result of accuracy tested to different age	12
2.2	Tell Me More advantages and disadvantages	14
2.3	Advantages and disadvantages of Rosetta Stone	17
2.4	Pimsleur advantages and disadvantages	19
2.5	Comparison between Tell Me More, Rosetta Stone and Pimsleur.	19
3.1	Software requirement	30
3.2	Hardware requirement	30
4.1	M-File in ILoQ	37
4.2	GUI components' description	41
5.1	Functionality testing result	56

LIST OF FIGURES

FIGURE NO	TITLE	PAGE
2.1	Overview of speech recognizer training process	9
2.2	Overview of speech recognition	9
2.3	GUI of main module	10
2.4	Design of Speech Studio interface	10
2.5	Interface design of VQ Classroom	11
2.6	Voice pattern in Tell Me More	13
2.7	An example of method that use in Rosetta Stone software	16
2.8	Waterfall SDLC Methodology	22
3.1	Waterfall methodology in ILoQ development	28
3.2	Flowchart for ILoQ's development	29
3.3	Flowchart diagram of ILoQ	32
3.4	Sequence diagram of interface in ILoQ	33
3.5	Sample of ILoQ's interface	34
3.6	Reading interface	34
4.1	Starting MATLAB using Start Menu	38
4.2	Default MATLAB Desktop	39
4.3	GUIDE Quick Start Interface	40
4.4	Blank GUI template	40
4.5	ILoQ's interface flow	42
4.6	Main interface of ILoQ design	43
4.7	Page connection coding	43
4.8	Close dialog box	44
4.9	Source code for Exit button	45

FIGURE NO	TITLE	PAGE
4.10	Design of Listening interface	45
4.11	Source code to play an audio	46
4.12	Speaking interface design	47
4.13	'Alif' Speaking interface	48
4.14	Speak button and wait bar coding	49
4.15	Source code to display audio wave	50
4.16	Design of Reading interface	51
4.17	Source code to view the image	52
5.1	Audio signal of Alif pronunciation	54
5.2	Result from comparing two audio files	55

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Gantt Chart	64
B	User Manual	66

ABBREVIATION

IloQ	-	Independent Learning of Quran
MATLAB	-	MATrix LABoratory
SDLC	-	Software Development Life Cycle
VTLN	-	Voice Tract Length Normalization

CHAPTER 1

INTRODUCTION

1.1 Background Information

Independent Learning of Quran (ILoQ) - Alphabet is a computer application that can help users especially beginner to learn how to pronounce the alphabet of Quran. ILoQ has three elements for education which are listening, pronunciation and reading. Users can view and hear the pronunciation of alphabet. Besides, user can pronounce the alphabet in order to indentify either he/she has pronounced the alphabet correctly or not. The application used speech recognition in order to detect the accuracy of reciting Quran. The application will function when users pronounce the word by using microphone.

ILoQ also provides the image for users to recognize or read the basic of Quran's alphabets. Reading is good method in learning because reader can obtain understanding by learning how to pronounce and identify the words. Additionally, listening also can be described as a suitable scheme in education. As a listener, a person has to give close attention to the matter. So, people that are active listeners use new information more productively and as a result active listener exhibit

better performance in concentration and memory. Because of the effectiveness of listening technique, ILoQ will provide the voice that already been set in the application as a reference on how to deliver the pronunciation of Quran's alphabets for listener. This can teach users to pronounce the word properly and correctly.

ILoQ is an application package that can help beginners to learn the basic Quran's alphabet ike 'Alif, Ba, Ta'. The application should be user friendly and useful for those who want to learn and read Quran properly or accurately.

1.2 Problem Statement

Quran is a sacred writings of Islam revealed by Allah to the prophet Muhammad during his life at Mecca and Medina [1]. Quran consists of 114 surah and are formed from Arabic alphabet. Quran alphabet consists of 28 basic letters starting from '*alif*' until '*ya*'. Quran alphabet should be learned first before people learn to recite Quran.




In Malaysia and worldwide, the Muslims have to learn Quran and need guidance or instructor which could be parents or '*ustaz*' to teach and guide the learners. Learners can be divided into two groups which are kids as a group of the beginners and intermediate learner and also elderly as a higher level learner. For the elderly, they need to attend the class for re-learn on how to recite Quran properly. For the kids, they need parent to teach them to recite and learn Quran from the beginning. But, some of the parents are busy at work and have no time to teach. So, as an alternative, kids are sent to Quran class. That means, parents have to spend extra money to pay for the fee and elderly learners have to spend spare time to go to the Quran class. However, with the suitable applications in computer, people can teach themselves a new thing such as foreign language, Quran and others without needs a teacher or going to extra class.

Nowadays, the world urbanized with thousands of technologies and developments. As the human race become more complex and busy, computers turn out to be human's best friends. Most of people activities are done by application in computers such as purchase things online, sharing files using peer-to-peer (P2P) application and also as a study tools. Application software is computer software that designed to help the user to perform singular or multiple related specific tasks. Depending on the activity for which the application was designed, an application can manipulate text, numbers, graphics or a combination of these elements. Examples include enterprise software, accounting software, office suites, graphics software and media players.

According to Mondofacto's dictionary, independent learning can be defined as programs of study developed for a particular learner who wishes to study alone [3]. Independent learning is also a proactive learning with or without other peoples' aid. Approximately, self-determining learning is same as self learning which both are about learn independently. Self-determining learning helps learners to make informal choices and to take responsibility for deciding what learners need to do in order to learn. One of the strategy features in independent learning is media consumption. Media expenditure is a method that can attract people to learn new things. Self education is a continuous learning process that people have to learn, unlearn and relearn to get things right.

Speech recognition is the process by which a computer or other type of machine identifies spoken words [4]. Speech technology offers great promise in the field of automated literacy and reading tutors especially for children. In such applications speech recognition can be used to track the reading position, distinguish oral reading miscues, assessing comprehension of the text being read by estimating if the prosodic structure of the speech is appropriate to the conversation structure of the story, or by engaging the user in interactive dialogs to assess and train comprehension [5]. Despite such promises, application that uses speech recognition also face some risk of higher error rates for user due to variability in vocal region of length, formant frequency, pronunciation and grammar.

Learning Quran independently needs the better guide to assure learner can understand and know how to recite Quran properly. Learning Quran is same as learning other foreign languages that need audio software in order to demo the right way to utter the words. Today, most of the Quran software is not user friendly. Some of the software just provides one way learning which user only can hear the pronunciation and could not measure the accuracy of user's articulation. Besides, most of independent learning software is just focus on learning foreign language and less for the teaching Quran.

As a solution, Independent Learning of Quran (ILoQ) is needed to design. ILoQ will be suitable for beginners and consist of  - Alif,  - Ba,  - Ta etc which are the basic words for the beginners before read the Quran. This software includes voice recognition as interactive learning. User could pronounce the word and will be measured by the software either the pronunciation is correct or not. So, learner will be more confident to recite Quran independently and accurately.

1.3 Aim and Objectives

The aim of this project is:

- To develop Independent Learning of Quran (ILoQ) software using voice recognition.

In order to achieve the aim, the objectives are:

- i. To do a research about speech recognition.
- ii. To identify and advance the media consumption that use for independent learning.
- iii. To search the way to measure the accuracy of Quran's alphabet pronunciation.
- iv. To develop interactive software considering the above three.

1.4 Scope

Basically, this software will focus on these scopes.

i. Software

- Windows XP Professional
- MATLAB 7.8.0

ii. Hardware

- Processor: Intel Core 2 Duo CPU
- RAM: 2.99GB, 993 MHz
- Audio Drivers (Microphone and Speaker)

iii. User

- Beginner in recite Quran.

iv. System

- ILoQ is developing base on speech recognition technologies.
- ILoQ focus on three education elements such as listening, reading and speaking.
- ILoQ is able to check the accuracy of reciting Arabic's alphabets.

1.5 Organization of Thesis

This thesis consists of 6 chapters. After this chapter 1 completed, chapter 2 was continued and followed by another chapters.

Chapter 2 is literature review that will explain the selected project. This chapter consists of two sub researches such as studies on current or existing system and explains about hardware and software requirement that have been used to implement this project. For this chapter, the related information can get via book, internet, article, journal and others.

The methodology used for developing the application will be provided in chapter three. Chapter 3 also details out the system development life cycle besides software and hardware specification that are needed for this project development.

Chapter 4 is implementation which all process that involved in development phase must be documented. In general, this topic explains project development that has designed. This chapter explains about coding of system and explanation about system.

Chapter 5 is result and discussion. This chapter will explain discovery or result analysis the facts such as result, project constraints and suggestion on how to fix this project.

The last chapter is conclusion. Chapter 6 will summarize about project that has developed. This chapter consists of the following information such as project summarization, data summarization and surveys that with objective and problem statement of this project. It also should consist of summarization of project's methodology, implementation that have used, suggestion or comment how to fix or improve this project for the future.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The introduction of speech recognition and independent learning has been discussed in the previous chapter. This chapter is about the literature reviews that have been reviewed based on the research topic. Literature is a report of a compilation and review means is in written forms on the research topic which write by those people with authority in the field.

As discussed before, learning Quran is same as learning foreign language which learner has to learn the basic of Quran's alphabet. From the study, some of the weakness and advantage of the language learning software in the market today has been discovered to be implemented and enhanced in the ILoQ development project. Besides, the methods and techniques that involve in speech recognition also will be discussed in this chapter.

2.2 Studies on Related Research

2.2.1 Design of Educational Software for Automatic Speech Recognition (ASR) Technique [6]

i. Introduction

This research is to develop a set of interactive and practical education software to guide and assist students in studying, and performing experiments for speech recognition. This research is proposed for the improvement over the limitations and problems of the traditional teaching method in speech recognition. This educational toolkit may be able to help students and others in getting an appropriate starting knowledge and experience in speech recognition topics.

ii. Speech Recognizer Overview

This software is including several speech recognition techniques to concentrate on such as Linear Predictive Coding (LPC), LPC-based Cepstrum, and Mel-Frequency Cepstrum Coefficients (MFCC). These three feature extraction techniques are functional to reducing the rate which is apply in the first stage of processing. Vector Quantization (VQ) also is integrated in this toolkit. VQ is to reduce storage and computation time for spectral analysis information.

Dynamic Time Warping (DTW) technique is a time-alignment algorithm that can handle variety rate of word pattern problems. DTW is use to compare the speech pattern in order to determine the similarity. Hidden Markov model (HMM) is one of the statistical model of characterizing the spectral properties of the frames of a pattern also is one of the speech education learning in this toolkit.

Figure 2.1 shows the overview of speech recognizer training process. For speech recognizer training purposes, the digitized speech samples are processed to produce speech features by using the feature extraction technique. Speech features

can go through VQ stage to create a sequence of codebook indexes. Then, the speech features or codebook indexes will be stored in the template library.

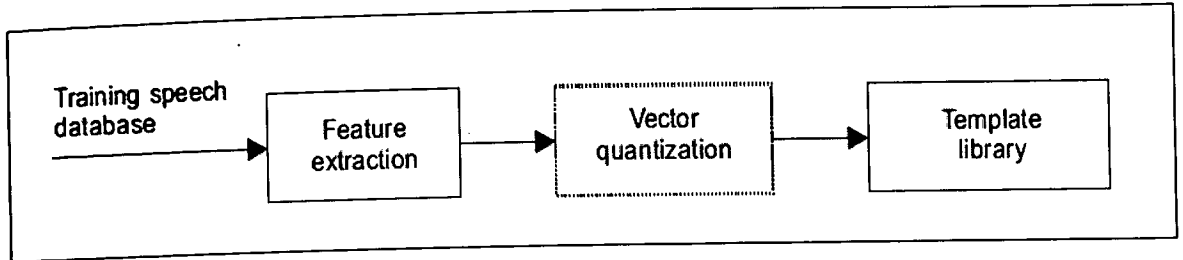


Figure 2.1: Overview of speech recognizer training process.

Figure 2.2 illustrate the recognition purposes which is the process to produce the speech feature are same with speech recognizer process. The features or indexes will be compared with the indexes that stored in the template library. The best-matched template will be determined so that an appropriate action can be done. However, VQ is an optional step for DTW matching technique but is essential for discrete density HMM.

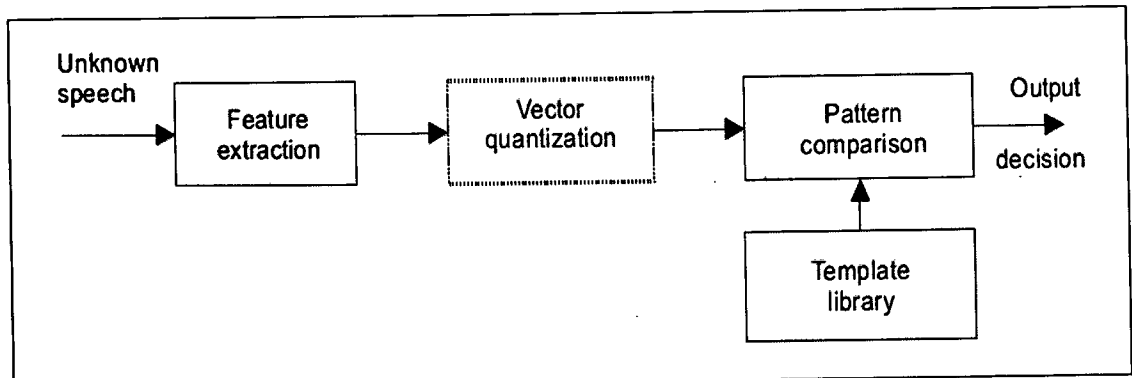


Figure 2.2: Overview of speech recognition.

iii. Interface and Software Design

Figure 2.3, 2.4 and 2.5 are shown the design of software's interfaces.

- Main module.

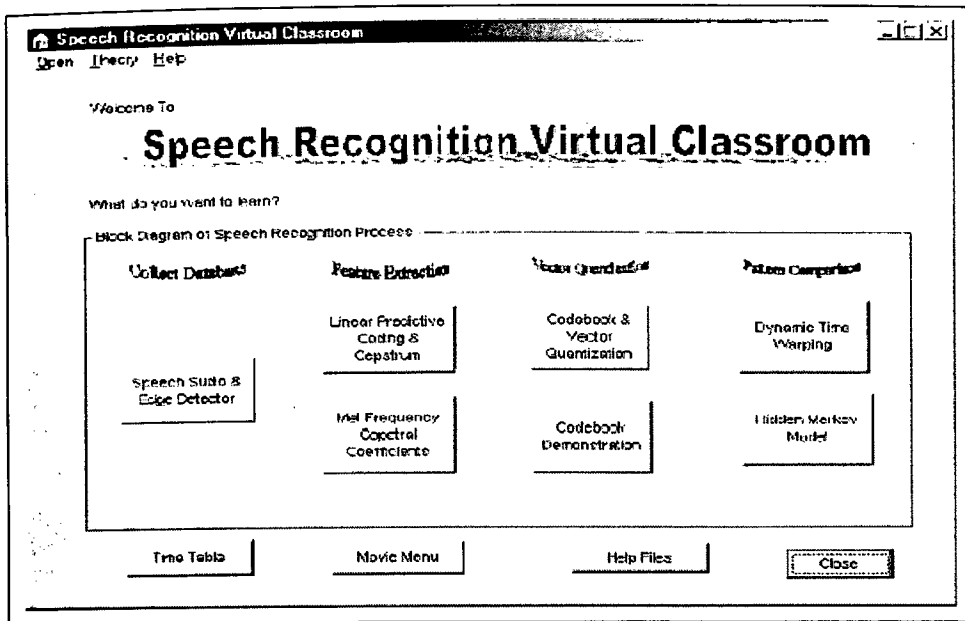


Figure 2.3: GUI of main module.

- Speech studio.

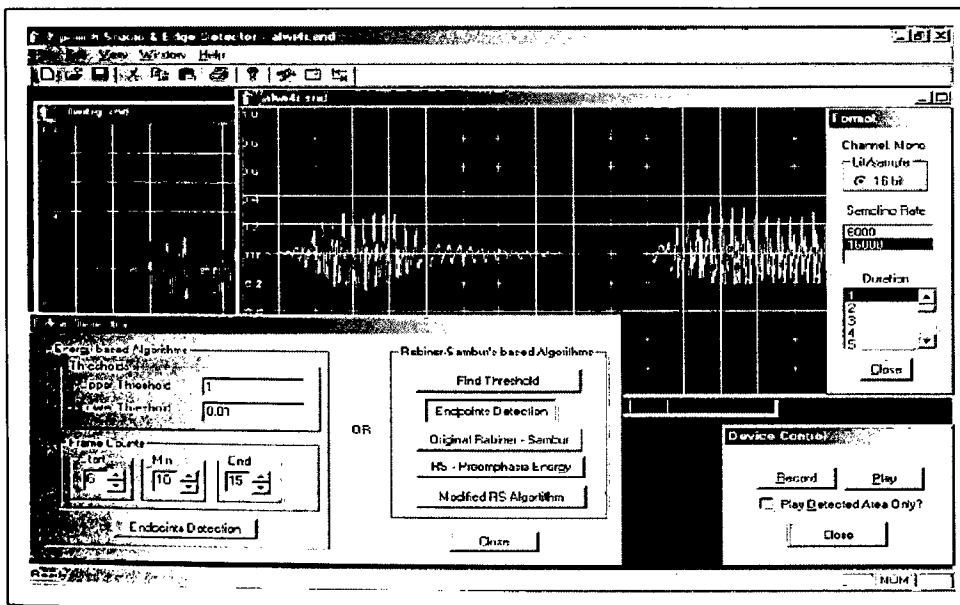


Figure 2.4: Design of Speech Studio interface.

- VQ classroom.

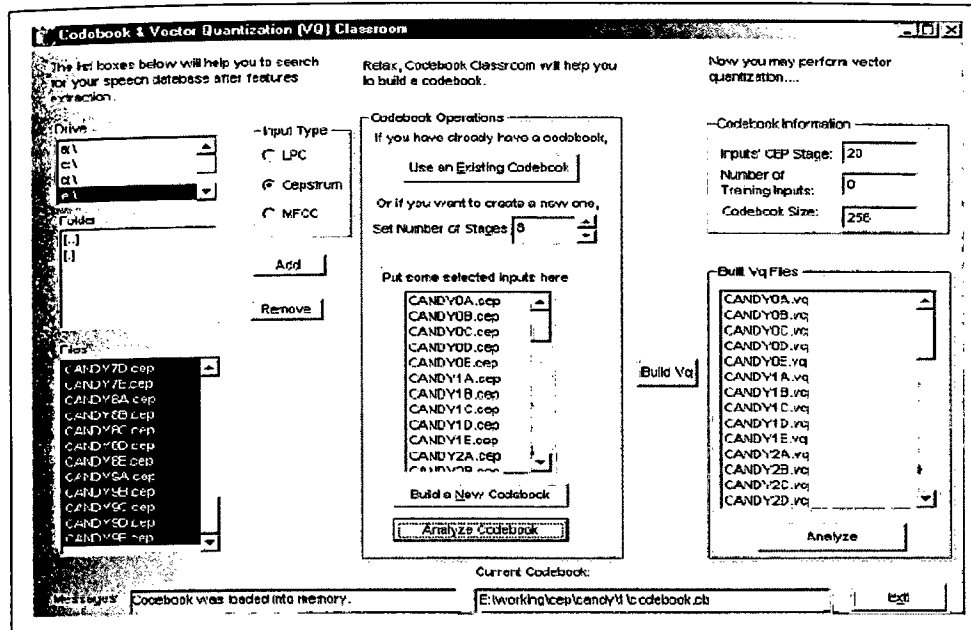


Figure 2.5: Interface design of VQ Classroom.

2.2.2 Adaptation Techniques for Children's Speech Recognition [7]

i. Introduction

This article is discussed about the suitable technique to apply for the children's speech recognition application. Currently, most of speech software and applications have been trained and evaluated on adult speech. There are some differences between adult and children's speech. Children have higher pitched voices than adults due to shorter vocal tract. Besides, the experienced in articulating sounds like adults is less and cause the difference when pronounce.

ii. Adaption techniques

A speech recognizer might be divided into four parts: the digitizing, feature extraction, acoustic modeling and language modeling part. Appropriate methods and techniques are important in order to normalize some spectral characteristics of the recorded speech, adapt the acoustic models to better speech match and also to reduce the sensitivity against background noise.

- Vocal tract length normalization (VTLN)

Each person has different length of vocal tract. Short vocal tract tend normally has a high-pitched voice and vice versa. These differences will cause the acoustic mismatch problems and affect the performance. VTLN function as a method to extend the tube by a factor alpha scale the frequency of the spectrum.

Table 2.1 shows the word accuracies between children and adults recognizers after the test has been done. Both are improve after applying VTLN.

Table 2.1: Result of accuracy tested to different age.

Set-up	Word accuracy at an age of 7 years	Word accuracy at an age of 13 years
Recogniser for adults	62%	95%
Adult recogniser + VTLN	76%	96%
Recogniser for children	85%	96%
Child recogniser + VTLN	89%	97%

- Voice transformation

Voice transformation technique is use to normalize the signal. The word error rate experiment has been done between adults and children. The result shows that children has higher error rate which is 43% compared to adults, 19%. After voice transformation has implement, the rate for children reduce to 31%.

2.3 Study on Current Existing Software

2.3.1 Tell Me More [8, 9, 15]

TELL ME MORE language programs are software packages and can be ordered as CD-ROM sets (with free headset) or as online courses from Auralog. Each package is inclusive and covers basic, intermediate, and advanced learner levels. Languages that available in this software are, Spanish, English, French, German, Italian, Dutch, Chinese, Japanese and Arabic.

Tell Me More is full of graphics, photographs, sound bites and videos and also the speech recognition software. This software does an outstanding job of picking up on the tonal mistakes made when user say the words. This language learning software can even show user how to form the words by using diagrams if users are having difficulty. The resulting sound into the microphone can be compared and displayed as a voice pattern alongside the native speaker's. The voice pattern is shown in Figure 2.6.

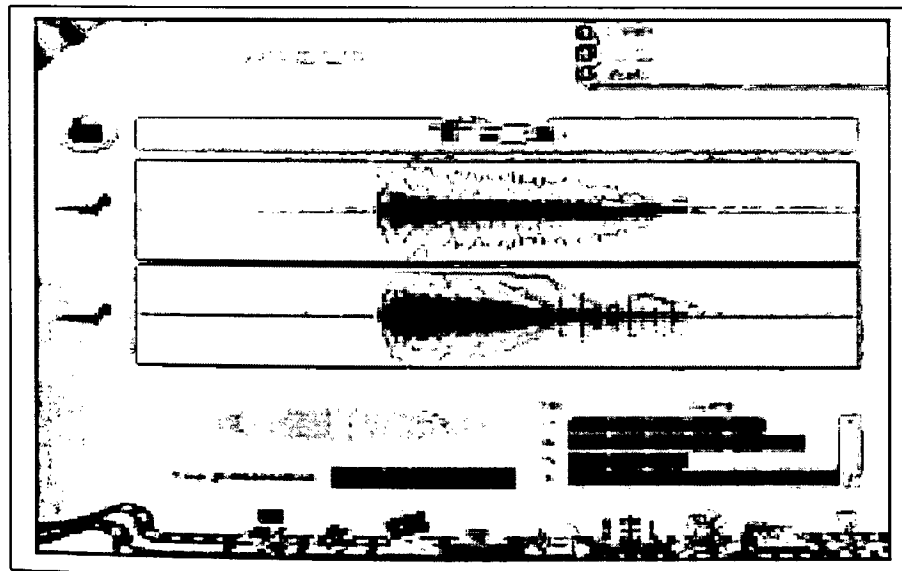


Figure 2.6: Voice pattern in Tell Me More.

The questions in the dialogue section are based on various images of objects and actions in TELL ME MORE software. User will responds to each question by