VOICE SYNTHESIZER SYSTEM USING MICROCONTROLLER

NOOR KAMILA ISMAIL

(2005509311)

BACHELOR OF ENGINEERING (HONS) ELECTRICAL

UNIVERSITI TEKNOLOGI MARA (UiTM)

NOVEMBER 2008

ACKNOWLEDGEMENTS

Without the help and support of a number of people it would not have been possible for me to finish this project. First and foremost, I would like to express my gratitude to Allah the Almighty for giving me the opportunity to finish this project. I am greatly indebted to my supervisor, *Assoc. Prof. Mohd Uzir Kamaludin*, for providing his kind guidance throughout the development of this study. His comments have been of greatest help at all times. He has also provided all the necessary resources such as the equipments I needed throughout this project.

This episode of acknowledgement would not be complete without mentioning my parents, who taught me the value of hard work by their own example. They rendered me enormous support during the whole tenure of my research.

Finally, I would like to thank all whose direct and indirect support helped me complete my theses in time, who contributed to this project, especially my fellow classmates, for their help and support.

ABSTRACT

This thesis is prepared for the Voice Synthesizer System using Microcontroller project chosen for Final Year Project II.

Objective: This study is carried out to design and implement a system that can produce synthetic voice using microcontroller PIC 16F877. It is also done to investigate the quality of the voice produce by the UM5100 voice processor chip. This project is also carried out to test and determine how to generate words such as "hello", "good" and "thank you" using this system.

Methods: The research methodology proposed for this project starts with the hardware development, writing a program to produce data needed to experiment with the voice processor chip and then loading the program and storing it into the PIC. Later, a system testing is done for data collection and lastly documenting the results into technical papers, thesis and presentation slides.

Results: From the test results collected, it can be concluded that the data output using delay method failed to produce the expected output. Using the data output using READ line delay method, it can be concluded that each 8-bit data sends to the voice processor chip successfully produced the expected 128 allophones sound. There are 256 voice sounds that can be produced by the voice processor. Repetition at $80 shows that there are 128 allophones can be recognized. The allophones waveforms are impossible to be mapped since there are 128 varieties of sounds and the lack of sources to analyze the waveforms. Production of words is impossible at this stage because the allophones are not clearly tabulated.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1.0 Introduction to Speech Synthesis

Speech synthesis is the process of determining the operating rules and techniques for reproducing speech. [1] It can also be interpreted as the artificial production of human speech. A system that is used for this purpose be it in the form of a software or a hardware, is called a speech synthesizer. Speech synthesis is first motivated by a desire to improve the efficiency of digital communication. This was done in the second half of the 18^{th} century. For decades, scientist have studied the human voice and tried to produce it artificially using different mechanical and electrical models. During the 1970's, further development of speech synthesis was closely associated with computer technology in general.

Today, the software based speech synthesizer synthesized speech by concatenating pieces of recorded speech that is stored in a database. Each of these systems is differed by its size of stored speech units. A speech unit is the system that stores phones and diphones for the speech synthesizer system. Phones are single speech sounds. These sounds cannot be successfully concatenated into words and sentences, since the acoustic properties of these minimal distinctive segments of speech vary as a function of their context. [3] This variation is necessary for intelligibility and naturalness. Diphones is an adjacent pair of phones or the transition between a pair of phones. It consists of the second half of one speech sound and the first half of the subsequent which results in a large number of elements. These elements have to be carefully selected to achieve a high degree of naturalness, even without having a complete description and understanding of the acoustics of speech production. However, these methods lack the flexibility of synthesis by rule which does not use