Higher Technical Institute
ELECTRICAL ENGINEERING DEPARTMENT

DIPLOMA PROJECT

VOICE MANIPULATION

BY
ZANTIS GEORGIOS

E/1181

JUNE 1999

HIGHER TECHNICAL INSTITUTE

ELECTRICAL ENGINEERING DEPARTMENT

DIPLOMA PROJECT

VOICE MANIPULATION

E - 1181

ZANTIS GEORGIOS

JUNE 1999

Contents

Acknowledgments

I would like to especially thanks my parents for giving me all the time I needed to fulfill the project requirements.

Also, I would like to dedicate this project to all my friends I met over Internet Relay Chat (IRC) for their encouraging support.

And, finally, I would like to thanks my HTI supervisor Mr. Theopemptou Charalambos for his endurance over my huge number of questions I asked him.

Summary

Title : Voice Manipulation

From : Zantis George

Like it or not, the speech it the way human beings are able to communicate. And since the term speech exist it can be analyzed and modified like every other thing seen by people. A lot of natural sounds onced needed to be used for film production, documentary, shows and a lot more. The exact purpose of this project is to provide these sound without the need to be on location of which the required sound is produced. Furthermore, these place might be inappropriate for the purpose they are needed so the final product wouldn't be able to be produced. In order to construct such a device, there is a need to understand how these sounds are initially created and simulate the nature phenomenon with an equivalent electrical device. In addition to this, some sounds don't even exist in nature and they are man-made in order to please of human hearing utilizing the full range of it with beautiful sounds. So, at the end we manage, using technology, to save money from having to go on location to do our job and find new ways of entertainment. At the end of the line we have made our life simpler and more convenient for our own good.

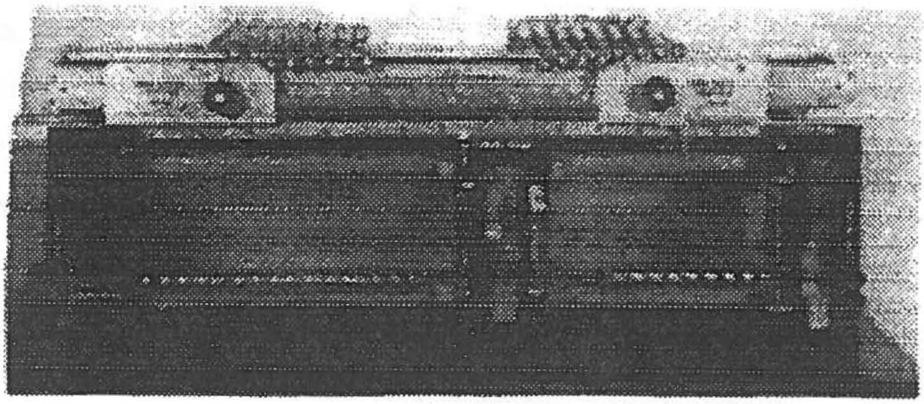

figure 1 A front view of the Electrical Vocal Track developed by H. K. Dunn (1950).

Introduction

Generally speaking, speech processing relies on basic research in the speech and hearing sciences, some of which is centuries old, and much of which is ongoing. Only a part of the speech processing engineers have the time, opportunities or knowledge to become expert in these fundamental sciences, so the field of speech processing remains inherently open for more research. Nevertheless, the speech processing engineer needs a sound working knowledge of basic concepts from these areas in order to intelligently analyze and model speech, and to discuss findings with researchers in other fields.

In order for communication to take place, a speaker must produce a speech signal in the form of a sound pressure wave that travels from the speaker's mouth to a listener's ears. Although we focus on the production of speech, hearing is an integral part of the so-called speech chain. At the same time, the talker can continuously monitor and control the vocal air pressure he creates by receiving a feedback from his ears. Any delay in this feedback may cause difficulty in proper speech production. A utilization of the delay in sound production is appeared in the echo and reverberation effects. Of course here the talker has the direct feedback of his voice but extra sounds of the original wave and mixed together. So, when the talker say something, he get the required feedback to control his voice in addition of a series of similar wave coming out after some delay. This is so called the echo effect and it can be heard at open places of which there is something to reflect the sound. An approach used in the past to simulate this effect was with the use of a rotating continuous magnetic tape. This tape was passing from a magnetic head which was writing the sound. Then, there was an other head reading back the sound and playing it after the original. For the reverberation effect, a series of reading heads were used. By mixing the original sound with the reproduced one, the reverberation effect was created. An other gadget invented in the late 30's was the Voder. This unit could produce speech with its 14 keys, energy switch and pitch-control pedal. Later, the computers started to appear and speech production was easy and could be heard as clear as normal speech. In addition to this, special ICs (integrated circuits) were constructed just to be able to handle frequencies in the range of human speech (0,1 - 3,4 Khz). These ICs, could modify directly the source voice and