



VOICE ASSISTANCE FOR PEOPLE WITH SPEAKING ABNORMALITIES

Jyothish Yadav M, Anooj H.N, Mathews K.Varghese, Surekha Mariam Varghese
Computer Science and Engineering Department,
Mar Athanasius College of Engineering,
Kothamangalam, Kerala, India

Abstract— In our society we see around a lot of people having problem with speaking. They do talk, but the actual content and the meaning of the words sometimes gets diverged. This is because they do suffer from problems like twisted tongue, abnormal lips, tooth etc. Some people have voice problems by birth. This can be due to defectiveness in any of the organs influencing the production of sound. It is difficult for other people to understand speeches made by this affected person. An application which understands this speech and produces the correct output has been developed.

Keywords— Fingerprint, Training, Feature points

I. INTRODUCTION

In our society we see around a lot of people having problem with speaking. Some people have voice problems by birth. This can be due to defectiveness in any of the organs influencing the production of sound. It is difficult for other people to understand speeches made by this affected person. An application which understands this speech and produces the correct output has been developed. An application which understands this speech is developed. It takes the affected persons voice as input and produces correct speech as text and voice. Here the input from the user is collected and trained to the system. Training means a fingerprint of the audio sample is created by using a fingerprinting algorithm. Later upon recognition the fingerprints are compared and output is decided. This output is given as both text and speech which can be understood by a normal person. The voice of output can be male or female as per need. Usually the sound is given as that of the system that it works. It thus makes the life of the affected person much easier as he can now easily interact with normal people as they can understand what he speaks.

II. SYSTEM DESIGN

A. Design of Input –

Human generated voice is converted into computer format. Correct input design have to be beginning for the correct working of the system else it will produce an inaccurate result. Input which has error need to be controlled.

Data must be entered in the system in a simple format. The main goal is to enter the data without any error and in a simple way. Here the method in which input is given to the system should be decided. Microphone is the most common.

When giving input it should be given carefully. Considering the cost required for the system it is the collection of input which forms an integral part. Everyone users interact first with this input method so it can lead to error if not good.

The samples of an affected person are collected and train the samples. Train the samples and convert them into fingerprints. An option called train database is provided to train the database. On clicking this option can train the database. After training the database the input voice can be given in two ways. It can be given through a microphone or a saved file in the system.

B. Design of Output –

The output to be produced should be in an organized and an easily understandable manner. On only producing the right output people will understand if the system works properly. Output should be clear as the person listening to it should understand it.

GUI is the main output of the system. GUI is the interface that appears before doing any operation. It makes the program even more attractive. In Java, GUI interface is created by using some commands. Functions that are written and associated with a specific component or with the GUI figure are called callbacks. Callbacks control how your GUI responds to events such as button clicks, slider movements, and menu item selections. Specify the primary callback function for a component using the Callback property. Guide command written on the command prompt will create new GUI and then can be designed according to user customization.



III. EXPERIMENT AND RESULT

A. Reading Audio

Using microphone we extract the audio required which is the primary task. A file in the system can also be used.

B. Spectrograms

The first thing to do after getting the input is deciding how it should be processed: FFT and spectrograms. The audio currently entered using the microphone is of no use in its current format which contains only amplitude vs time. The frequencies needed should be understood from the input data. It forms the integral part of audio fingerprinting.

To draw input as a waveform it is passed through an FFT. It then showed the magnitudes for different frequencies for chunks of the time domain data.

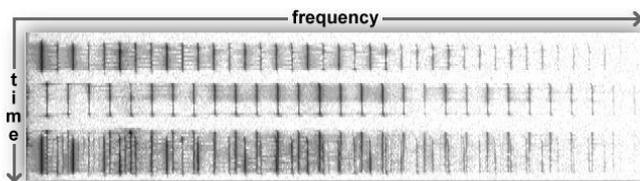


Fig 1. Sample of a spectrogram

C. Reading in wave files

After reading the audio from the microphone it is converted to a code from which the data could be directly read.

First duplicate the rendering class of spectrogram and modify the signal so as to derive the feature points which is used for comparison. A high score system with a simple band where the frequency domain data is split into group is derived. The frequency with highest magnitude is found out by looping each band. A feature point is one that has a high magnitude. Two images are saved which shows the feature points of a spectrogram to see if they could be compared after hindrance. To find the feature points which overlay these two were compared using appropriate software. An example of comparing feature points is given below where green is the points which overlay perfectly.

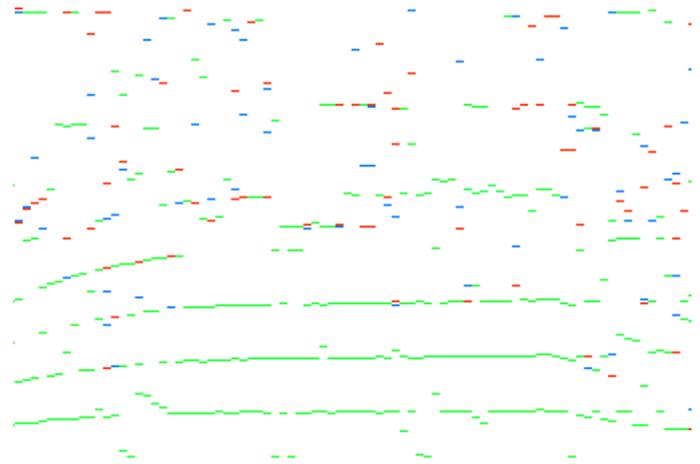


Fig 2. Comparing feature points

D. Producing output

Samples of affected person is collected and converted into fingerprint using the above method and stored. The input is given and the input is also converted to fingerprint using the same method. The input audio fingerprint is compared to with every sample fingerprint and the sample with maximum similarity is given as output as speech and text. The sample input file consists of enough data to produce the correct speech and text.

IV. CONCLUSION

One of the most difficult situations for a person is the difficulty in communicating with another person due to birth defects or any accidents. This project helps them in overcoming this issue. The device developed helps the person to easily overcome this difficulty. The person has to speak to the device and the device converts his speech to fingerprints and compares with all the fingerprints in the most matching cluster. The fingerprint with most matching percentage is given as output as both speech and text as the other person can understand easily.

V. REFERENCE

- [1] CLOUDGARDEN. 2010. "Jigloo SWT/Swing GUI Builder for Eclipse and WebSphere". [online].