SIGNIFICANCE OF ENERGY DISTRIBUTION USING PRAAT

Sohail 17BEC0187

Abid Ali 17BEC0513

Hussain 17BEC0863

Praat is an open source program created by Paul Boersma and David Weenink, from the Institute of Phonetic Sciences-University of Amsterdam

This Program allows users to visualise, annotate and extract the components of signal in a very simple way

Here, we have extracted MFCCs of a simple voice signal to extract its unique component which varies with respect to person to person

For this we should understand how MFCCs are extracted

1. First the Signal is Divided into frames with a rate of 25ms or (20-40ms)
2. The we find the energy spectral density which gives us an idea of power existing at each frame
3. Then the signal is made to pass through Mel filter banks which filter out unnecessary frequencies
4. Show the perception of sound by simulating a Human Ear
5. Reduce the noise
6. Compress the signal to an extent
7. Then Log operation is performed on the signal as the change in a signal is not perceived in a linear way, its rather exponential
8. Then the frames are combined but this leads to overlapping so, we need to discretise(perform DCT) the signal so that the overlap is ignored and we get a limited number of co-efficients

BLOCK DIAGRAM

How to extract MFCCs using PRAAT:

1) Praat performs all these operations and displays the coefficients

2) For this we first select New-> Record mono sound for the real time observation or open an audio file from the Open menu

3) Then select the object and select the option view and edit

4) You will observe a dialog box showcasing spectrogram along the Amplitude/time plot of the signal

5) Now Using Cut command from the Edit menu we can remove the silent parts (almost no amplitude) of the Signal to increase the accuracy of the MFCCs

4) After editing the object, Select the object then click Analyse spectrum, then MFCC

5) Then specify the windowing or framing length of each window along with the specification of the Mel filter bank

6) Then click ok, then select inspect

7) The parameters used along with the number of frames and the first co efficient is displayed

8) To view the MFCCs of the each frame select open at each frame.

For the step by step analysis follow this link:

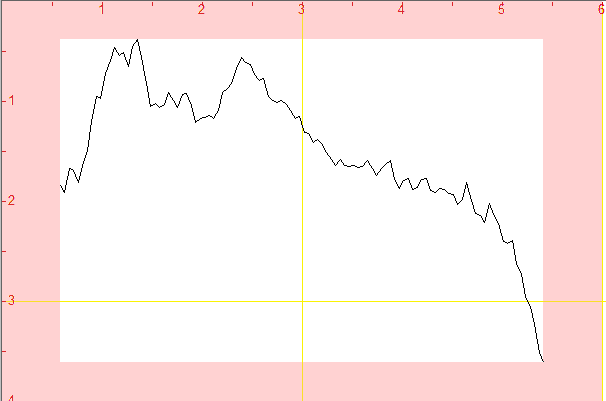
EXPECTED RESULT:

A set of coefficients which will be unique for every person's voice can be utilized for identification of a person's voice.

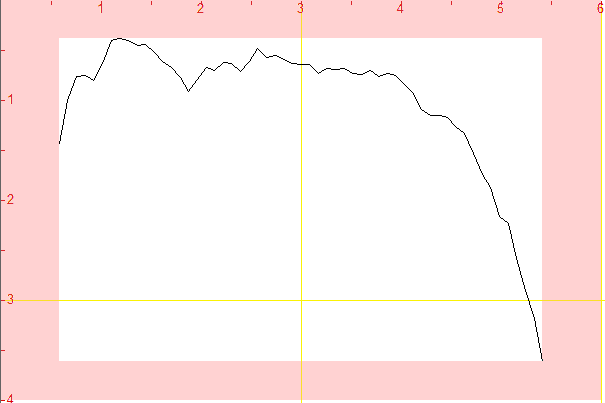
RESULT:

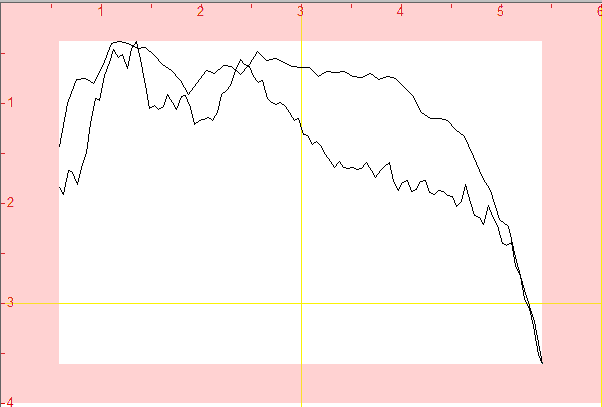
After taking two voice samples and after trimming them we put through the following steps: Then for the sample

Male-“hello” we plotted the amplitudes with respect to time and thus was the graph obtained:



And for Female-“hello” we plotted the amplitudes with respect to time and thus was the graph obtained:

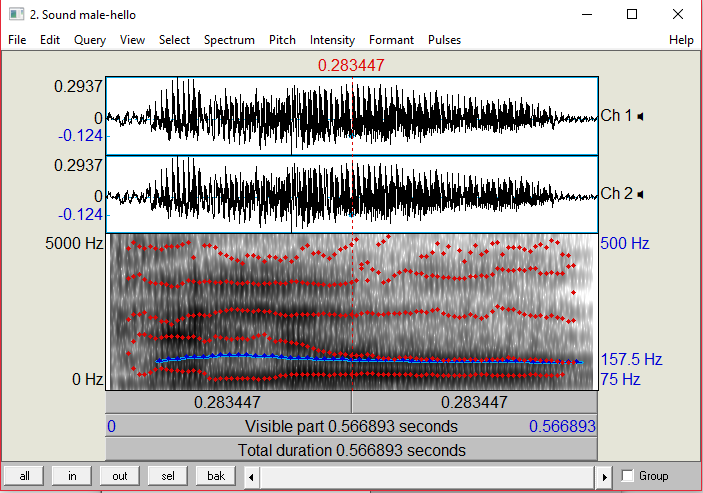


And as we compared the both

There was a lot of difference between its coefficients. Hence its uniqueness.

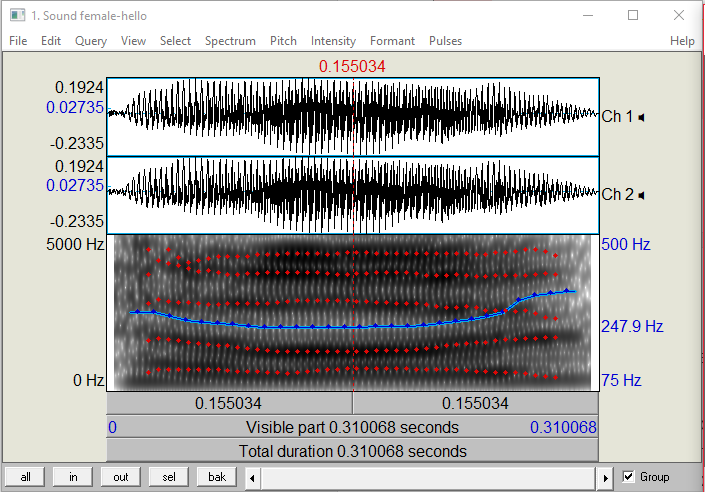
As you see the MFCCs are the operated coefficients of Power Spectral Density of the Signal these also depend upon the power present at a particular frequency leading us to Significance of Energy with respect to the signal as we see the MFCCs and LPCCs actually represent the Energy distribution with respect to the Mel frequency if we look in the spectrogram of the audio samples we see

For male sample:



The red dots represent the formants and the blue represents the intensity.

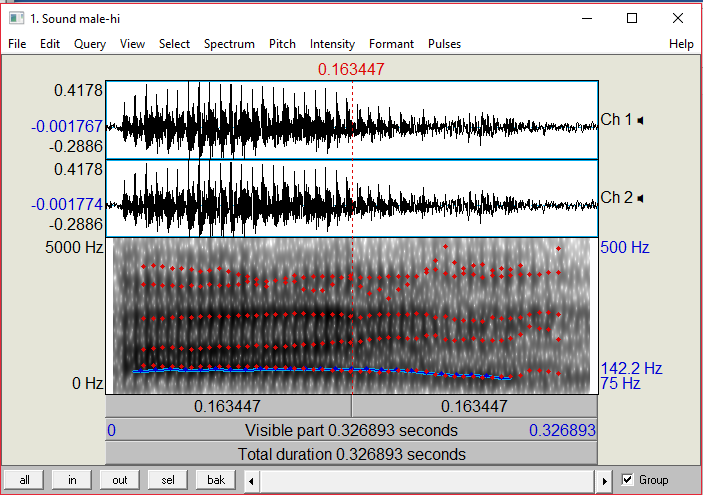
For Female sample:



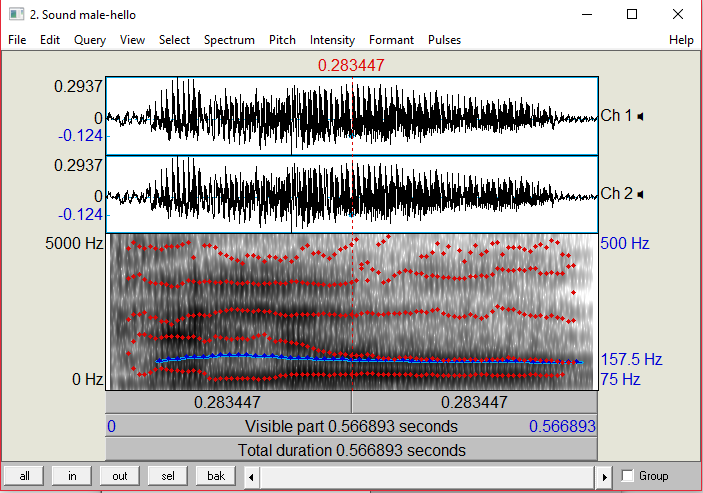
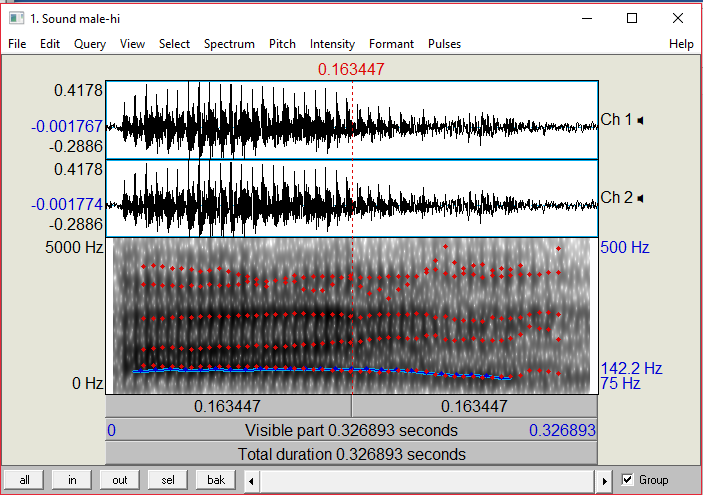
The red dots represent the formants and the blue represents the intensity.

As we can see there is no much difference in the alignment of formants. Because same word has been spoken by different people

But if we change the word the position of the formants and acoustic energy concentration varies across the frequencies Due to which we are able speak different words. To prove this we took a different audio sample of a male saying “HI” and observed its spectrogram:



And if we compare it side by side to the spectrogram of hello we see different alignment of these formants



Hence signifying the importance of energy across the signal