CS : 215
Signal & Data Communication Laboratory

Experiment: VII-B

Mritunjay Aman (1912170)

14.04.2021

Spectral Analysis

Aim

To use Spectral Decomposition and Analysis to recognize the vowel present in the given sample audio.

Theoretical Background

Formants In speech science and phonetics, a formant is the broad spectral maximum that results from an acoustic resonance of the human vocal tract. In acoustics, a formant is usually defined as a broad peak, or local maximum, in the spectrum.

Vowels can uniquely be defined using a series of formant frequency with suitable variance.

Discrete Fourier Transform
The Discrete Fourier Transforms converts a sequence of Complex Numbers $\{x_1, x_2, \ldots, x_n\}$ into another sequence of Complex Numbers $\{X_1, X_2, \ldots, X_n\}$ and is defined as :

$$X_n = \sum_{n=0}^{N-1} x_n . \exp \frac{j2\pi k}{N} n$$

Discrete Fourier transforms (DFTs) are extremely useful because they reveal periodicities in input data as well as the relative strengths of any periodic components. And thus they are used to covert Signal to its Frequency Domain for efficient Analysis.
The fast Fourier transform is a particularly efficient algorithm for performing discrete Fourier transforms of samples containing certain numbers of points.

Methodology

1. The Audio Sample is read for the given Sample Range { 21000-26000 }.
2. The first 100 samples of the Signal obtained is plotted against time.
3. The DFT of the Signal is computed and plotted.
4. The formant Frequency and time distribution curve is they compared with that of other vowels and conclusion is drawn.

Code

```
clc;
clear all;

% Reading the Audio File
```

```
load handel.mat;
y1 = y;
filename = "handel.wav";
audiowrite(filename,y,Fs);
samples = [21000,26000];
clear y Fs;
[y, Fs = audioread(’handel.wav’, samples);
sound(y, Fs);
info = audioinfo(’handel.wav’);

% Plotting the Samples
n = 26000 - 21000 + 1 ;
t = (0:n-1)/Fs;
stem(t, y, ’b.’, ’linewidth’, 1.5);
title("First 100 Samples of the audio sample handel");
ylabel("Audio Signal");
xlabel("Time");
set(gca, ...
      ’Box’, ’on’, ...
      ’Ytick’, [-0.8:0.2:0.8], ...
      ’TickDir’, ’in’, ...
      ’XGrid’, ’on’, ...
      ’YGrid’, ’on’, ...
      ’FontSize’, 15, ...
      ’FontName’, ’Calibri’)

axis([-0.0005 100/Fs-1/Fs -0.8 0.8]);
```

```
% Fourier Tansform and Signal Spectrogram
fTransf = abs(fft(y));
freq = (0:n-1)*Fs/n;
stem(freq, fTransf, 'k.', 'Marker', 'None', 'Linewidth', 1.5);
title("Magintude Spectrum of the Sample");
ylabel("Magnitude");
xlabel("Frequency(Hz)");
set(gca, ...
      'Box', 'off', ...
      'TickDir', 'out', ...
      'YGrid', 'on', ...
      'FontSize', 12,...
      'FontName', 'Calibri');
axis([0 8000 0 250]);

% Fourier Tansform and Signal Spectrogram
fTransf = abs(fft(y));
freq = (0:n-1)*Fs/n;
subplot(2,1,1);
stem(freq, fTransf, 'k.', 'Marker', 'None', 'Linewidth', 1.5);
title("Magintude Spectrum of the Sample");
ylabel("Magnitude");
set(gca, ...
      'Box', 'off', ...
      'TickDir', 'out', ...
      'YGrid', 'on', ...
      'FontName', 'Calibri');
axis([0 4650 0 250]);
```

```
subplot(2,1,2);
segmentlen = 100;
noverlap = 0;
NFFT = 200;
colormap('jet');
sample = y1(21000:26000);
spectrogram(sample, segmentlen, noverlap, NFFT, 'xaxis', Fs); w
```

Input Data Description

Sample audio file : 'handel.mat'
Range: 21e+3 - 26e+
Spectrum Range = less tha 8000Hz

Plots

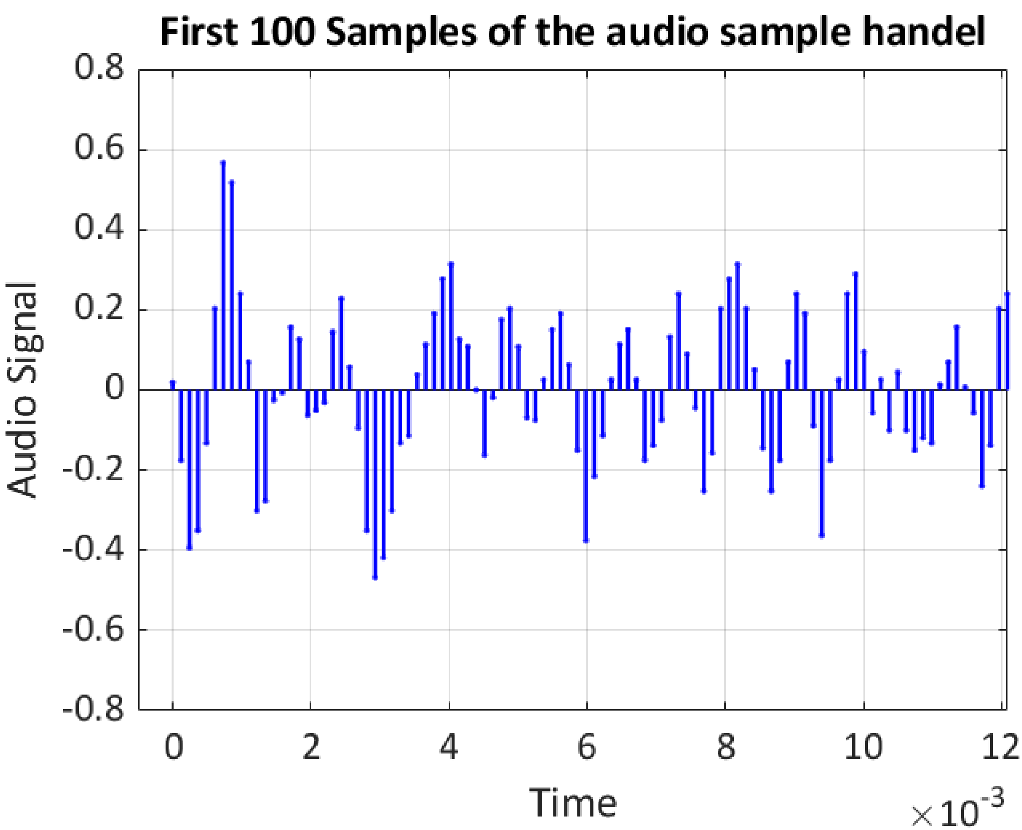


Figure 1: First 100 samples of given audio

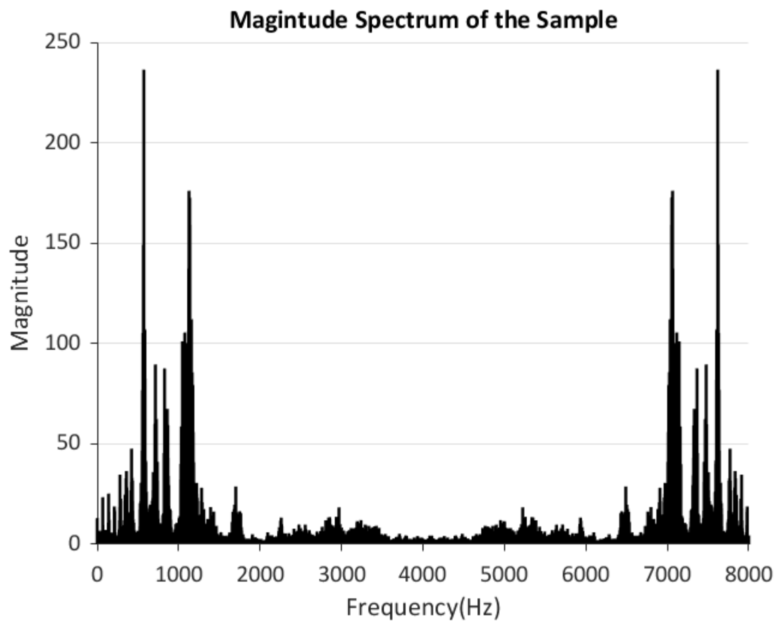


Figure 2: Spectrum for Frequency < 8KHz

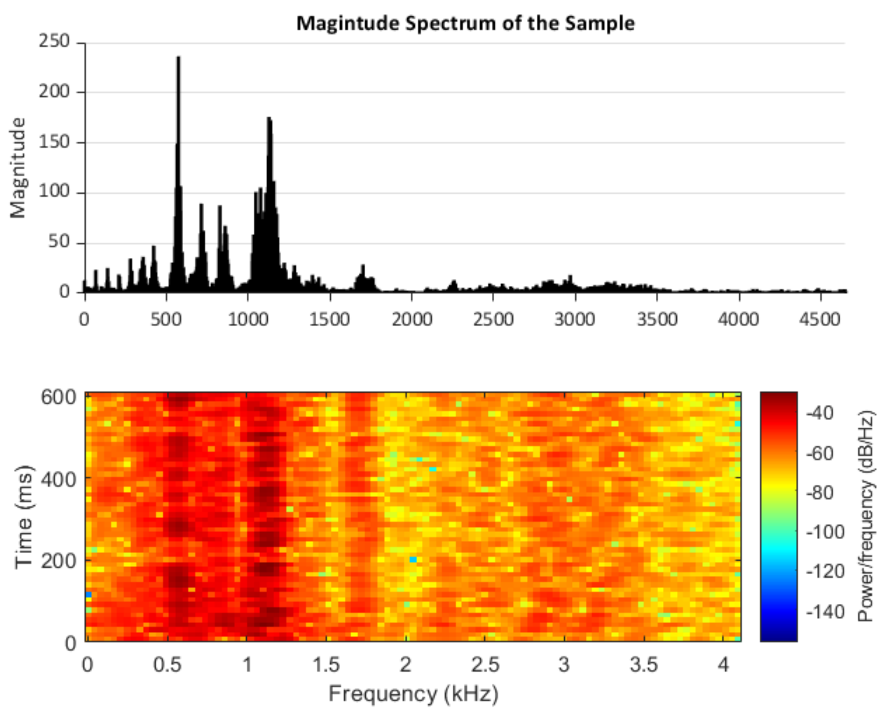


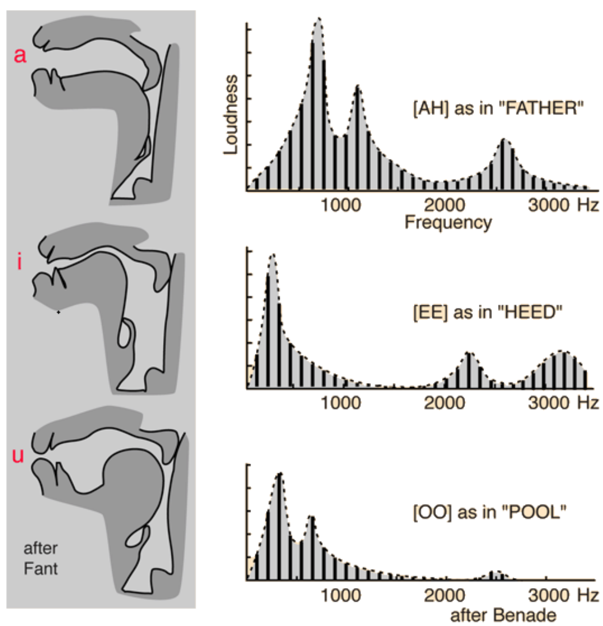
Figure 3: Broad peaks shown using Spectrogram for varying Frequencies

Result

Vowel	F1(Hz)	F2(Hz)	F3(Hz)
iː	280	2620	3380
ɪ	360	2220	2960
e	600	2060	2840
æ	800	1760	2500
ʌ	760	1320	2500
ɑː	740	1180	2640
ɒ	560	920	2560
ɔː	480	760	2620
ʊ	380	940	2300
uː	320	920	2200
ɜː	560	1480	2520

Adult male formant frequencies in Hertz collected by J.C.Wells around 1960.
Note how F1 and F2 vary more than F3.

URL : https://www.phon.ucl.ac.uk/courses/spsci/iss/week5.php



URL : http://hyperphysics.phy-astr.gsu.edu/hbase/Music/vowel.html

Comparing the spectrogram and Magnitude Spectrum obtained earlier against this chart of formant values, we conclude that the vowel present in audio is similar to $/a:/$. The frequency curve also closely matches to that of "ah" formant.

Conclusion

- The audio on focused listening gives an "ah" sound and thus closely resembles $/a:/$ formant.
- Also, the frequency values centre around F1 = 700Hz and F2 = 1200Hz, and thereby confirm our hypothesis. The slight variations encountered in F1, F2, and F3 is attributed the high pitch nature of audio sample and the fact that sample is taken from opera and subject is singing.
- **URL** : https://www.phon.ucl.ac.uk/courses/spsci/iss/week5.php
- **URL** : http://hyperphysics.phy-astr.gsu.edu/hbase/Music/vowel.html