NATIONAL INSTITUTE OF TECHNOLOGY SILCHAR

Cachar, Assam

B.Tech. IVth Sem

Subject Code: CS215

Subject Name: Signals and Data Communication

Submitted By:

Name : Subhojit Ghimire

Sch. Id. : 1912160

Branch : CSE – B

1. **Load the "handle.mat" dataset using "load handle.mat" command in MATLAB. Set "y" equal to samples 21000 to 26000 from the dataset. This part of the sound file is a vowel sound. Run a spectral analysis to determine which of several vowels this part of the sound file is. Run a script to see the spectrum of the vowel sound. The script should plot two parts; first shall plot first 100 samples of "y" versus t and the second part should plot magnitude of FFT of "y" versus values of frequencies from 0 to just less than 8000Hz. Compare the spectrum to the published spectra for several vowels.**
   Your report should comment about what vowel you think the dataset is, and then listen to the dataset (use "sound") and comment again on what vowel you think the dataset is. Your report should also mentions about URL of the published spectra of vowel used by you.

➔ **AIM: TO PERFORM SPECTRAL DECOMPOSITION ON AN AUDIO SAMPLE AND IDENTIFY THE VOWELS OCCURING IN IT.**

THEORITICAL BACKGROUND:

Discrete FT: The discrete Fourier transform which transforms a sequence of N complex numbers $x_n = x_0, x_1, \ldots, x_{N-1}$ into another sequence of complex numbers $X_k = X_0, X_1, \ldots, X_{N-1}$ is defined by:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot \exp -\frac{j2\pi}{N} kn$$

Usually, the x_n are samples of a signal and the DFT shows the frequency density of the signal.

Formants: Formants are defined as broad peaks or local maximums in the frequency spectrums. Humans, despite having different vocal tract lengths (and hence, different resonant frequencies), can produce sounds perceived to be in the same category.
To solve this solution, Hermann coined the term 'formant'. A vowel can be uniquely defined by a series of formant frequency (with acceptable variance).

METHODOLOGY:

1. The first 100 samples of handle.mat are plotted.
2. The fourier transform and spectrogram of the samples are computed and plotted.
3. The coefficients are calculated for a small interval using LPC function and are then solved to find complex roots.
4. The frequency and the corresponding bandwidth are calculated and sorted from the complex roots (only those having positive imaginary parts).
5. The triplet values F_1, F_2 and F_3 are compared against different vowel formant data.

CODE:

```
clear all;
clc;

load handel;
y = y';

vow = y(21000:26000);
time = (0:length(vow)-1)/Fs;
```

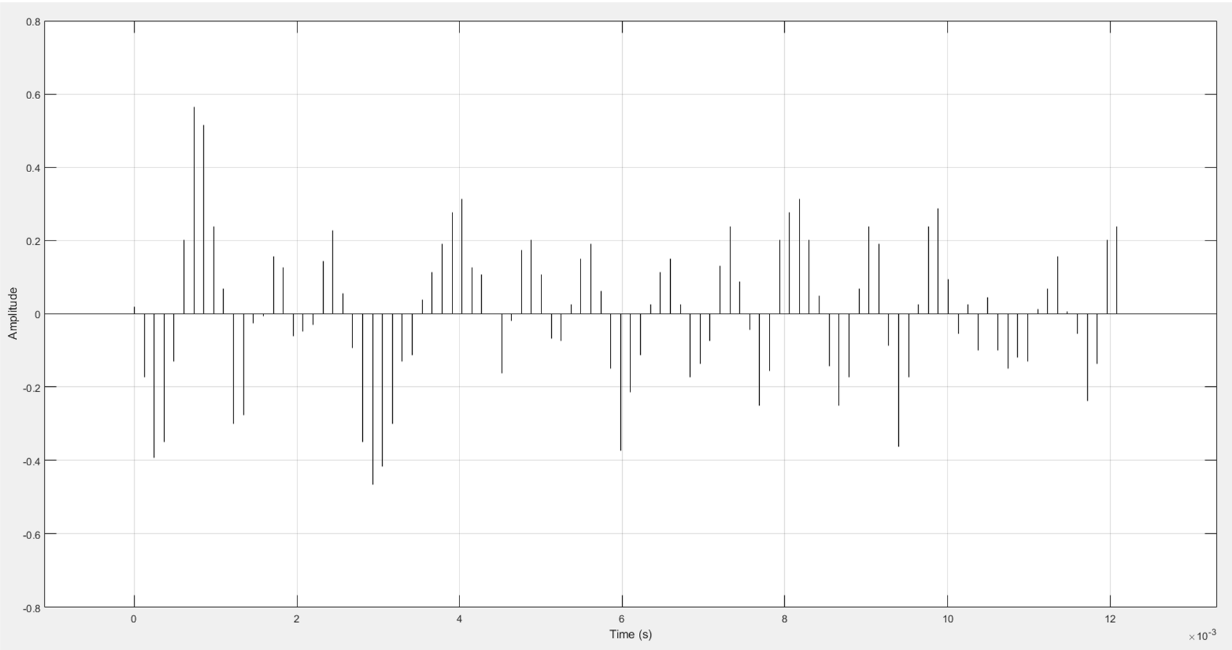
```
ogPlt = axes;
stem (time(1:100), vow(1:100), 'k', 'Marker', 'None');
ogPlt.PlotBoxAspectRatio = [2 1 1];
axis ([-time(10) time(110) -0.8 0.8]);
grid on;

xlabel ('Time(s)');
ylabel ('Amplitude');

ft = fft(vow);
N = length (vow);
freq_shift = (-N/2:N/2-1)*(Fs/N);

ftPlt = subplot (2, 1, 1);
stem (freq_shift, abs(fftshift(ft)), 'k', 'Marker', 'None');
grid on;

ftPlt.XLim = [0 freq_shift(end)];
ftPlt.Position = [0.1200 0.4512 0.67 0.4];
ftPlt.XTick = [];
ftPlt.YTick = 50:50:300;
ylabel (ftPlt, 'FT Amplitude');

spectPlt = subplot (2, 1, 2);
colormap ('jet');
spectrogram (vow, 200, [], [], 'xaxis', Fs);

dt = 1/Fs;
formants = zeros(5, 3);

for i = [0.1 0.2 0.3 0.4 0.5]
        Ibegin = round (i/dt);
        Iend = round ((i+0.1)/dt);

        testRange = vow(Ibegin:Iend);
        testRange = testRange.*hamming(length(testRange))';
        testRange = filter (1, [1 0.63], testRange);

        poly = lpc(testRange, 8);
        rts = roots (poly);
        rts = rts (imag(rts) >= 0);
        argz = atan2 (imag(rts), real(rts));

        [frqs, indices] = sort(argz.*(Fs/(2*pi)));
        bw = -0.5*(Fs/(2*pi))*log(abs(rts(indices)));

        nn = 1;
        for kk = 1:length (frqs)
```

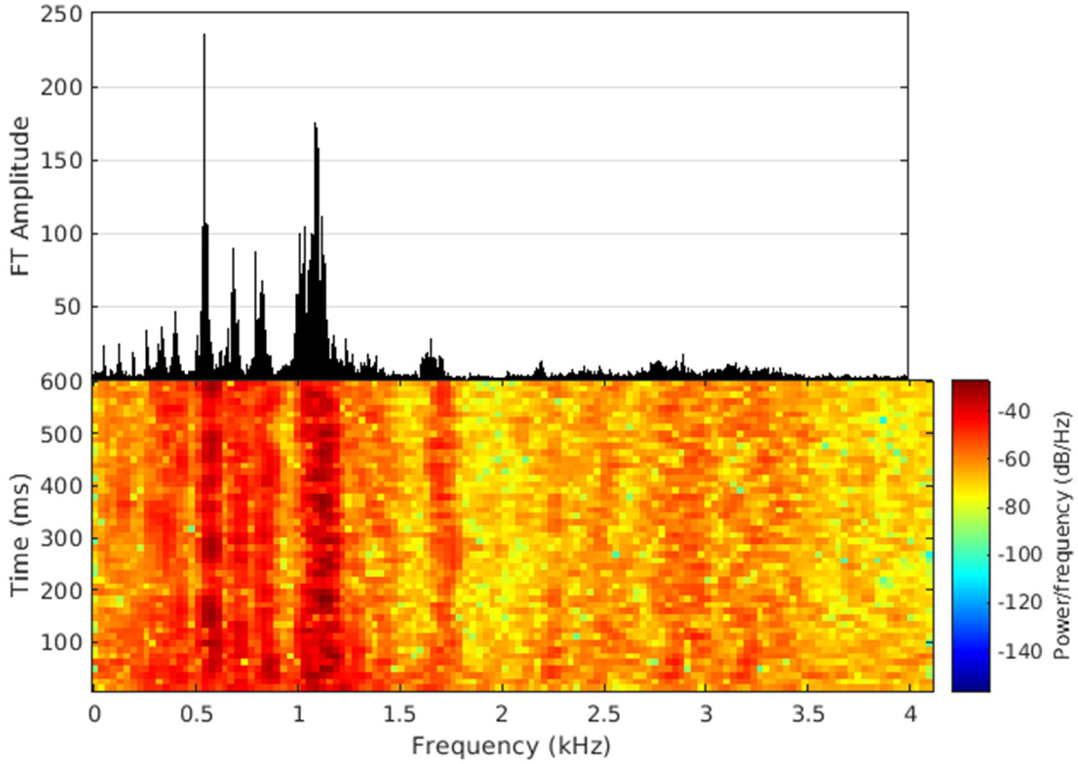
```
                        if (frqs(kk) > 90 && bw(kk) < 400)
                                formants (i*10, nn) = frqs(kk);
                                nn = nn+1;
                        end;
                end;
        end;
        writematrix (formants, 'formants.txt');
```

INPUT DATA DESCRIPTION:

Samples considered from "handle.mat" equals 21e+3 – 26e+3
The spectrogram window is taken to be 200 samples.
Interval for formant calculations is 0.1 seconds.

RESULT:





CONCLUSION/DISCUSSION:

From the graphs plotted above, the broad peaks form at the lower frequencies, while the higher frequencies are rather sparse. It is a wild guess that the vowel ʌ is the most dominant sound. Upon further considerations and looking at the generated fricant results for 0.1 intervals and comparing it with the average formant locations (Hz) for vowels in American English, we can make out that the vowels /i/ /æ/ /a/ /U/ /u/ /ʌ/ are also present.

There is a range of values centered around F_1 and F_2 frequencies, although, the vowel ʌ is clearly the most dominant. However, in F_3 frequency, there is a significant difference in values.