VOICE ACOUSTIC AND DETECTION OF VOICE DISORDER

by

H Tejswaroopreddy (21bec1606)

R Aashish (21bec1766)

A project report submitted to

SUKRITI

Associate Professor

SCHOOL OF ELECTRONICS ENGINEERING

B. Tech. ELECTRONICS AND COMMUNICATION ENGINEERING

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Vandalur – Kelambakkam Road

Chennai – 600127

ABSTRACT

Digital signal analysis are extremely useful in the diagonisis and study of the voice disorders .The microphone is useful to caputure the voice an acoustic signal can be produced and stored in the computer . This distinct signal pattern is an excellent indication of the quality of the voice and can be used to compare the deviant voices with the healthy voices .

This study focuses of on the detection of the voice disorder function of laryngeal functions –specifically on the acoustic analysis of voice fold vibration and the pitches of the voices analysed using matlab computer program, acoustic signal of voice can be analysed by fast fourier transform. The comparison is made between the pattern of the normal voice and the disorder voice can be detected.

1. INTRODUCTION

1.1 OBJECTIVES AND GOALS

Voice disease are increasingly dramatically due to unhealthy social habits and voice abuse the diagnosis of voice disorders, one limitation exists with the glottal area waveform design. If a problem exists with the patient's vocal tract, analysis using the glottal opening technique could not determine the cause of the disorder.

- In this case, only the acoustic technique could detect the problem.

1.2 APPLICATIONS

- The speech acoustics of the patient may use by the healthcare professional for the easy detection of diseases that the patient was affected
- The information is recorded more accurately, transmitted properly, and also delivered and saved with much more precision. The information is also transmitted in real-time across departments, due to which both the departments remain updated with factual information. With the advancement in technology, speech recognition systems are getting more authentic and are expected to save time abd money as well.
- The Speech recognition system can capture speech much faster than typing in real-time and provides more flexibility to the physician.

1.3FEATURES:

A quiet environment is essential for the proper operation of speech and voice recognition technology. Too much background noise might degrade speech and voice recognition results. Hence, this is one of the primary obstacles to efficiently deploying speech and voice recognition technology

2.IMPLEMENTATION

2.1 DESIGN AND IMPLEMENTATION

ANALYTICAL APPROACHES ACOUSTIC ANALYSIS Noise is always contained in a person's voice, present in small fluctuations in frequency, amplitude, and wave shape known as perturbations (30). Perturbation in a voice signal is defined as a minor disturbance, or temporary change from an expected behavior. Perturbations do not alter the qualitative appearance of a pattern, but are only small irregularities. The graphs below show acoustic signals from two subjects with normal voices. The second patient shows significant levels of noise present in the intensity peaks, although both voices are of similar perceived quality.

SPECTROGRAM Fourier analysis provides a general tool for examining the global frequency distribution. However, a meaningful Fourier spectral analysis is restricted to linear systems and periodic or stationary data. On the other hand, the spectrogram, or short Fourier Transform, aims to modify the global representation of the Fourier analysis and provide local time and frequency distribution (33). In the clinical analysis of voice disorders, the periodicity (regularity), or deviation from it, aperiodicity (irregularity), of the acoustic signal or glottal waveform between vibratory cycles is one of the key indices to describe most pathological voices. Therefore, the regularity of successive vocal fold vibrations over sufficient number of cycles needs to be analyzed. For this purpose, it is far more effective to obtain the instantaneous time frequency distribution rather than a global (or average) distribution as usually derived from the Fourier spectral analysis. The spectrogram graphs the instantaneous frequency distribution with time.. The spectrogram uses darker colors to illustrate higher intensity levels, and is able to show how the amplitude of the voice spectrum changes over time. The spectrogram equation is simply the Fourier Transform equation multiplied by a window function, w(t). This window function adds a time aspect to the previous Fourier Transform equation. The given signal is analyzed at each time interval using this small window. The window is moved after every calculation, until the entire signal is analyzed. The spectrogram can then display the relationship of frequency, intensity, and time.

The spectrogram is denoted by:

$$S(t,f) = \int_{-\infty}^{\infty} [f(t) w^*(t-t')] e^{-i2\pi ft} dt \,,$$

The spectrogram is more useful in analyzing the disorders for this project since it obtains a threedimensional plot of frequency, intensity, and time. The average intensity bands of the spectrogram correspond to the peaks in each FFT graph, as shown by the figure below. The figure shows an example of a typical spectrogram of a sustained vocalization from the same normal voice as the above FFT. The higher-order harmonics represent the fundamental frequency and the vocal tract effect, and correspond to the harmonics of the previous FFT plot.

2.2 BLOCK DIAGRAM:

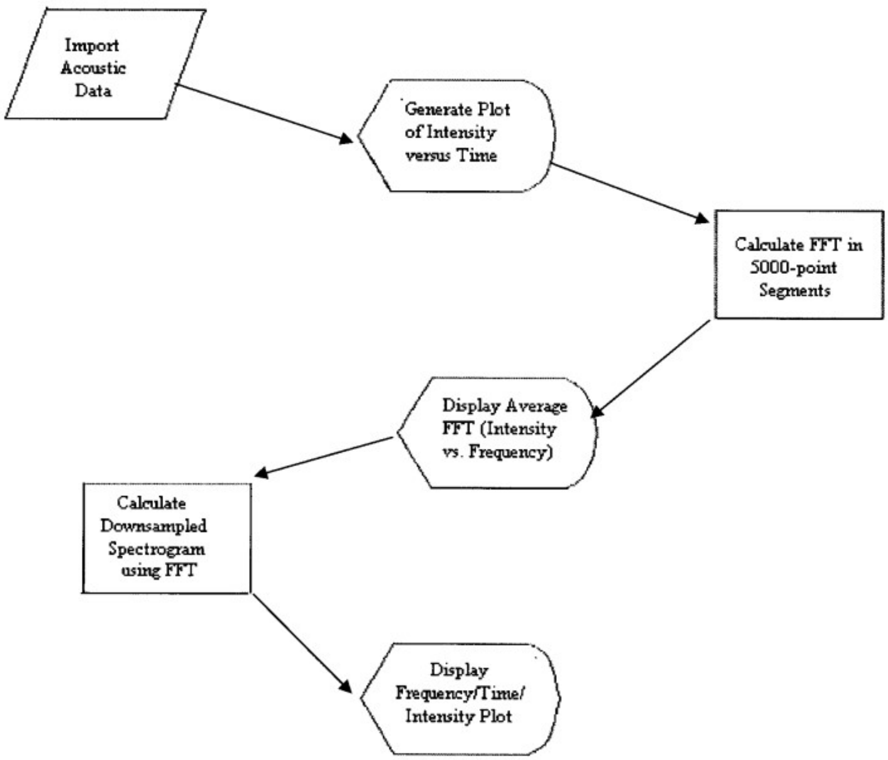


Figure 20. Flow chart of MatLab Computer Program

3.1 SOFTWARE –CODING AND ANALYSIS

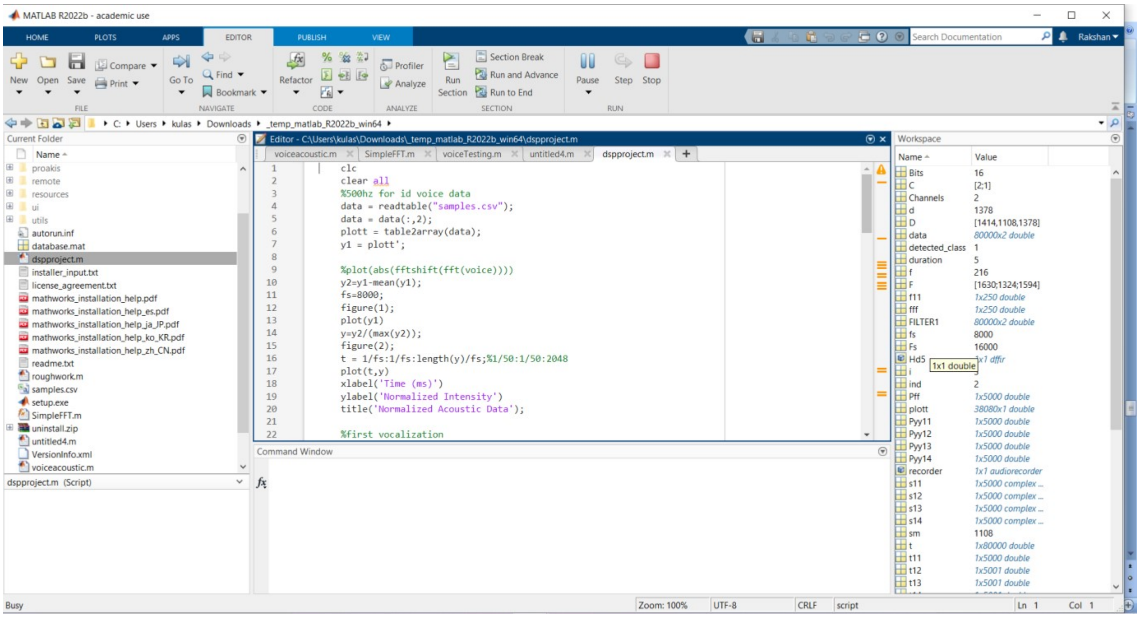
MatLab (by The MafhWorlcs, Inc.) offers technical computing for numerical computations. It is very useful for signal processing and computation, as it performs numerous calculations very quickly. The version used for this project was MatLab 7.0. For each patient's vocalization, the corresponding data file showing vocal intensity versus time was imported to the MatLab program.

FOURIER TRANSFORM As mentioned in the above section, the Fourier Transform displays the frequency components of a given signal.

```
clc
clear all
%500hz for id voice data
data = readtable("samples.csv");
data = data(:,2);
plott = table2array(data);
y1 = plott';

%plot(abs(fftshift(fft(voice))))
y2=y1-mean(y1);
fs=8000;
figure(1);
plot(y1)
y=y2/(max(y2));
figure(2);
t = 1/fs:1/fs:length(y)/fs;%1/50:1/50:2048
plot(t,y)
xlabel('Time (ms)')
ylabel('Normalized Intensity')
title('Normalized Acoustic Data');

%first vocalization
x=(1:26000);
t=x/fs
Y=y(1:26000);
plot(t,Y) %plots first 26000 frames of vocalization
xlabel('Time (ms)')
ylabel('lntensity')
title('Acoustic Data for First Vocalization');
```

```
    %1,2,3,4 segment of first vocalization%
    x11 =(1:5000)
    t11 =x11/fs;
    Y11 =y(1:5000)
    x12 =(5000:10000)
    t12 =x12/fs;
    Y12 =y(5000:10000)
    x13 =(10000:15000)
    t13 =x13/fs;
    Y13 =y(10000:15000)
    x14 =(20000:25000)
    t14 =x14/fs;
    Y14 =y(20000:25000)
    plot(t11 ,Y11)
    xlabel('Time (ms)')
    ylabel('lntensity')
    title('Acoustic Data for First Segment');

    s11 =fft(Y11,5000);
    Pyy11 =s11.*conj(s11 )/5000;
    f11 =(50000/5000)*(0:249); %calculates fft for first segment
    s12 =fft(Y12,5000);
    Pyy12 =s12.*conj(s12 )/5000;
    s13 =fft(Y13,5000);
    Pyy13 =s13.*conj(s13 )/5000;
    s14 =fft(Y14,5000);
    Pyy14 =s14.*conj(s14 )/5000;
    plot(f11 ,Pyy11 (1 :250))
    xlabel('Frequency (Hz)')
    ylabel('lntensity')
    title('FFT for the First Segment');


    %average fft for first vocalization%
    Pff=(Pyy11 +Pyy12+Pyy13+Pyy14)/4;
    fff=(50000/5000)*(0:249);
    plot(fff,Pff(1 :250));
    xlabel('Frequency (Hz)')
    ylabel('lntensity')
    title('Average FFT for First Vocalization');

%second vocalization
x2=(26000:38000)
t2=x2/50000
Y2=y(26000:38000);
plot(t2,Y2) %plots second 12000 frames of vocalization
xlabel('Time(ms)')
ylabel('Intensity')
title('Acoustic Data for Second Vocalization');
%
%1,2,3,4 segment of second vocalization%
x21 =(26000:31000)
t21 =x21/50
Y21 =y(26000:31000);
plot(t21 ,Y21) %plots fourth 5000 frames of vocalization
xlabel(Time (ms)')
ylabel('Intensity')
title('Acoustic Data for First Segment');
s21 =fft(Y21 ,5000);
```

```
Pyy21 =s21.*conj(s21 )/5000;
f21 =(50000/5000)*(0:249); %calculates fft for first segment and plots first 250 points
plot(f21 ,Pyy21 (1 :250))
xlabel('Frequency (Hz)')
ylabel('Intensity')
title('FFT for the First Segment');

%average of all vocalizations%
PffF=(Pff+Pff2)/2;
fffF=(50000/5000)*(0:249);
plot(fffF,PffF(1 :250));
xlabel('Frequency (Hz)')
ylabel('Intensity')
%title('Average FFT');
End

%Spectrogram
Yd=DOWNSAMPLE(y, 10)
specgram(Yd,D,5000)
ylabel('Frequency (Hz)')
xlabel(Time (s)')
%title('Spectrogram');
```



4. INFERENCE, RESULT, AND CONCLUSION

COMPARISON OF A NORMAL VOICE WITH AND WITHOUT NOISE

As mentioned previously, the amount of noise present in a person's voice is not indicative of the vocal quality. An example of this is shown below. Both samples are from the same normal patient, but the sample on the right contains vocal noise. This patient performed short vocalizations of the vowel leI. The acoustic signals have been zoomed to show only 20 ms of the vocalization so that the signal pattern can be easily seen. The noise is seen in the small fluctuations in intensity. As shown by the FFT and the spectrogram, both are healthy voices

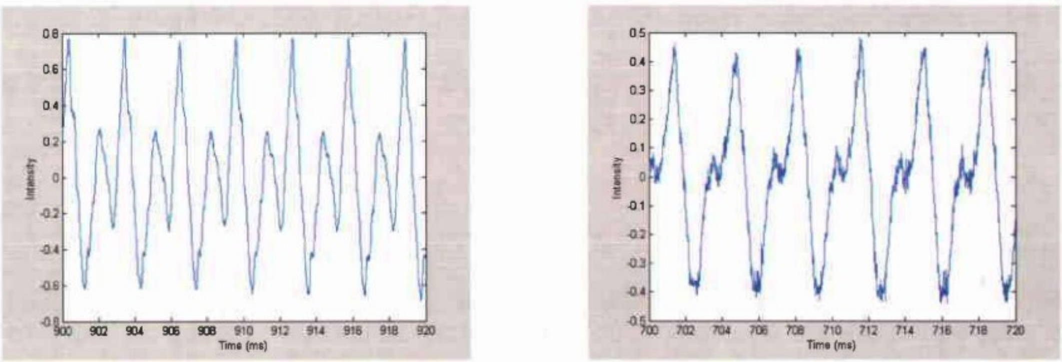


Figure 21. Acoustic Signal of vocalizations from the same patient with a normal voice without noise (left) and with noise (right)

The FFT plots for these signals are shown in the figures on the next page.

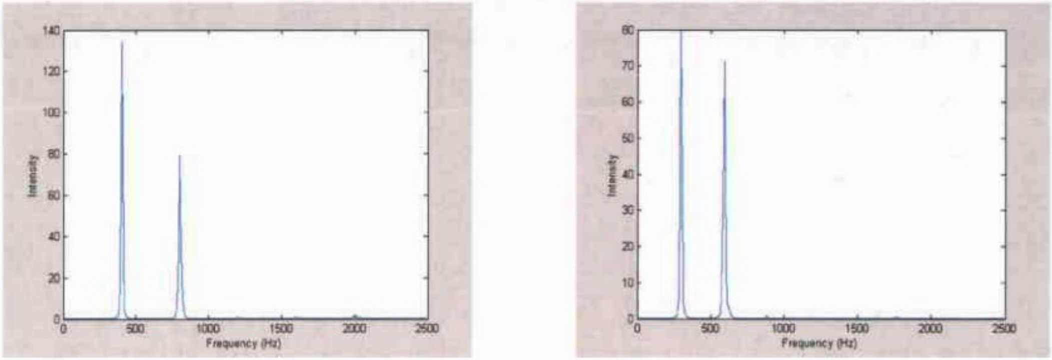


Figure 22. FFT Comparison of normal vocal samples without noise (left) and with noise (right)

REFERENCES

1. Author Name , (Year),"Title ", Computers and Electrical Engineering, Vol 44, pp 45, pp 153-171.
2. (minimum 5-recent journals)

3.Boone, DR., McFarlane, S.C. (2000). The Voice and Voice Therapy. (6th ed.). Boston, MA: Allyn and Bacon.

4.Morrison, M., Rammage, L. (1994). The Management of Voice Disorders. London: Chapman & Hall Medical.