A

Project Based Lab Report

on

Speech Signal Analysis

A mini project work on Signal Processing (13-ES-205L) submitted to
KL University under the partial fulfilment of
II /IV B.Tech

By

B.SIDHARTH SARAT RAJ (14006128)

CH.CHANDU(14006131)

R.SRIKANTH(14006130)

MD.ISMAIL PATHAN(14006131)

Under the guidance of

MIS.PRIYANKA MADEM (Asst. Prof.)

Dept. of Electronics and Communication Engineering
K.L. UNIVERSITY
Green fields,Vaddeswaram-522502,
Guntur District

2014-15

Department of Electronics & Communication Engineering

K.L.UNIVERSITY

CERTIFICATE

This is to certify that the project entitled **"Speech Signal Analysis"** is the bonafide work carried out by **B.SIDHARTH SARAT RAJ(14006128), CH.CHANDU(14006129), R.SRIKANTH(14006130), MD.ISMAIL PHATAN(14006131)** students of II year B.Tech, EEE Department, College of Engineering, K.L.University, in the "Signal Processing Laboratory" for the academic year 2014-2015.

Signature of the Project guide   Signature of Course Coordinator

Head of the department

ACKNOWLEDGMENT

Our sincere thanks to **MS.PRIYANKA MADEM** in the Lab for his outstanding support throughout the project for the successful completion of the work.

We express our gratitude to **Dr.K.Thirupathi Rao,** Head of the Department for Computer Science and Engineering for providing us with adequate facilities, ways and means by which we are able to complete this project.

We would like to place on record the deep sense of gratitude to the honourable Vice Chancellor, K L University for providing the necessary facilities to carry the concluded project.

Last but not the least, we thank all Teaching and Non-Teaching Staff of our department and especially my classmates and my friends for their support in the completion of our project.

Place: KL University

Date: 11-11-2014

B.SIHARTH SARAT RAJ(14006128)

CH.CHANDU(14006129)

R.SRIKANTH(14006130)

MD.ISMAAIL PATHAN(14006131)

CONTENTS

ABSTRACT

Objectives:

(a) Load, display and manipulation of speech signals.

(b) Estimate the fundamental frequency of a section of speech signal from its waveform using autocorrelation.

(c) Estimate the fundamental frequency of a section of speech signal from its spectrum using cepstrum.

(d) Compute and plot the spectrum of speech signals.

This process is shown in the following block diagram.

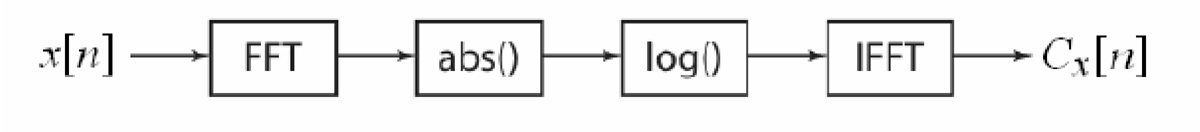


Fig 1: Computation of Cepstrum of a Signal

Task1: Fundamental frequency estimation-time domain: Auto-correlation

The perception of pitch is more strongly related to periodicity in the waveform itself. A means to estimate fundamental frequency from the waveform directly is to use *autocorrelation*. The autocorrelation function for a section of signal shows how well the waveform shape correlates with itself at a range of different delays. We expect a periodic signal to correlate well with itself at very short delays and at delays corresponding to multiples of pitch periods. We can estimate the fundamental frequency by looking for a peak in the delay interval corresponding to the normal pitch range in speech.

Task2: Fundamental frequency estimation- frequency domain: Cepstrum

A reliable way of obtaining an estimate of the dominant fundamental frequency for long, clean, stationary speech signals is to use the *cepstrum*. The cepstrum is a Fourier analysis of the logarithmic amplitude spectrum of the signal as shown in Fig.1. If the log amplitude spectrum contains many regularly spaced harmonics, then the Fourier analysis of the spectrum will show a peak corresponding to the spacing between the harmonics: i.e. the fundamental frequency. Effectively we are treating the signal spectrum as another signal, then looking for periodicity in the spectrum itself.

The cepstrum is so-called because it turns the spectrum inside-out. The x-axis of the cepstrum has units of frequency, and peaks in the cepstrum (which relate to periodicities in the spectrum) are called harmonics. To obtain an estimate of the fundamental frequency from the cepstrum we look for a peak in the frequency region corresponding to typical speech fundamental frequencies.

Task3: Repeat the above tasks-1 and 2 for noisy speech signals.

Task4: Repeat the above tasks-1 and 2 for noisy musical signals.

Task5: Repeat the above tasks-1 and 2 for noisy musical speech signals.

CHAPTER 1

INTRODUCTION

Speech processing is the study of speech signals and the processing methods of these signals. The signals are usually processed in a digital representation, so speech processing can be regarded as a special case of digital signal processing, applied to speech signal. Aspects of speech processing includes the acquisition, manipulation, storage, transfer and output of speech signals.

Speech signal processing refers to the manipulation, acquisition, storage, transfer and output of vocal output by a computing machine.

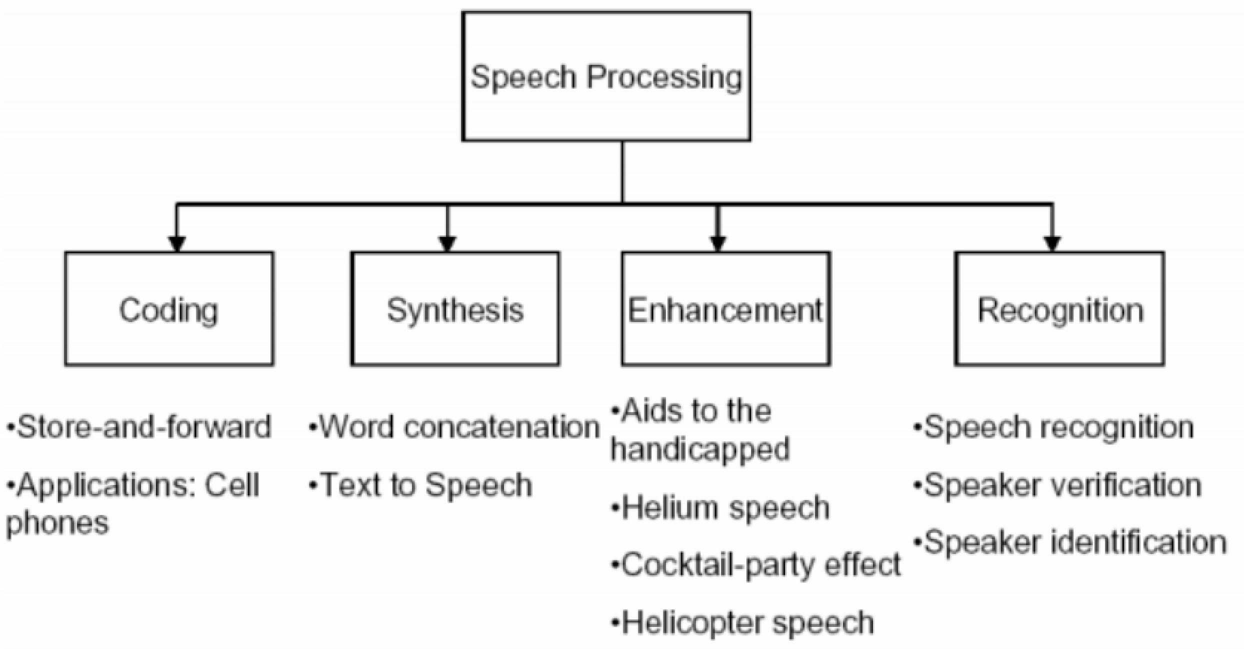


Fig 2: Converting Speech to Digital Signals

The general problem of fundamental frequency estimation is to take a portion of signal and to find the dominant frequency of repetition. Difficulties arise from (i) that not all signals are periodic, (ii) those that are periodic may be changing in fundamental frequency over the time of interest, (iii) signals may be contaminated with noise, even with periodic signals of other fundamental frequencies, (iv) signals that are periodic with interval T are also periodic with interval 2T, 3T etc, so we need to find the smallest periodic interval or the highest fundamental frequency; and (v) even signals of constant fundamental frequency may be changing in other ways over the interval of interest.

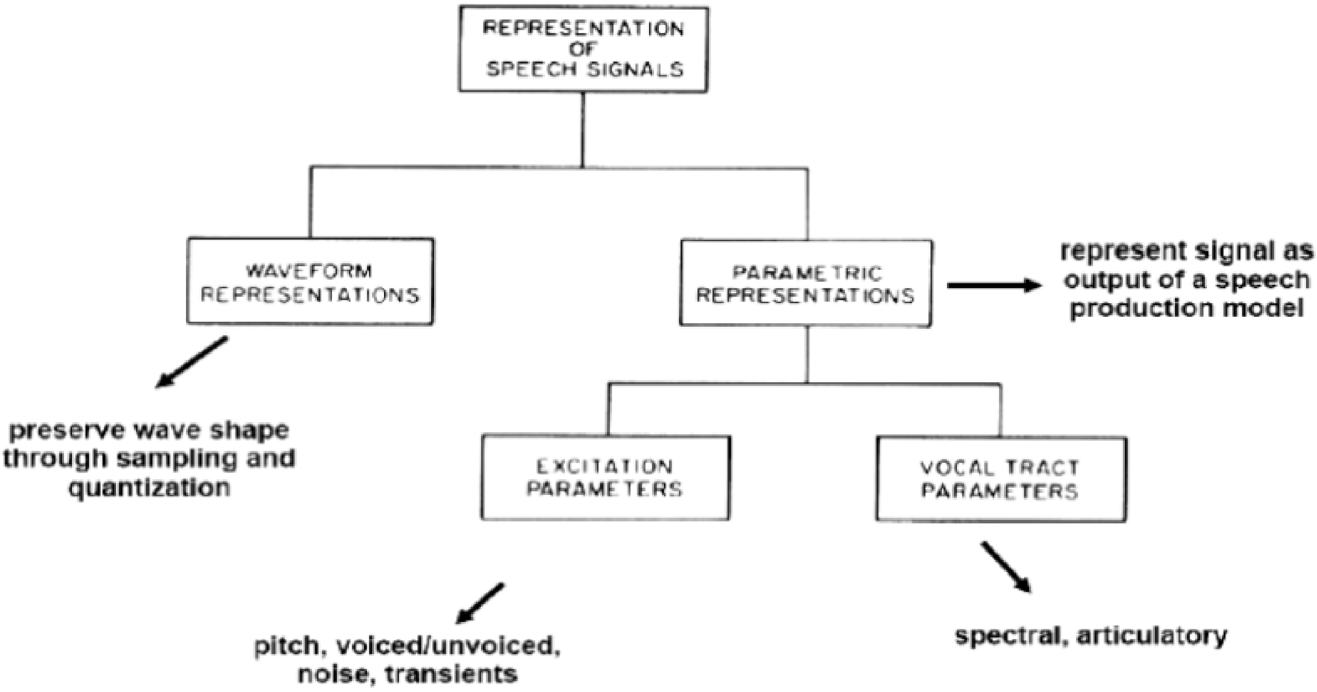


Fig 3: Representation of Speech Signals

A reliable way of obtaining an estimate of the dominant fundamental frequency for long, clean, stationary speech signals is to use the cepstrum. The cepstrum is a Fourier analysis of the logarithmic amplitude spectrum of the signal. If the log amplitude spectrum contains many regularly spaced harmonics, then the Fourier analysis of the spectrum will show a peak corresponding to the spacing between the harmonics: i.e. the fundamental frequency. Effectively we are treating the signal spectrum as another signal, then looking for periodicity in the spectrum itself.

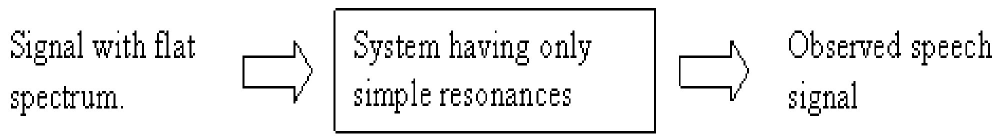
The cepstrum is so-called because it turns the spectrum inside-out. The x-axis of the cepstrum has units of quefrency, and peaks in the cepstrum (which relate to periodicities in the spectrum) are called rahmonics.

To obtain an estimate of the fundamental frequency from the cepstrum we look for a peak in the quefrency region corresponding to typical speech fundamental frequencies.

The cepstrum looks for periodicity in the log spectrum of the signal, whereas our perception of pitch is more strongly related to periodicity in the waveform itself. A means to estimate fundamental frequency from the waveform directly is to use autocorrelation. The autocorrelation function for a section of signal shows how well the waveform shape correlates with itself at a range of different delays. We expect a periodic signal to correlate well with itself at very short delays and at delays corresponding to multiples of pitch periods.

Estimation of formant frequencies is generally more difficult than estimation of fundamental frequency. The problem is that formant frequencies are properties of the vocal tract system and need to be inferred from the speech signal rather than just measured. The spectral shape of the vocal tract excitation strongly influences the observed spectral envelope, such that we cannot guarantee that all vocal tract resonances will cause peaks in the observed spectral envelope, nor that all peaks in the spectral envelope are caused by vocal tract resonances.

The dominant method of formant frequency estimation is based on modelling the speech signal as if it were generated by a particular kind of source and filter:



This type of analysis is called source-filter separation, and in the case of formant frequency estimation, we are interested only in the modelled system and the frequencies of its resonances. To find the best matching system we use a method of analysis called Linear Prediction. Linear prediction models the signal as if it were generated by a signal of minimum energy being passed through a purely-recursive IIR filter.

Speech Processing Applications:

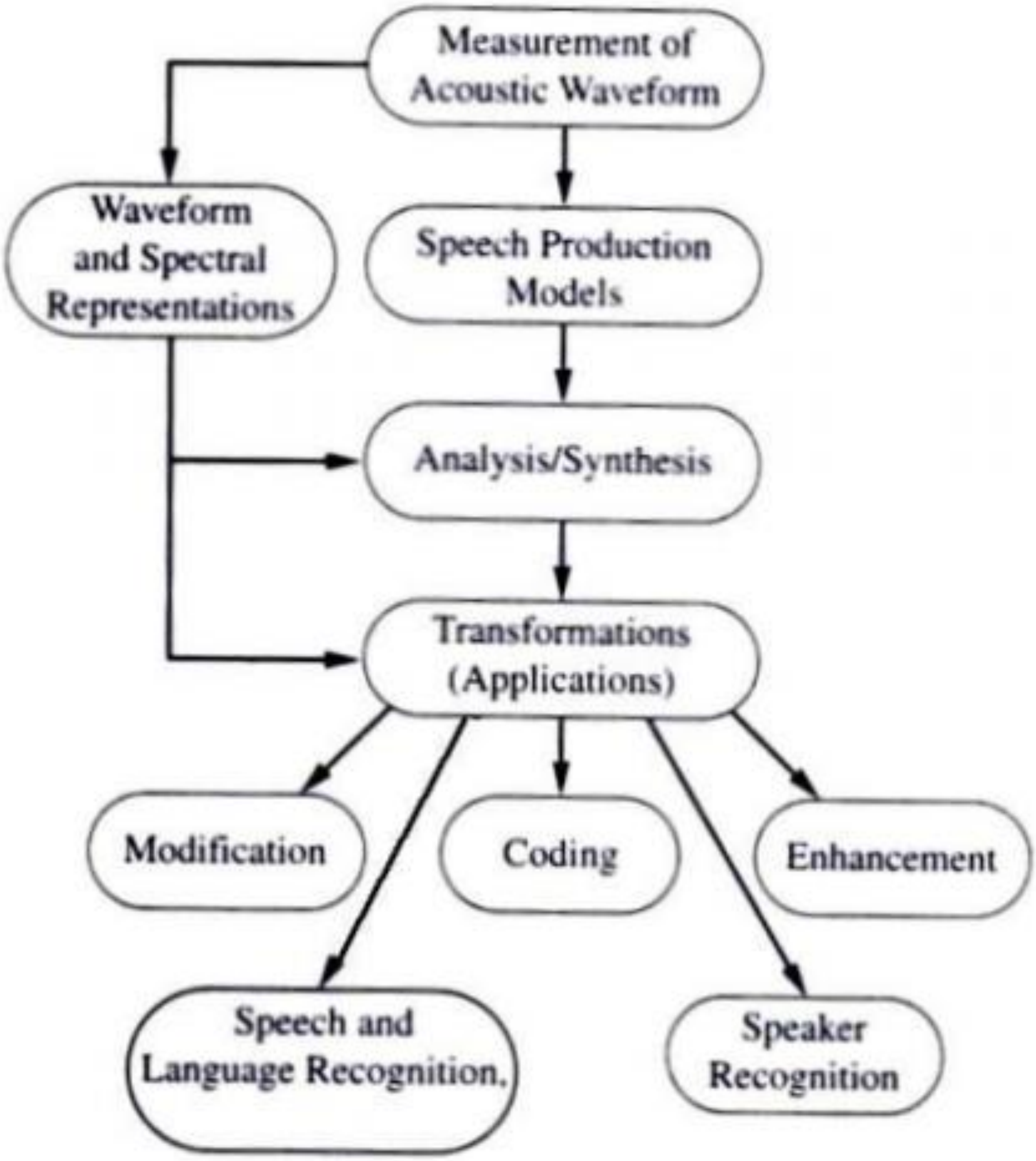


Fig 4: Discrete-time speech signal processing overview.

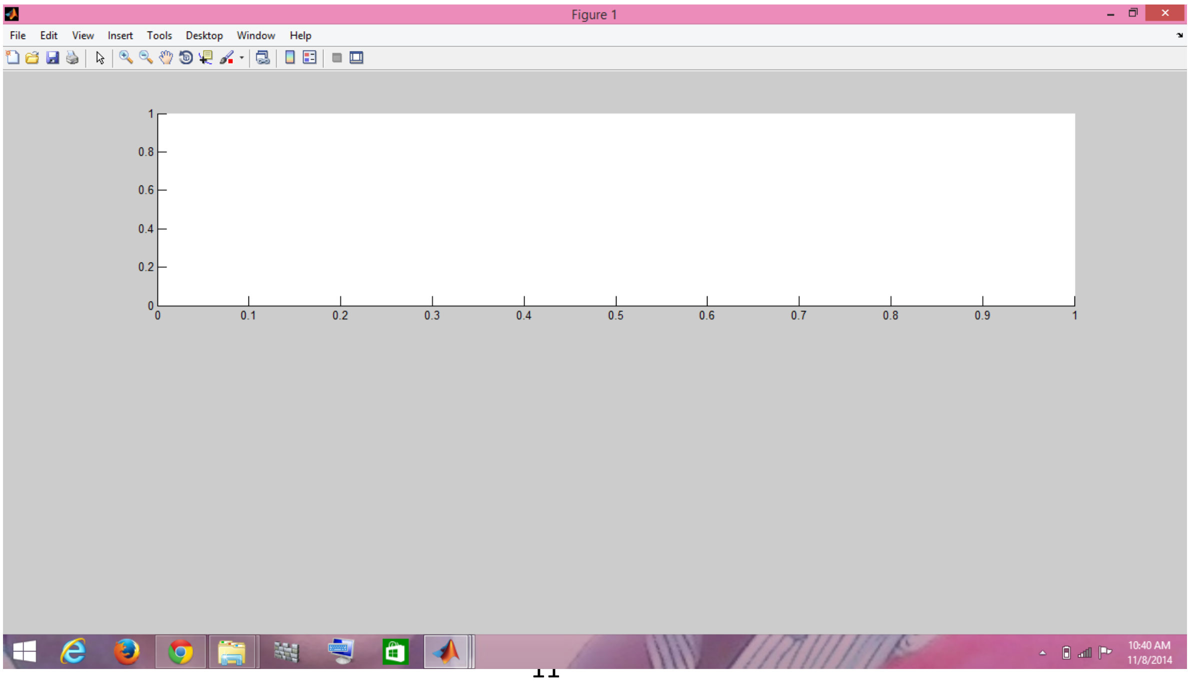
Applications within the text include speech modification, coding, enhancement and speaker recognition.

CHAPTER 2

TASKS & THEIR SIMULATION RESULTS

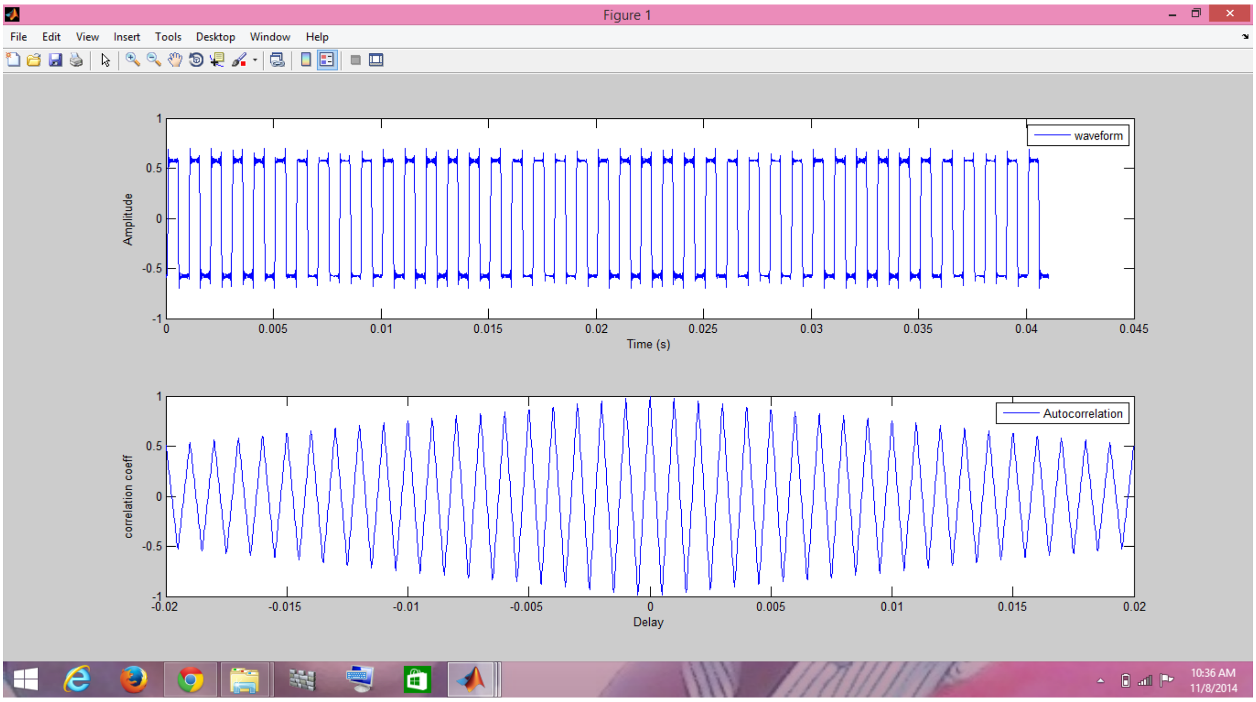
1.A. Task 1: Load, display and manipulation of speech signals.

```
[x,fs]=wavread('C:\Users\MY LAPTOP\Desktop\teja.wav',[24120 25930]);
X=resample(x,5000,fs);
fs=5000;
t=(0:length(X)-1)/fs;
subplot(2,1,1);
plot(t,x);
legend('waveform');
xlabel('Time (s)');
ylabel('Amplitude');
ncoeff=2+fs/1000;
a=lpc(X,ncoeff);
[h,f]=freqz(1,a,512,fs);
subplot(2,1,2);
plot(f,20*log10(abs(h)+eps));
legend('LP Filter');
xlabel('Frequency  (Hz)');
ylabel('Gain (dB)');
```



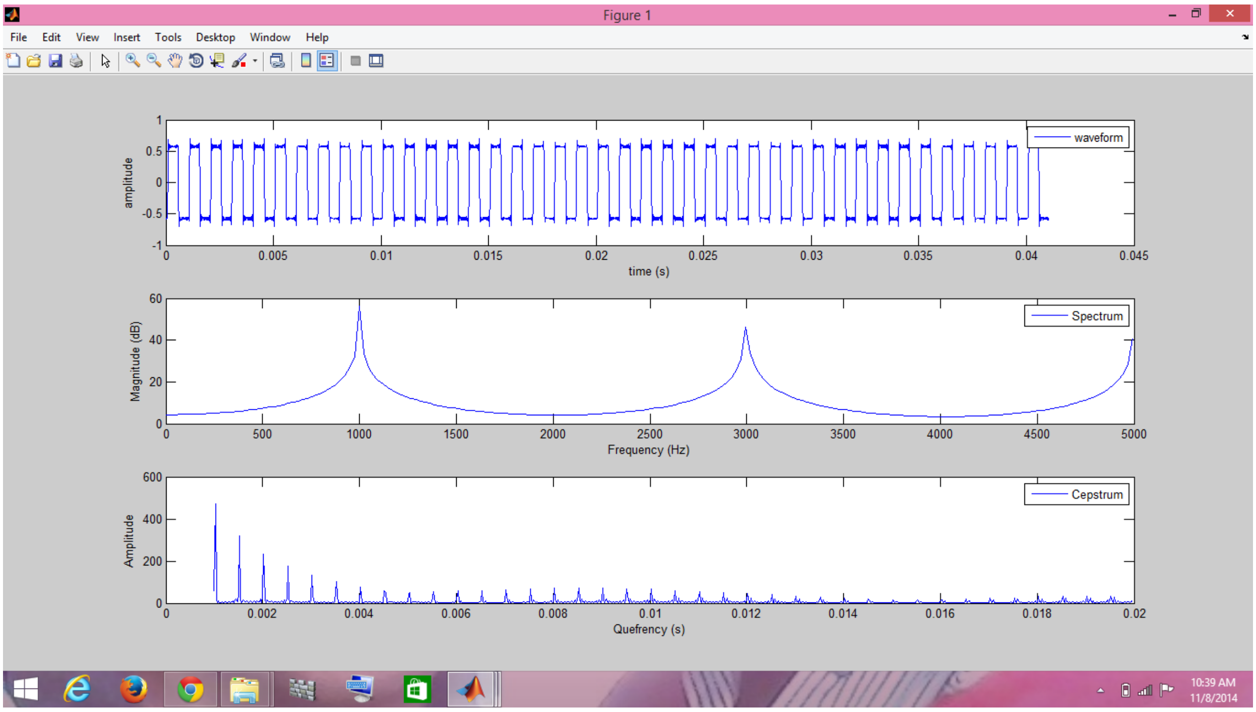
1.B. Task 2: Estimate the fundamental frequency of a section of speech signal from its waveform using autocorrelation.

```
[x,fs]=wavread('C:\Users\MY LAPTOP\Desktop\teja.wav',[24120 25930]);
ms20=fs/50;
t=(0:length(x)-1)/fs;
subplot(2,1,1);
plot(t,x);
legend('waveform');
xlabel('Time (s)');
ylabel('Amplitude');
r=xcorr(x,ms20,'coeff');
d=(-ms20:ms20)/fs;
subplot(2,1,2);
plot(d,r);
legend('Autocorrelation');
xlabel('Delay');
ylabel('correlation coeff');
```



1.C. Task 3: Estimate the fundamental frequency of a section of speech signal from its spectrum using cepstrum.

```
[x,fs]=wavread('C:\Users\MY LAPTOP\Desktop\teja.wav',[24120 25930]);
ms1=fs/1000;
ms20=fs/50;
t=(0:length(x)-1)/fs;
subplot(3,1,1);
plot(t,x);
legend('waveform');
xlabel('time (s)');
ylabel('amplitude');
Y=fft(x);
hz5000=5000*length(Y)/fs;
f=(0:hz5000)*fs/length(Y);
subplot(3,1,2);
plot(f,20*log10(abs(Y(1:length(f)))+eps));
legend('Spectrum');
xlabel('Frequency (Hz)');
ylabel('Magnitude (dB)');
C=fft(log(abs(Y)+eps));
q=(ms1:ms20)/fs;
subplot(3,1,3);
plot(q,abs(C(ms1:ms20)));
legend('Cepstrum');
xlabel('Quefrency (s)');
ylabel('Amplitude');
```



1.D. Task 4: Compute and plot the spectrum of speech signals.

CHAPTER 3

CONCLUSIONS & FUTURE SCOPE

From prehistory to the new media of the future, speech has been and will be a primary form of communication between humans.

Nevertheless, there often occur conditions under which we measure and then transform the speech to another form, speech signal, in order to enhance our ability to communicate.

The speech signal is extended, through technological media such as telephony, movies, radio, television, and now Internet. This trend reflects the primacy of speech communication in human psychology.

The topic of speech signal processing can be loosely defined as the manipulation of sampled speech signals by a digital processor to obtain a new signal with some desired properties.

REFERENCES

- http://202.114.89.42/resource/pdf/75.pdf
- http://www.phon.ucl.ac.uk/courses/spsci/matlab
- http://www.phon.ucl.ac.uk/courses/spsci/matlab/lect10.html
- http://www.phon.ucl.ac.uk/courses/spsci/dsp/
- http://mi.eng.cam.ac.uk/~ajr/SA95/SpeechAnalysis.html