ELL205 Project Report

Signals and Systems

Ridhima Gupta 2021EE30719 Khushika Shringi 2021EE10665

Project Title: Cepstral analysis of speech signals

OBJECTIVE

To understand Cepstral Analysis of speech signals

To analyse speech signals by computing the Cepstrum

We went through various materials and references and gathered information regarding Cepstrum Analysis approach of analysing the speech signals.

Our Approach to Cepstral Analysis:

An output signal (Speech signal y(n)) can be thought of convolution of input excitation (excitation source x(n)) and system response (Vocal tract system response h(n)). When both of the components are unknown and we have to extract them, we do Cepstral Analysis of that signal. Speech has voiced and unvoiced sounds. Voiced sounds are produced by exciting the time varying system characteristics with periodic impulse sequence and unvoiced sounds are produced by exciting the time varying system with a random noise sequence.

$$y(n) = h(n) * x(n)$$

$$Y(w)=H(w).X(w)$$

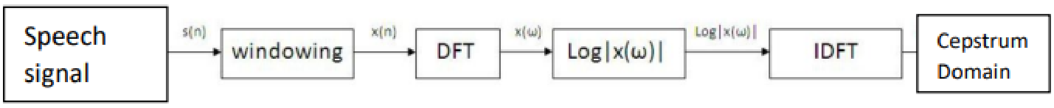
Cepstral analysis transforms the multiplied source and system components in the frequency domain to their linear combination in the cepstral domain known as *quefrency* domain.

$$\mathrm{Log}(|Y(w)|) = \mathrm{Log}(|H(w)|) + \mathrm{Log}(|X(w)|)$$

$$c(n)\ \{\text{Cepstrum domain}) = \mathrm{IDFT}(\mathrm{Log}(|Y(w)|)) = \mathrm{IDFT}(\mathrm{Log}(|H(w)|)) + \mathrm{IDFT}(\mathrm{Log}(|X(w)|))$$

IDFT→Inverse Discrete Fourier Transform.

The vocal tract components are represented by the slowly varying components concentrated near the lower quefrency region (low-time liftering) and excitation components are represented by the fast varying components at the higher quefrency (high-time liftering). We have two types of Cepstrum Real and Complex. Real Cepstrum Processing is done as follows:



Work Plan

1) Firstly, We have completed the theoretical and numerical part

2) Collection of Sample speech signals: We have collected a few speech signals for analysis

which includes voiced sounds like a,e,i.

Signals in the below link:

https://drive.google.com/drive/folders/1TWftkw7lwG75dVdOXHD49B6MJNcnTwLS?usp=sharing

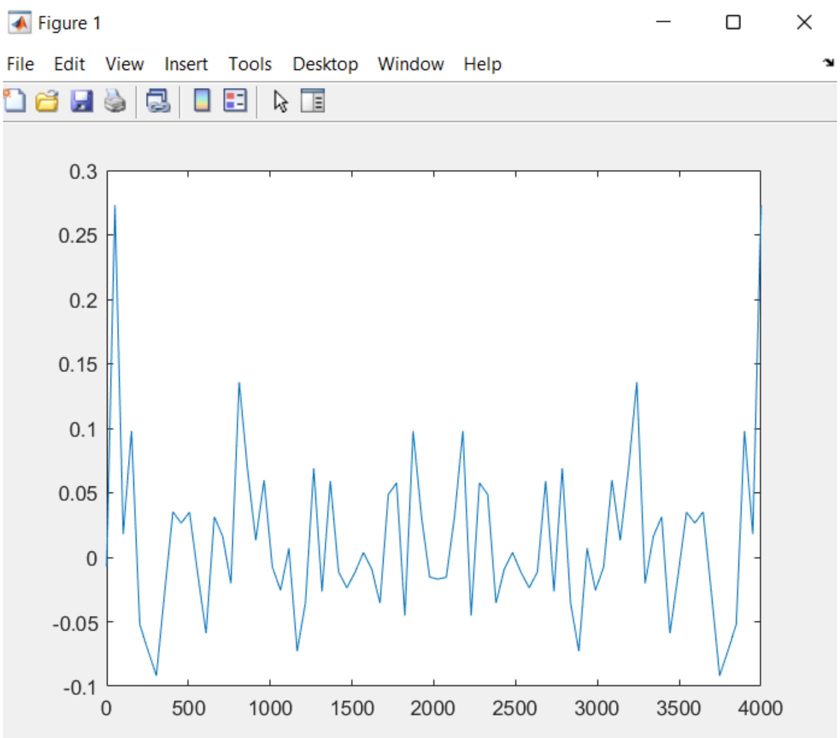
3)We learnt MATLAB simulation for Signal Processing.

4)We took various samples and through MATLAB we computed firstly the Fourier Transform of the sample , then took log , inverse Fourier and thus obtained the real cepstrum of our voiced sound.

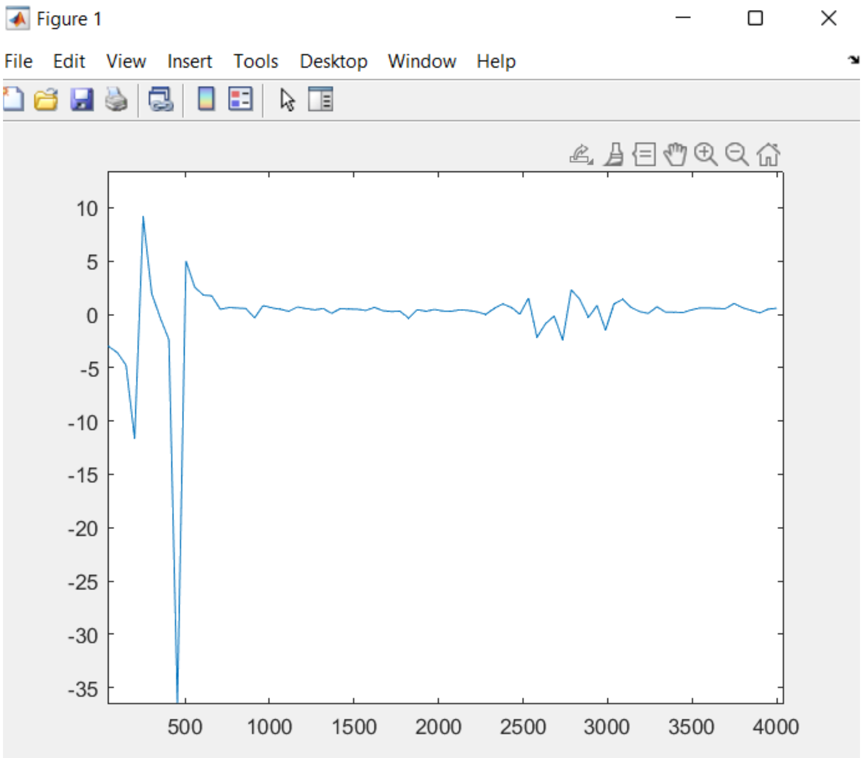
(We have shown here the results of a.wav sample)

OUTPUT :

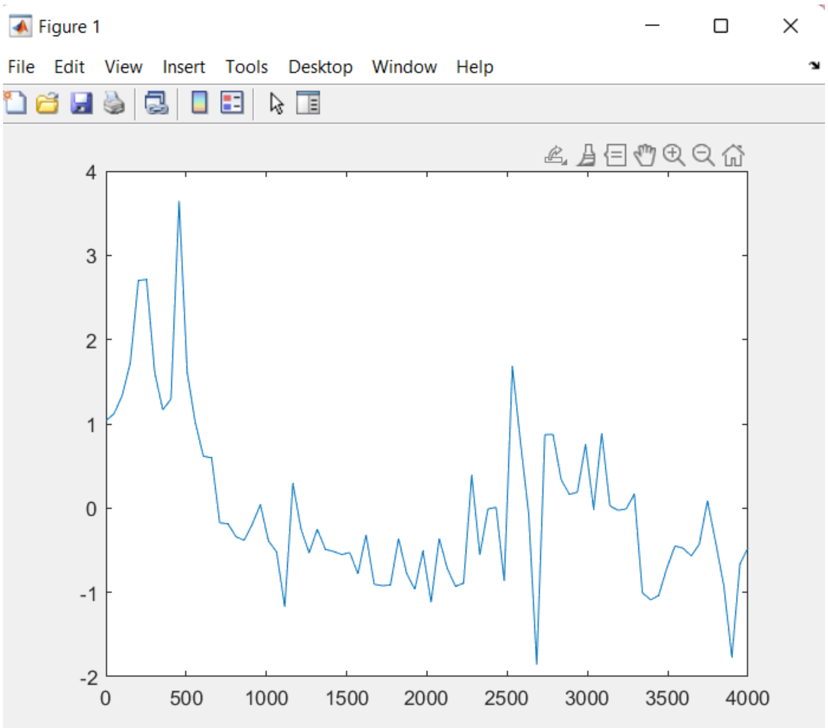
Speech Signal X(w)



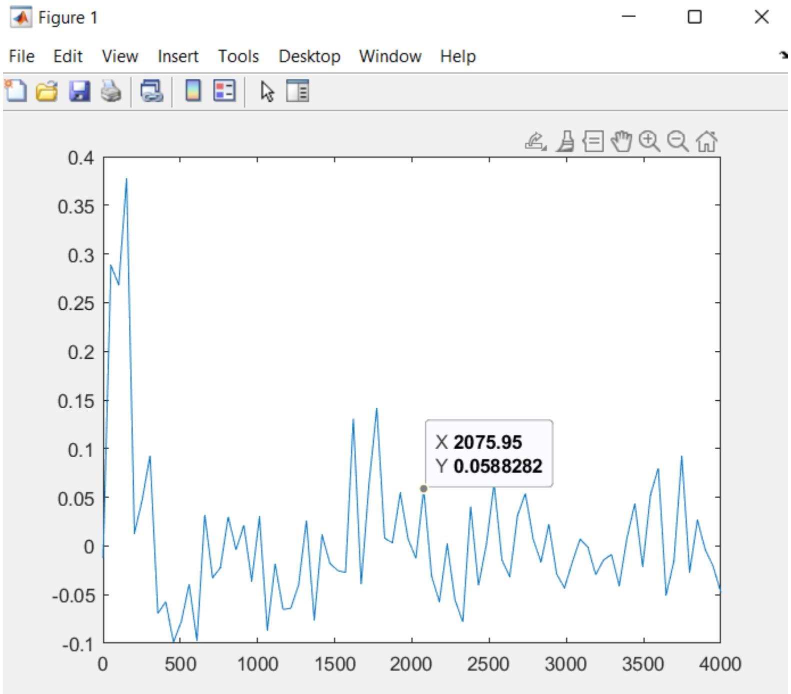
FOURIER TRANSFORM X(w)



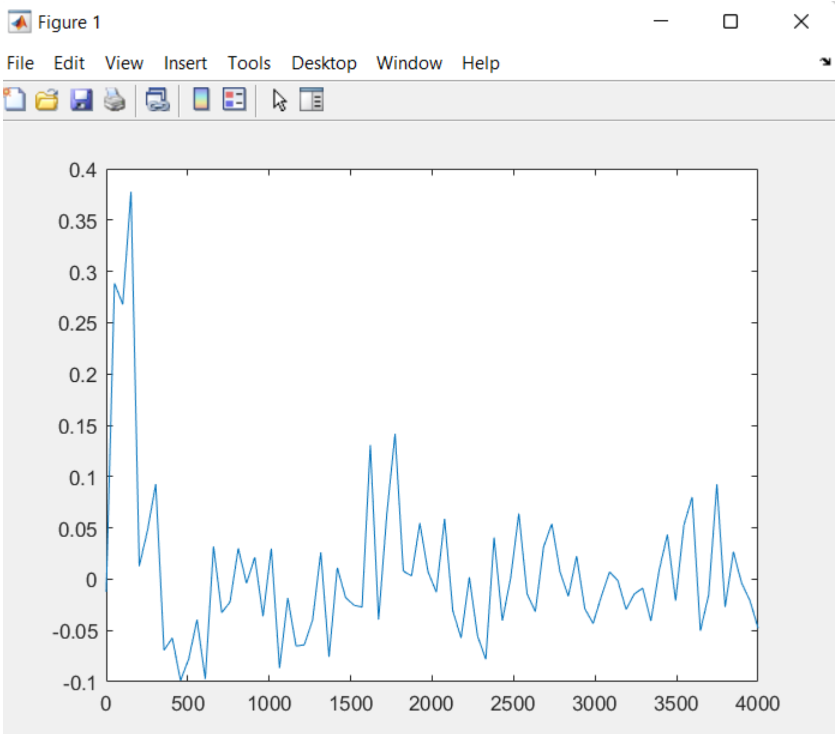
Taking Log-> log(abs(fft(X(w)))



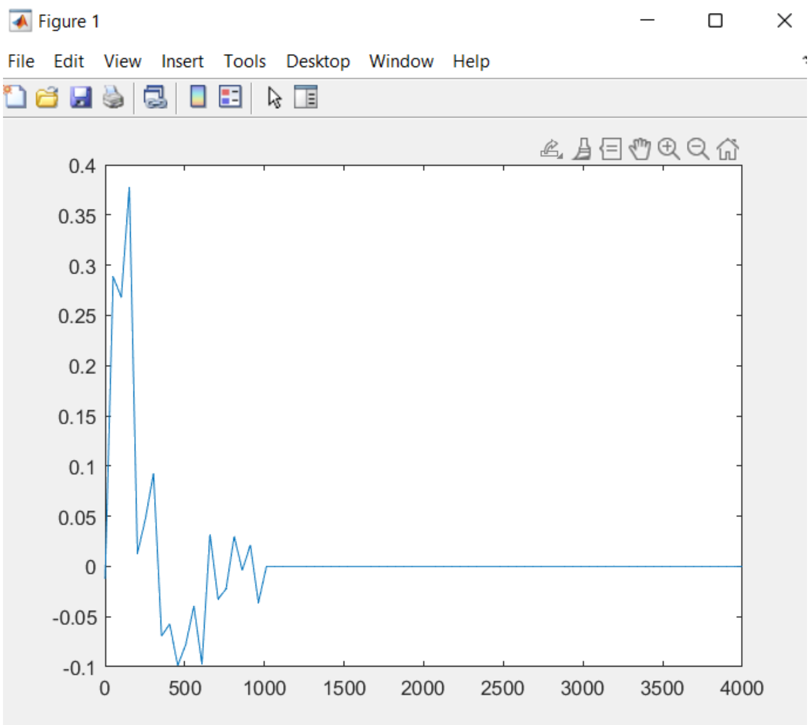
Inverse Fourier Transform



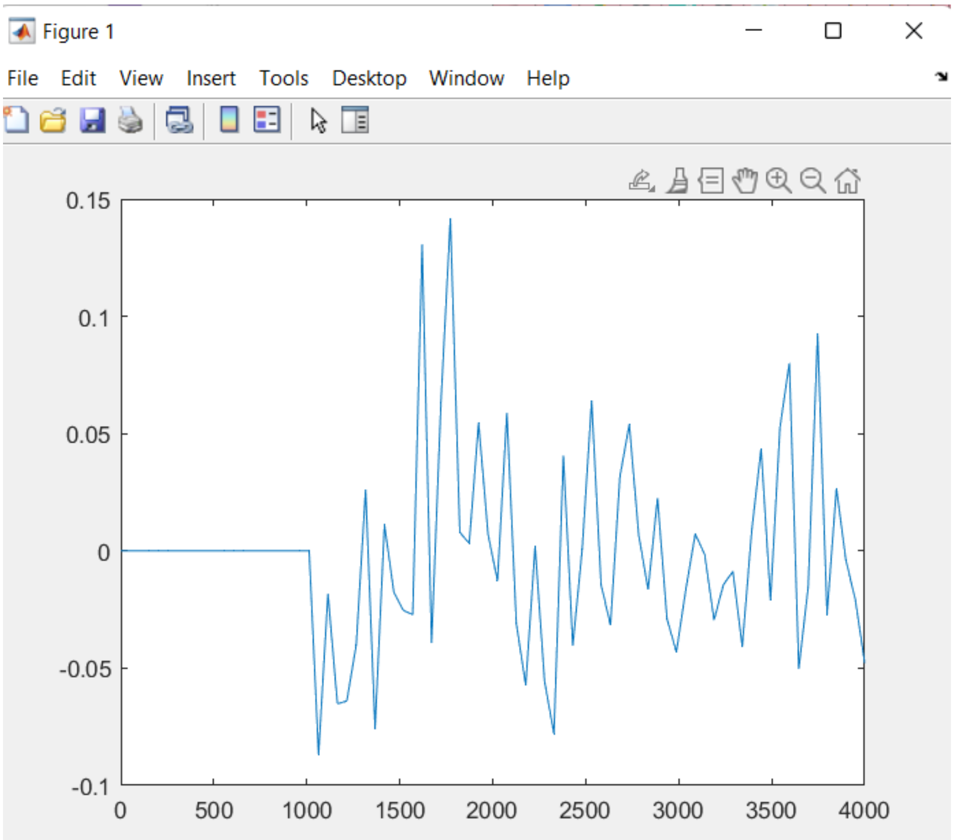
REAL CEPSTRUM from RCEPS function (same output)



Vocal tract Component(RCEPS(1:21))



SOURCE COMPONENT(RCEPS(21:160))



Following is the code used for above results.

MATLAB Code

```
[smpl,fs]=audioread('a.wav');%Read the audio speech signal input. We
took vowel 'a'.
smpl = resample(smpl(8801:9760,1),8000,48000);%Resampling the audio
signal
%First we calculated the real cepstrum of the input wave using the
inbuilt
%rceps model
ft=rceps(smpl);
%plot(linspace(0,4000,80),ft(1:80))
%Calculating the  Fourier Transform of the Sample
ft1=fft(smpl);
%plot(linspace(0,4000,80),ft1(1:80))
%Taking log of the transform to convert the product into sum of the
vocal
%tract and excitation component
lg=log(abs(ft1));
%plot(linspace(0,4000,80),lg(1:80))
%Taking Inverse Transform to get into the cepstrum domain
ift=ifft(lg);
%plot(linspace(0,4000,80),ift(1:80))
%The Source excitation occurs at higher quefrency
%So we  replaced the lower frequency samples with zeroes
ft(1:21,1) = 0;%Source Excitation
%plot(linspace(0,4000,80),ft(1:80))
%The vocal tract component occurs at lower quefrency. Thus Replacing
with zeroes at higher quefrency
ft(21:160,1) = 0;%Vocal Tract Component
%plot(linspace(0,4000,80),ft(1:80))
```

ANALYSIS & CONCLUSION

Thus , We extracted the source/excitation component(characterising words spoken and the pitch) and vocal tract components(characterising the quality of speech) from the given speech signal by linear separation of the convolved signal which is known as liftering. We didn't had any earlier knowledge of these components and their deconvolution would have been a tedious task due to their uncertainity.

But due to the linear separation in cepstrum domain(quefrency), We could trace back these components by knowing their concentration at higher and lower quefrency regions.

We concluded that we can apply this Cepstrum Analysis in various domains which includes recognition of the speech samples.Not only this we could even use it to detect certain respiratory diseases by analysis the coughing samples of an individual, a boon in COVID era for early detection. We can detect faults in the defective machines from their noise spectrum and can get the location of defect.Thus, Cepstral Analysis has a wide range of applications.

References:

1. https://ssp-iiith.vlabs.ac.in/exp/cepstral-analysis/index.html
2. https://en.wikipedia.org/wiki/Cepstrum
3. http://www.xavieranguera.com/tdp_2011/8-Cepstral-Analysis.pdf
4. https://www.youtube.com/watch?v=wEoNuCxMv0U
5. https://in.mathworks.com/help/signal/ref/rceps.html