Lab2 Fundamentals of Human Speech Production

张旭东 12011923

1.Introduction

The purpose of this lab is to continue learn MATLAB functionality for speech processing and test the effect of digital filtering on speech perception. In this experiment, several signal will be analyzed, including the phoneme, resonance peak and other parameters of each segment.

2.Lab result and analysis

2.1 Problem1

The purpose of problem1 is to learn and understand the process of `segmentation and labeling` of speech using the principle of phonetic.

After loading the speech file `s5.wav`, the sentence is: Oak is strong and also gives shade. The ARPAbet symbols and the IPA representation are shown as follow:

| Oak	is	strong	and	also	gives	shade
OW K	IH Z	S T R AO NX	AE EN D	AO L S OW	G IH V Z	SH EY D
[oʊk]	[ɪz]	[strɔːŋ]	[ənd , ənd]	[ˈɔːlsoʊ]	[gɪvz]	[ʃeɪd]

Fig.1 ARPAbet symbols and IPA representation

Then, A function `PlotAudioFile` used in Lab1 is used again to plot the sound and label it. The result is as follows:
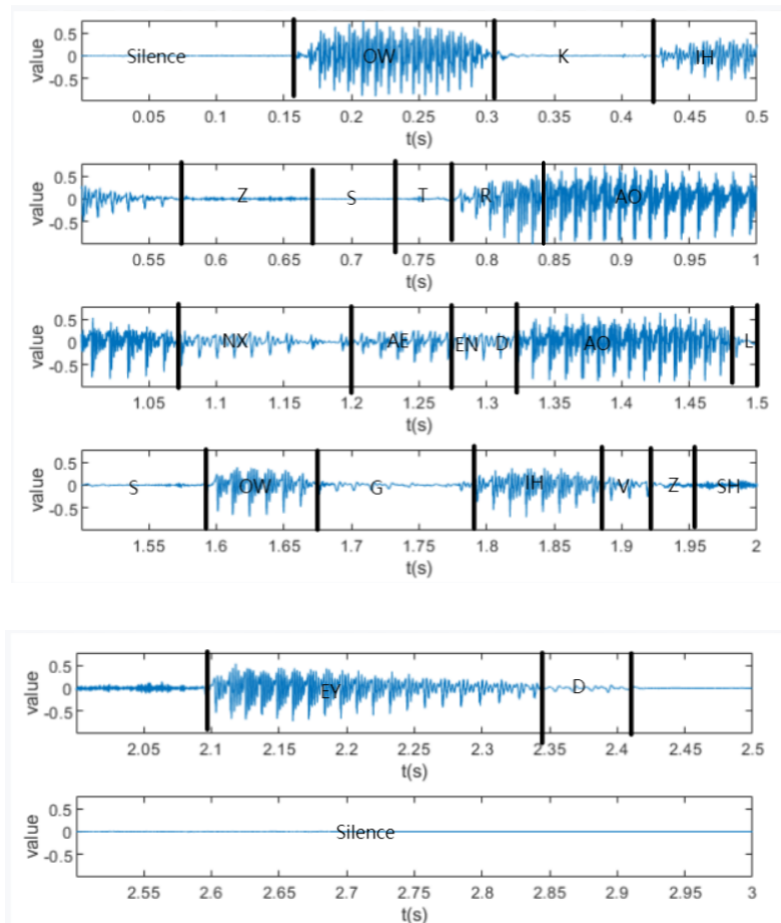
Fig.2 labeled waveform

What's more, We also use `colea` to plot and label each phoneme. The label result is as follows:
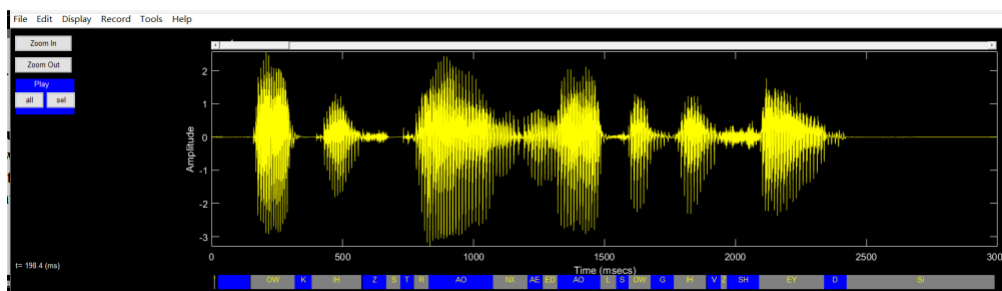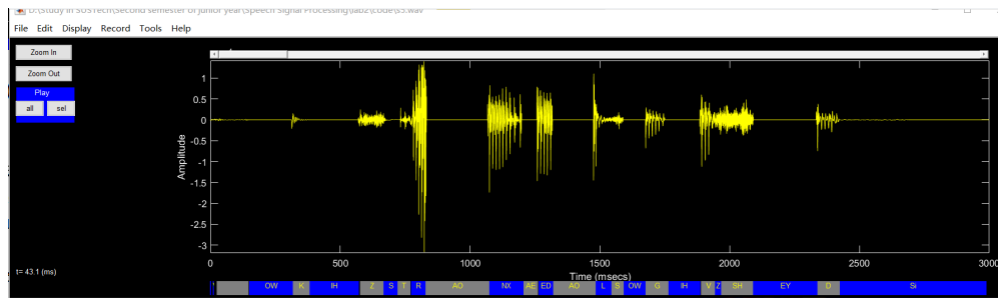

Fig.3 labeled waveform with colea

Then, we use `colea` to zero out the vowel regions of the sentence (replace the speech waveform with zero-valued samples during the vowel regions) for `colea`'s convenience. The result is as below:
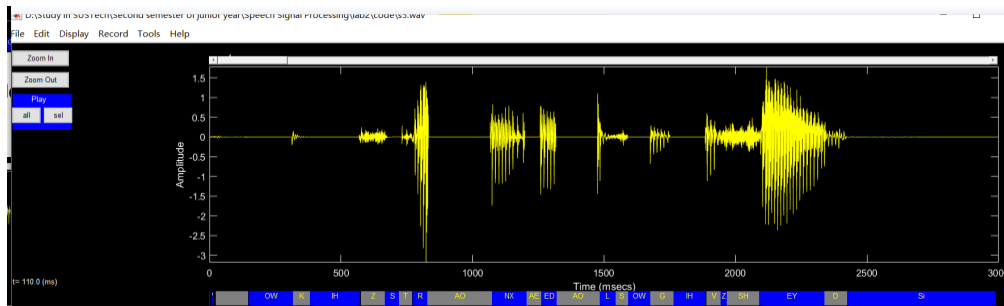
Fig.4 zero out the vowel regions without deleting diphthongs 'EY'

Fig.5 zero out the vowel regions with deleting diphthongs 'EY'

It is impossible to understand the meaning of the speech after hearing the speech which is handled by zeroing out the vowel region, which coincides with the theory in the lecture notes. Vowel has a big influence on hearing and understanding the the meaning of the speech.

Then, zeroing out all the consonant regions of the sentence is done.
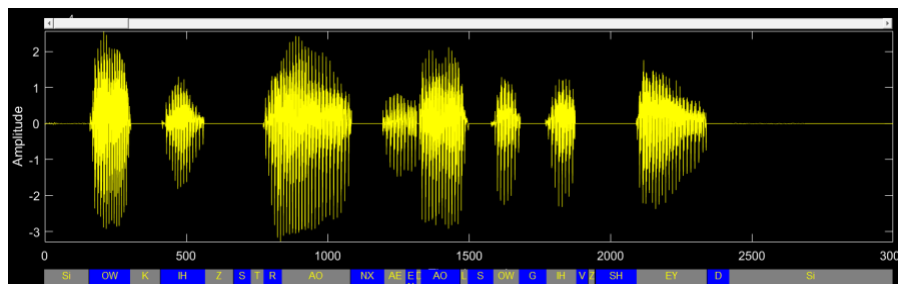
Fig.6 zero out the consonant regions

It is obvious that the latter sounds much better than the former and the information and meaning transferred by the speech can be generally understood, which implies consonant has little influence on hearing and understanding the the meaning of the speech.

Through the above comparison, the sentence without consonant has an advantage than that without vowel in conveying information.

2.2 Problem2

In this problem, the difference for peak of spectral envelope for several vowel sounds will be discussed and the difference of formant frequencies extracted from `FFT` and `Praat` analyses will be compared.

First, three vowels `/a/`, `/i/`, `/u/` , whose sample rate are 1.6 `kHz` are generated by using function `audiorecorder`.

```matlab
recorder=audiorecorder(16000,16,1);
%记录/a/
disp('start recording');
recordblocking(recorder,1);
disp('end recording');
a=getaudiodata(recorder);
filename='a.wav';
audiowrite(filename,a,16000);

%记录/i/
disp('start recording');
recordblocking(recorder,1);
disp('end recording');
i=getaudiodata(recorder);
filename='i.wav';
audiowrite(filename,i,16000);

%记录/u/
disp('start recording');
recordblocking(recorder,1);
disp('end recording');
u=getaudiodata(recorder);
filename='u.wav';
audiowrite(filename,u,16000);
```

Then, the waveform and spectrum of them are plotted.
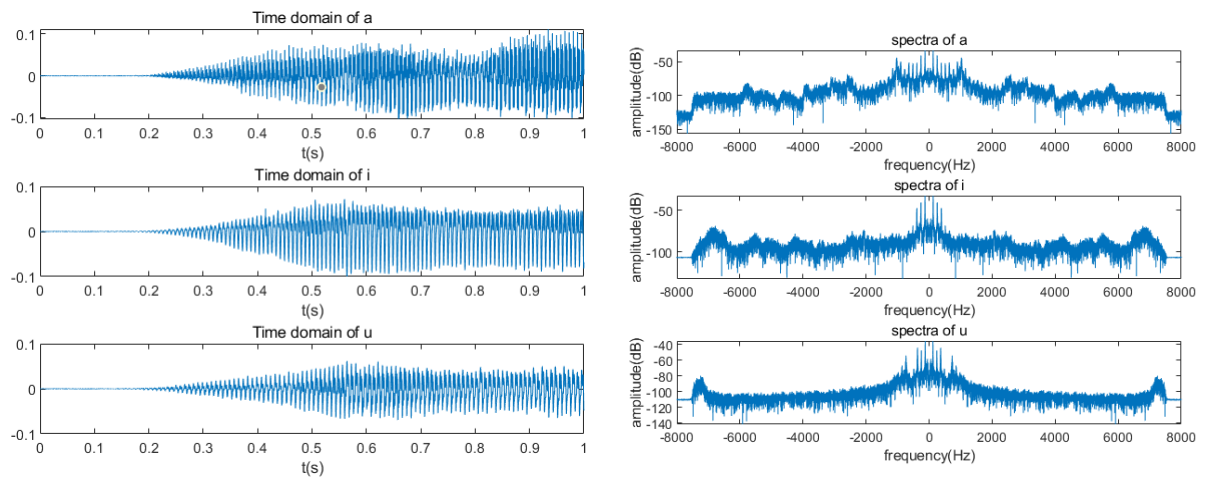
Fig.7 waveform and spectrum

To extract formant frequencies from FFT, the range of `x-axis` is set to $[0, 5000]$ because the maximum frequency of a male voice will not exceed 5,000 Hz.
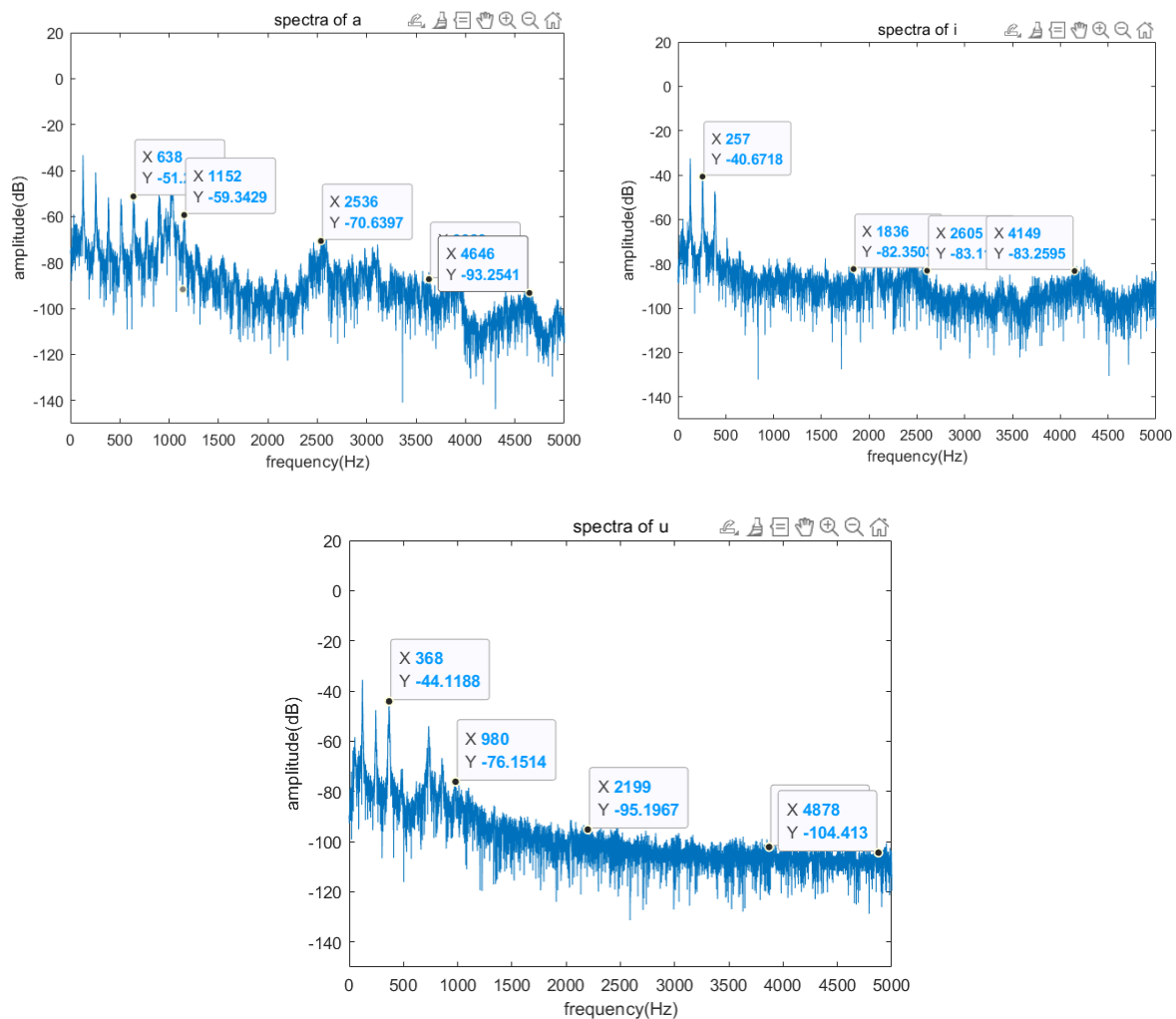
Fig.8 spectra analysis of sound

Then, we use `Praat` to analysis the three vowels to get the formant frequencies. To reduce the error, `Praat` don't record again.
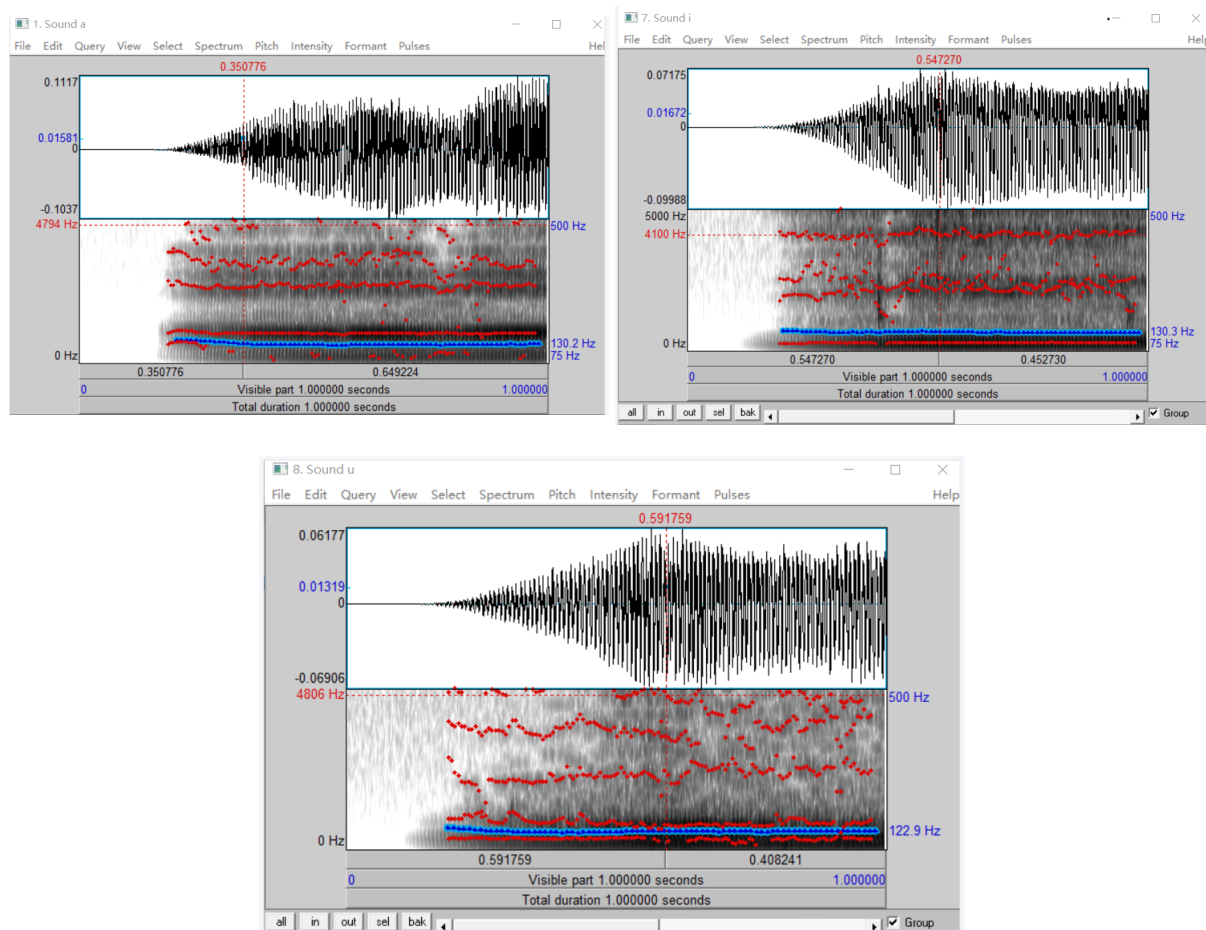
Fig.9 analysis from Praat

The formant frequencies extracted from `FFT` and `Praat` analysis are listed as below:

|		F0	F1	F2	F3	F4
a	MATLAB	638	1152	2536	3626	4646
	Praat	657.6	1141	2590	3586	4794
i	MATLAB	257	1836	2605	4149	
	Praat	265.1	1865	2620	4100	
u	MATLAB	368	980	2199	3870	4878
	Praat	347	908.3	2280	3777	4806

   It can't be ignored is that the value of formant frequency extracted from `FFT` is not very precise because the spectrum in MATLAB is continuous and it is hard to find out exactly where the formant frequencies are.

   Through the comparison of the value in the table, the formant frequencies extracted from `FFT` and `Praat` analysis are a little different. The reason is that the values of formant frequencies changes with time but we only select one segment to analyze.

2.3 Problem3

In this problem, the effect of filter will be discussed again and the process of designing a filter will be used again.

First, we filter the speech file `test_16k.wav` to bandwidths of 5.5, 4 and 3.2 `kHz` using the function `butter`.
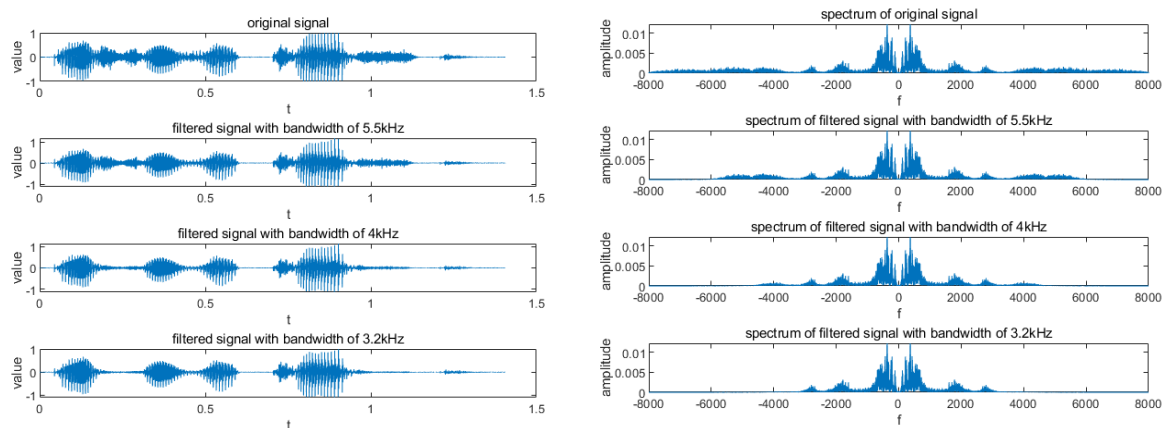

Fig.10 waveform and spectrum

From the second picture, it is obvious that the spectrum of filtered signal with bandwidth of 5.5 `kHz` has components in the frequency which exceeds 5.5 `kHz`. So as the rest two. The reason is that the magnitude response of `butterworth` filter has ripples in the frequency which exceeds 5.5 `kHz`. The details of `butterworth` filter have been discussed in report of `Lab1`.

After hearing the three results, it is obvious that three kinds of audio with different bandwidth are easy to understand. However, the smaller the cut-off frequency of low-pass filter, the more blurred the sound and the volume decreases due to the lowpass filter reduces the amplitude of the signal slightly.

Then, a function `fdatool` is used to design a digital high-pass filter, with appropriate cutoff frequencies to attenuated 60 `kHz` components by at least 40 dB.
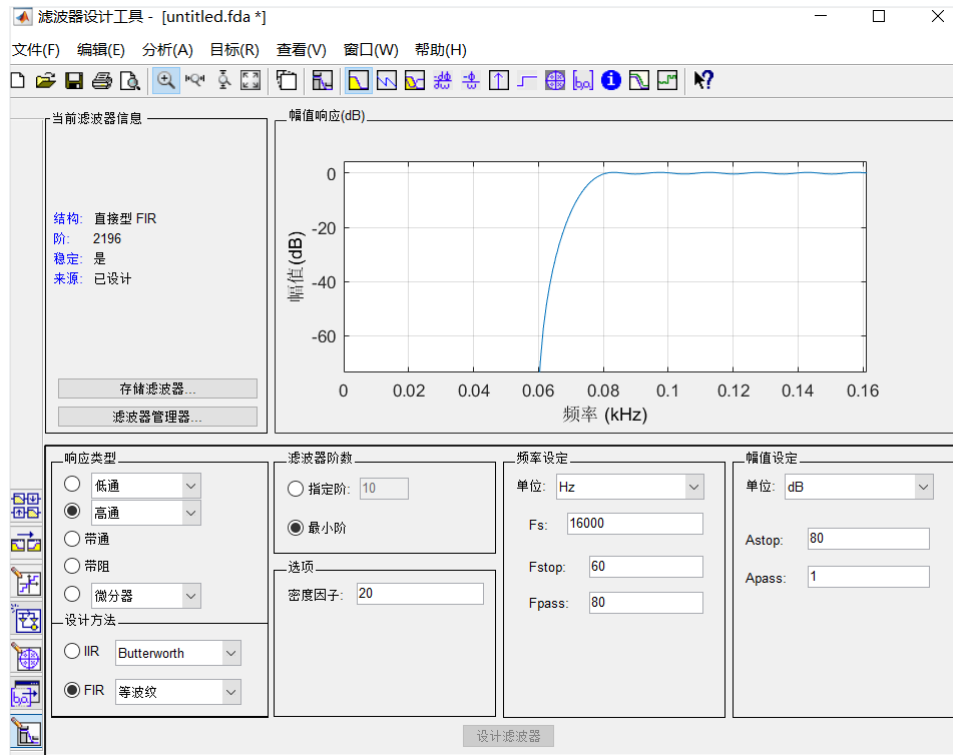
Fig.11 high-pass filter design

The code of the digital high-pass filter is below:

```
function Hd = highpass
%HIGHPASS  返回离散时间滤波器对象。

% MATLAB Code
% Generated by MATLAB(R) 9.10 and Signal Processing Toolbox 8.6.
% Generated on: 28-Feb-2023 22:42:41

% Equiripple Highpass filter designed using the FIRPM function.

% All frequency values are in Hz.
Fs = 16000;  % Sampling Frequency

Fstop = 60;              % Stopband Frequency
Fpass = 80;              % Passband Frequency
Dstop = 0.0001;          % Stopband Attenuation
Dpass = 0.057501127785;  % Passband Ripple
dens  = 20;              % Density Factor

% Calculate the order from the parameters using FIRPMORD.
[N, Fo, Ao, W] = firpmord([Fstop, Fpass]/(Fs/2), [0 1], [Dstop, Dpass]);
```
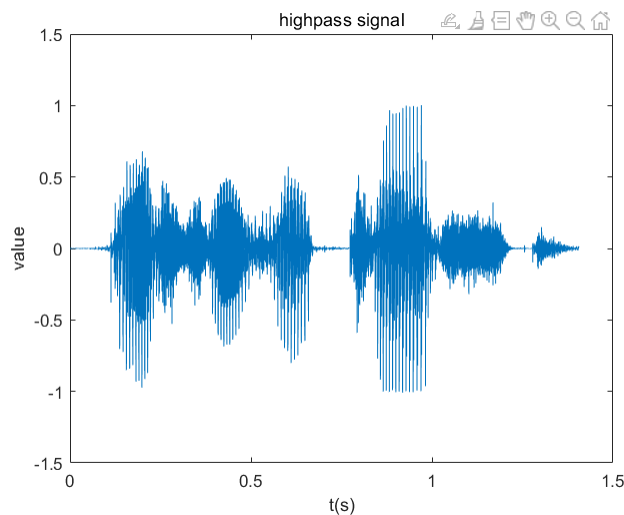
```
% Calculate the coefficients using the FIRPM function.
b  = firpm(N, Fo, Ao, W, {dens});
Hd = dfilt.dffir(b);

% [EOF]
```


Fig.12 filtered speech signal

Compared to the original signal, there is a little difference between them in hearing. This is because the components between -60 `kHz` and 60 `kHz` take relatively small part.

3.Conclusion

In this lab, the most important harvest is to have learned how to use `colea` and `Praat` to analysis and operate speech file. Besides, the influence of vowel and consonant on hearing and understanding the meaning of sentence has been discussed. What's more, the method of extracting formant frequencies is so convenient that it is easy to master.