Digital Signal Processing HW5 MATLAB Part

Name: Jingyang Zhang
Net ID: jz2807

1. Speech Recording. Record yourself saying a vowel sound and a consonant sound, respectively. You can create your recordings using a computer, smart phone, or other digital device. There are many audio recording utilities for both MS and Mac systems. You can also use MATLAB audiorecorder( ) function. A good thing with the audiorecorder( ) function is that you can specify the desired sampling frequency and number of bits per sample. For human speech, sampling rate of 8 KHz and bit depth of 8 bits are typically considered sufficient. However using a higher sampling rate and more bits/sample can give you better recording quality. Import your audio file into MATLAB using the audioread( ) function or other suitable function. To verify that your file is correctly imported into MATLAB, use the sound function to play the audio signals: sound(x, fs). You will need to specify the sampling rate fs. In MATLAB, plot both the entire signal as well as a short segment (e.g. 50 milliseconds) in the middle of the sound. Comment on how the two waveforms are different. The 'vowel' signal should be approximately periodic. The period is called the pitch period. What is the pitch period of your 'vowel' sound as measured in units of milliseconds?

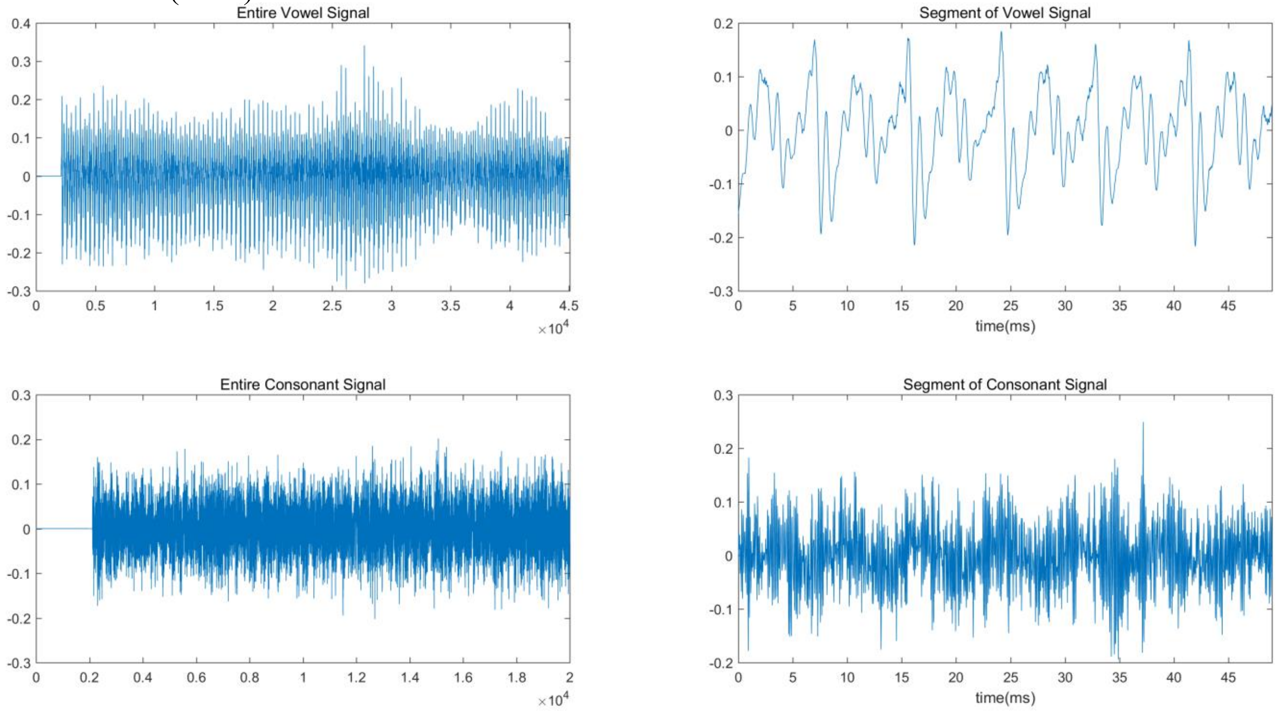
Solution:

.m file(s): jyz_HW5_1.m

Code:

```
clear
close all

[vowel,Fs1] = audioread('A.m4a');
[consonant, Fs2] = audioread('s.m4a');
% sound(vowel, Fs1)
% sound(consonant, Fs2)
vmiddle = vowel(20001:22205);
cmiddle = consonant(20001:22205);
range = 0:49/2204:49;
subplot(2,2,1)
plot(vowel)
xlim([0,45055])
title('Entire Vowel Signal')
subplot(2,2,2)
plot(range,vmiddle)
xlabel('time(ms)')
xlim([0 49])
title('Segment of Vowel Signal')
subplot(2,2,3)
plot(consonant)
xlim([0 20000])
title('Entire Consonant Signal')
subplot(2,2,4)
plot(range,cmiddle)
xlabel('time(ms)')
xlim([0 49])
title('Segment of Consonant Signal')
```

Result(Plots):



Comment:

The plot of the whole signal looks much denser than that of signal segment. The plot of Vowel Signal shows periodicity while Consonant signal does not. There are about 4 periods in 35 milliseconds, so the pitch period of my 'vowel' sound is about 8.75ms.

2. Speech Spectra. Use the MATLAB fft function to approximately compute the discrete-time Fourier transform (DTFT) of the 50 millisecond segments you plotted in the previous exercise. Note that it is best to specify a FFT length that is a power of 2 that is closest to the length of your segment. Plot the magnitude frequency-spectrum for both the sounds. Plot each spectrum on both linear and log scales. (For the log scale, use 20 log10 |X| for dB). Comment on your observations. The spectrum of 'vowel' sound should have distinct peaks that are should have peaks at approximately equally-spaced intervals (these are harmonics). Does the first peak frequency corresponds to the pitch period you observed in part 1? Also plot the short term Fourier transform using the "spectrogram( )" function to see how the spectrum changes slightly over time.

Solution:

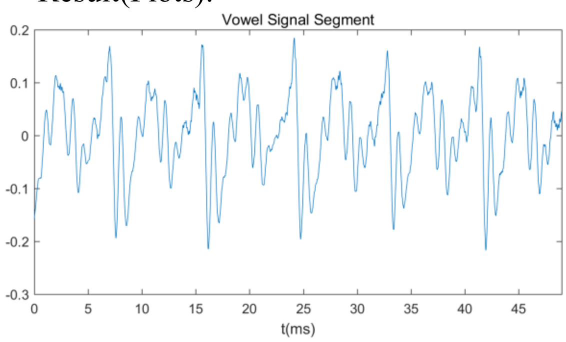
.m file(s): jyz_HW5_2.m

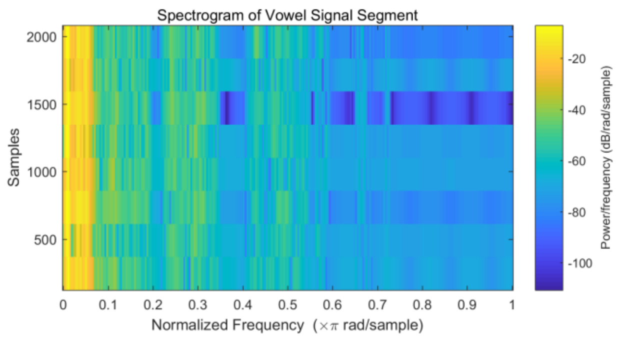
Code:

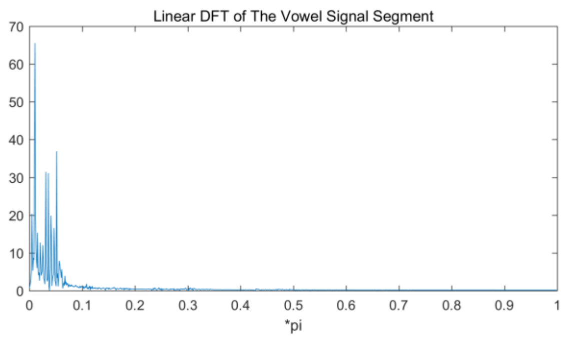
```
clear
close all

[vowel,Fs1] = audioread('A.m4a');
[consonant, Fs2] = audioread('s.m4a');
% sound(vowel, Fs1)
% sound(consonant, Fs2)
vmiddle = vowel(20001:22205);
cmiddle = consonant(20001:22205);
range = 0:49/2204:49;
vDTFT=fft(vmiddle(1:2048));
cDTFT=fft(cmiddle(1:2048));
```

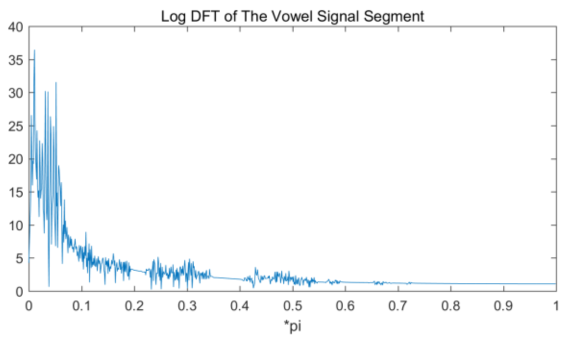
```
figure(1)
subplot(2,2,1);
plot(range,vmiddle)
title('Vowel Signal Segment');
xlabel('t(ms)')
xlim([0, 49])
subplot(2,2,2);
spectrogram(vmiddle); title('Spectrogram of Vowel Signal Segment');
subplot(2,2,3);
plot(0:1/1024:1-1/1024,(abs(vDTFT(1:1024)))), title('Linear DFT of The Vowel Signal Segment')
xlabel('*pi')
subplot(2,2,4);
plot(0:1/1024:1-1/1024,20*log10(abs(vDTFT(1:1024))+1)); title('Log DFT of The Vowel Signal Segment');
xlabel('*pi')

figure(2)
subplot(2,2,1);
plot(range,cmiddle); title('Consonant Signal Segment');
xlabel('t(ms)')
xlim([0, 49])
subplot(2,2,2);
spectrogram(cmiddle); title('Spectrogram of Consonant Signal Segment');
subplot(2,2,3);
plot(0:1/1024:1-1/1024,(abs(cDTFT(1:1024)))), title('Linear DFT of The Consonant Signal Segment')
xlabel('*pi')
subplot(2,2,4);
plot(0:1/1024:1-1/1024,20*log10(abs(cDTFT(1:1024))+1)); title('Log DFT of The Consonant Signal Segment');
xlabel('*pi')
```

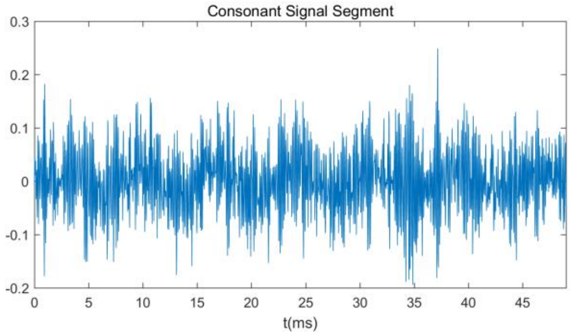
Result(Plots):

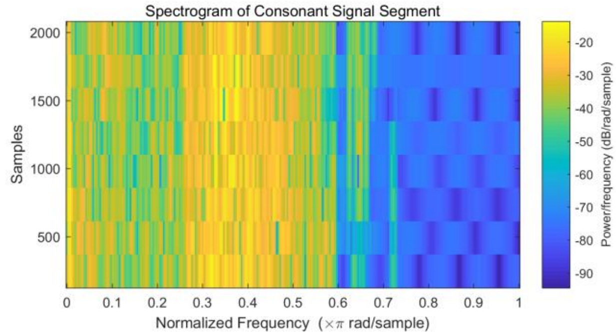


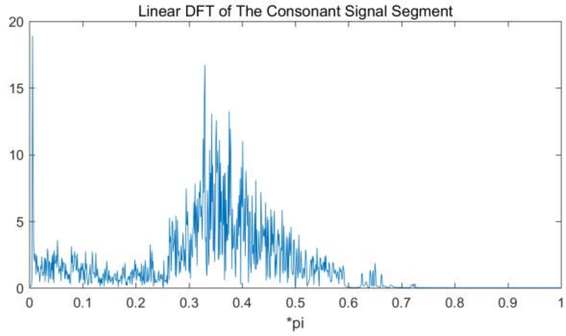


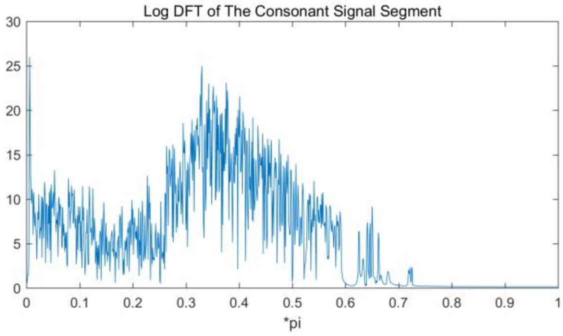




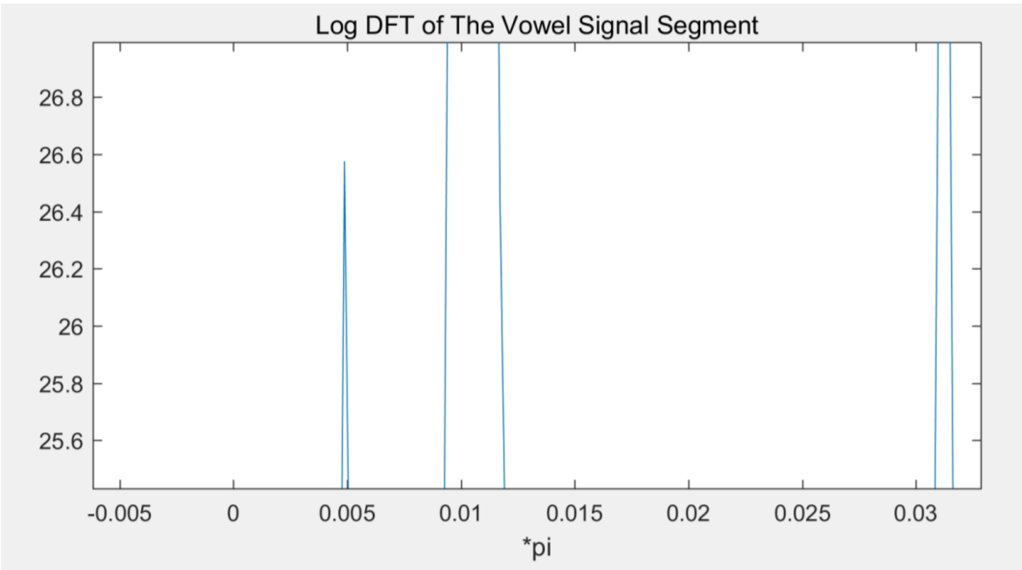








Zoom in the plot of '*Log DFT of The Vowel Signal Segment*', I get:



Comment: As we can see in the plots, the DFT of The Vowel Signal Segment has distinct peaks that have peaks at approximately equally-spaced intervals, while the DFT of The Consonant Signal does not. Zoom in the plot of '*Log DFT of The Vowel Signal Segment*', we can see the first peak is approximately at $\omega_0 = 0.005\pi$. We can get f_2:

$$f_2 = \frac{\omega_0}{2\pi} \times Fs \;=\; \frac{0.005\pi}{2\pi} \times 44100Hz \;\approx 110\text{Hz}$$

In the part 1, I draw the conclusion that the pitch period of my vowel sound is about 8.75ms through observation. Let's find f_1:

$$f_1 = \frac{1}{0.00875s} \approx 114\text{Hz}$$

So it is easy to find that

$$f_1 \approx f_2$$

The first peak frequency corresponds to the pitch period observed in part 1.

From the spectrums of 'spectrogram()' function, we can see the spectrum of vowel signal changes slightly over time, the spectrum of consonant signal changes more rapidly, but generally they remains steady.