ECE 251A Homework 8

Wenyu Zhang

TOTAL POINTS

100 / 100

QUESTION 1

1 HW8 **100 / 100**

✓ **- 0 pts** Correct

- 5 pts Part B1: Plot(s) incorrect/ missing

- 5 pts Part B2: Plot(s) incorrect/ missing

- 5 pts Results not commented properly

- 5 pts Frequency axis incorrect

- 10 pts Late submission

Homework #8 Speech Processing

Wenyu Zhang
ECE 251A

March 13, 2018

1 Text

1.1 Objective

In this homework, we do spectral analysis(e.g. spectral magnitude, linear prediction and autocorrelation analysis) of digital series of certain phonemes to extract their characteristics on frequency.

1.2 Background

In digital speech processing, sentences could divided into words, then phonemes as time series formation.
To do things like speech recognition, it's significant to analyze the characteristic of certain phonemes. Fortunately, phonemes do have characteristics on spectrum. To extract their properties with disturb of personal voice, inverse filtering is a good method(see "Digital Inverse Filtering- A New Tool for Formant Trajectory Estimation" by John D. Markel in 1972).
In our homework, we use inverse filtering to extract three phonemes' spectral characteristics and compare it with previous experiments(e.g. *http://ec-concord.ied.edu.hk/phonetics_and_phonology/wordpress/learning_website/chapter_3_consonants_new.htm*)

1.3 Approach

In this homework, we follow steps as below:
1. With sample time series of three phonemes(i.e. $/i/, /u/, /\eta/$), plot 256-point time series and 256-FFT of 256/64-points time series)
2. Autocorrelation method of linear prediction. With time series above, do all-zero filtering of time series and analyze the autocorrelation of filter coefficients on spectrum.

1.4 Results and Analysis

1.4.1 Part A. Conventional Spectral Analysis

1. 256-point time series.
The plot of 256-point series is shown in **Figure 1**.
Comment:
From the plot we can see the periodic characteristic of phonemes on time region. This indicates that those time series have certain frequency components on spectrum.

2. Spectrum of 256-point series.
The plot of magnitude of 256-point series' FFT is shown in **Figure 2**.
Comment:
With theoretical frequency components we know, we can check the peaks at certain frequency. From the plot, we can see the peaks exist near the theoretical value. However, as personal influence and noise to the sample, there are some ripples that vibrate through the whole frequency interval.

3. Spectrum of 64-point series. The plot of magnitude of 256-FFT on 64-point time series is shown in **Figure 3**.
Comment:
Compared it with **Figure 2**, we can see that as lower sample rate on frequency, the ripples become fewer so that the peaks become more clear while loss of spectral details.

1.4.2 Part B. Autocorrelation Method of Linear Prediction

1. E_p v.s. order p.
The plot of E_p v.s. order p of 256-point time series is shown in **Figure 4**. The plot of E_p v.s. order p of 64-point time series is shown in **Figure 6**.
Comment:
In general, the error decreases with the increase of filter order, which corresponds to the characteristic of inverse filter in linear prediction.
Compared **Figure 4** and **Figure 6**, we can see that the error of 256-point series prediction is smaller than 64-point series prediction for more spectral details.

2. Spectrum of $10log(\frac{1}{|\hat{A}(k)|^2})$.
The plot of spectrum of $10log(\frac{1}{|\hat{A}(k)|^2})$ of 256-point time series is shown in **Figure 5**. The plot of spectrum of $10log(\frac{1}{|\hat{A}(k)|^2})$ of 64-point time series is shown in **Figure 7**.
Comment:
Generally, the linear prediction filters the ripples and extract the frequency components we want. And the peaks are nearly at the theoretical frequency values.
When analyzing spectral characteristics of certain phoneme, we can see that the performance of /i/ is good.
The F3 of phoneme /u/ is not clear, which corresponds to the blurry of that frequency in spectrum analysis. The reason for this may be disturb in sampling this phoneme.
When it comes to /η/, we can see that the F1 and F3 are clear. However, the F2 is not clear for the flatness between F1 and F3. In fact, the theoretical value of F2 is also not available(see *http://ec-concord.ied.edu.hk/phonetics_and_phonology/wordpress/learning_website/chapter_3_consonants_new.htm*). This may be the characteristic of phoneme itself. As a result, we need to focus on F1, F3 and F4 if we want to extract the spectral characteristics of /η/.
Finally, compared **Figure 5** and **Figure 7**, we can see that spectrum of 64-point is more smooth with little disturb, while that of 256-point have higher magnitude at peaks, which keeps more details of both frequency components and disturb.

1.5 Summary

In this homework, we successfully do both spectrum analysis and linear prediction of three phonemes. After comparing the spectrum with theoretical value, we comments the performance and possible cause. We can see the correspondence of spectral characteristics between conventional spectrum and linear prediction, while the performance of linear prediction is better.

2 Plots

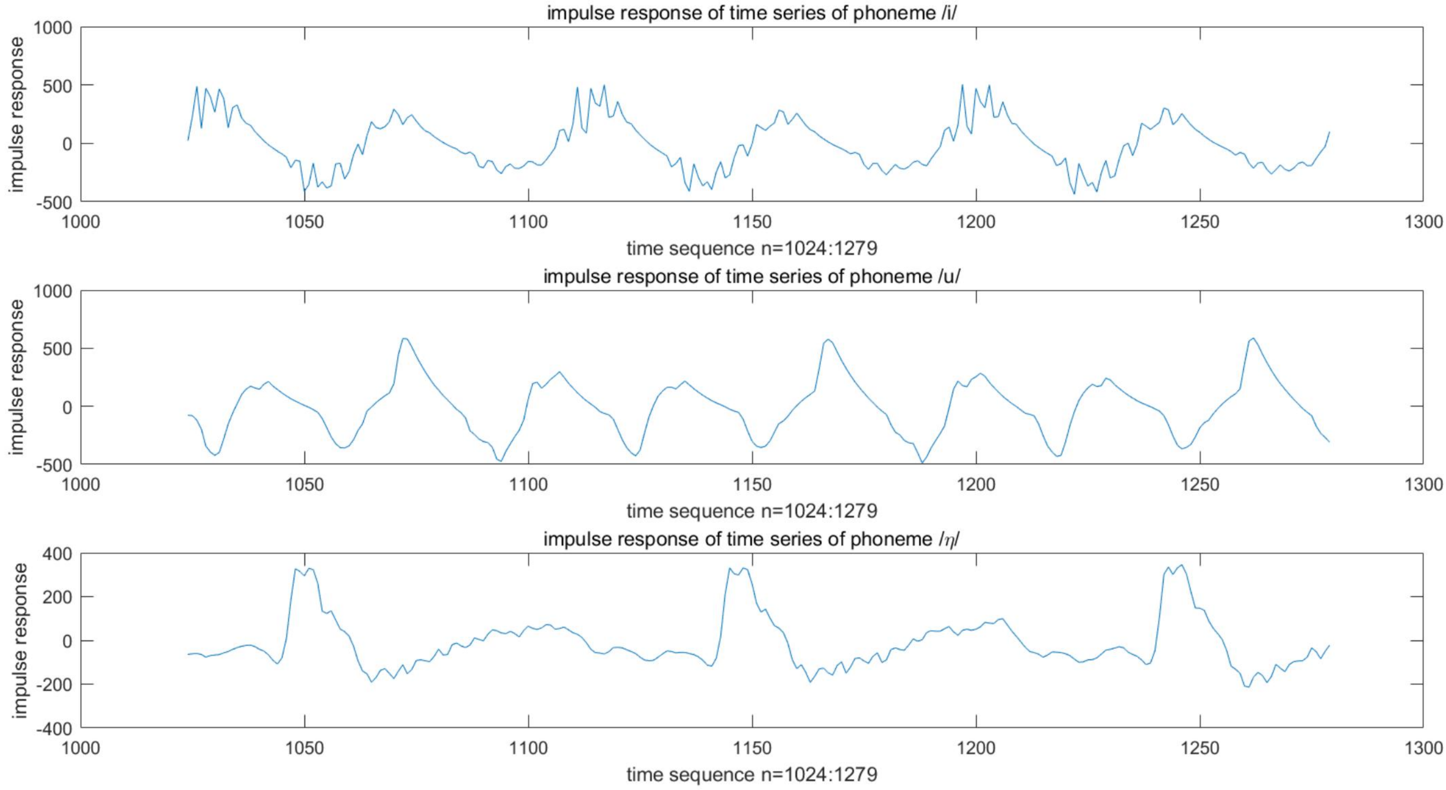

Figure 1: Time Series of Phonemes

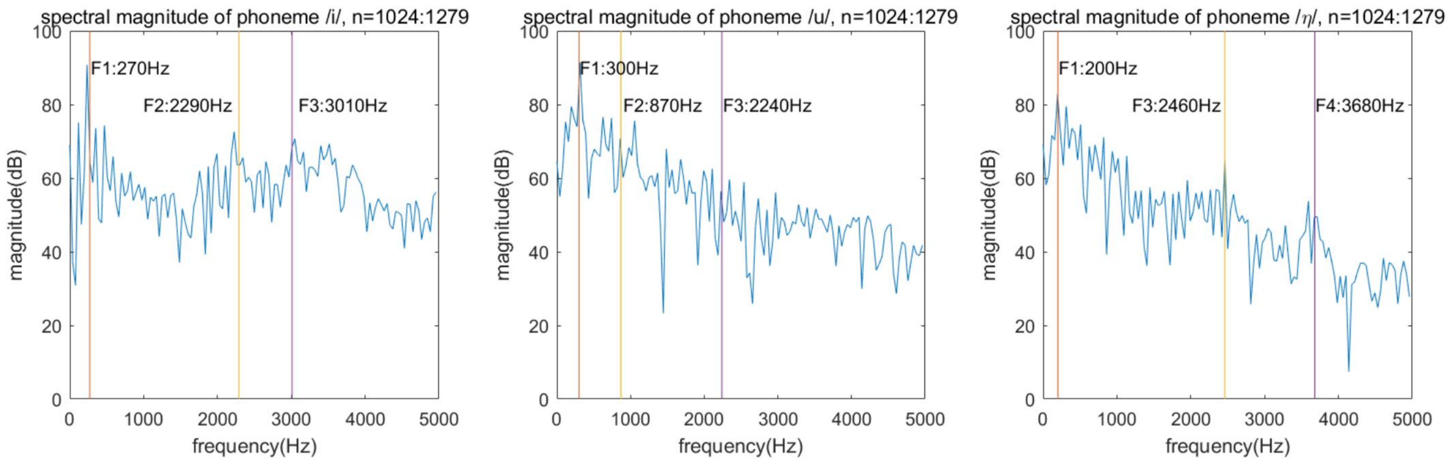

Figure 2: Spectrum of Phonemes, 256-point time series

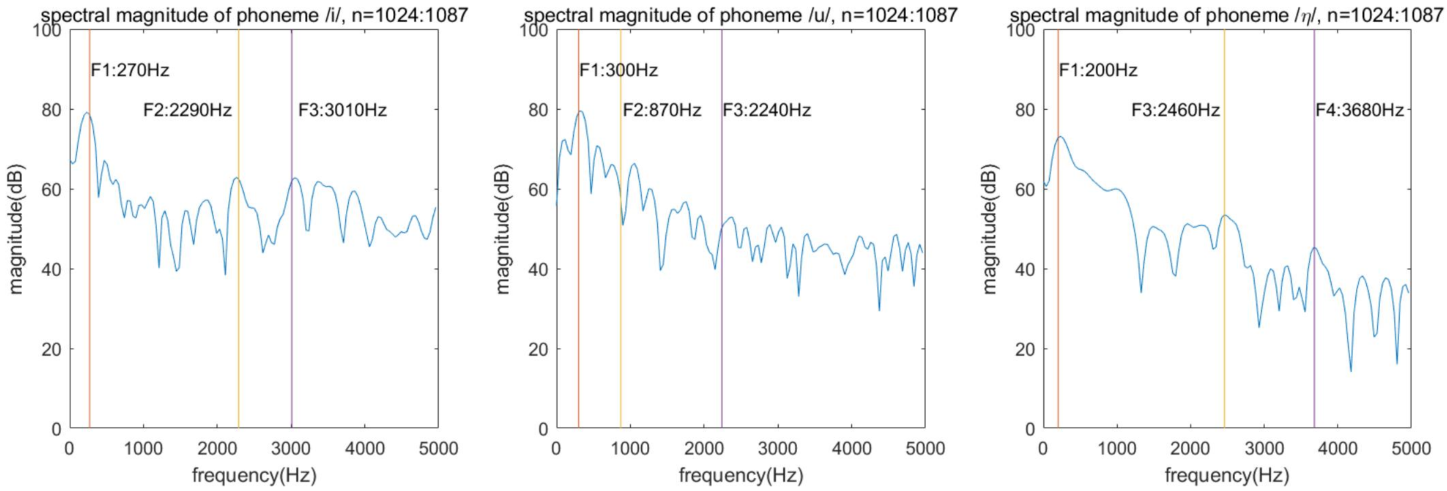

Figure 3: Spectrum of Phonemes, 64-point time series

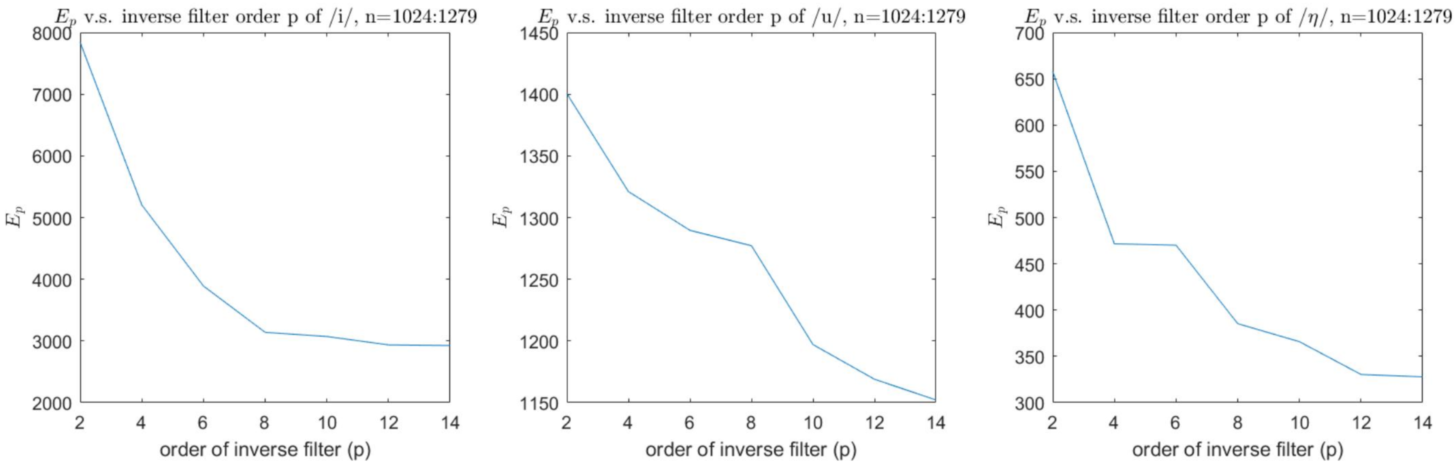

Figure 4: E_p v.s. Inverse Filter Order p, n=1024:1279

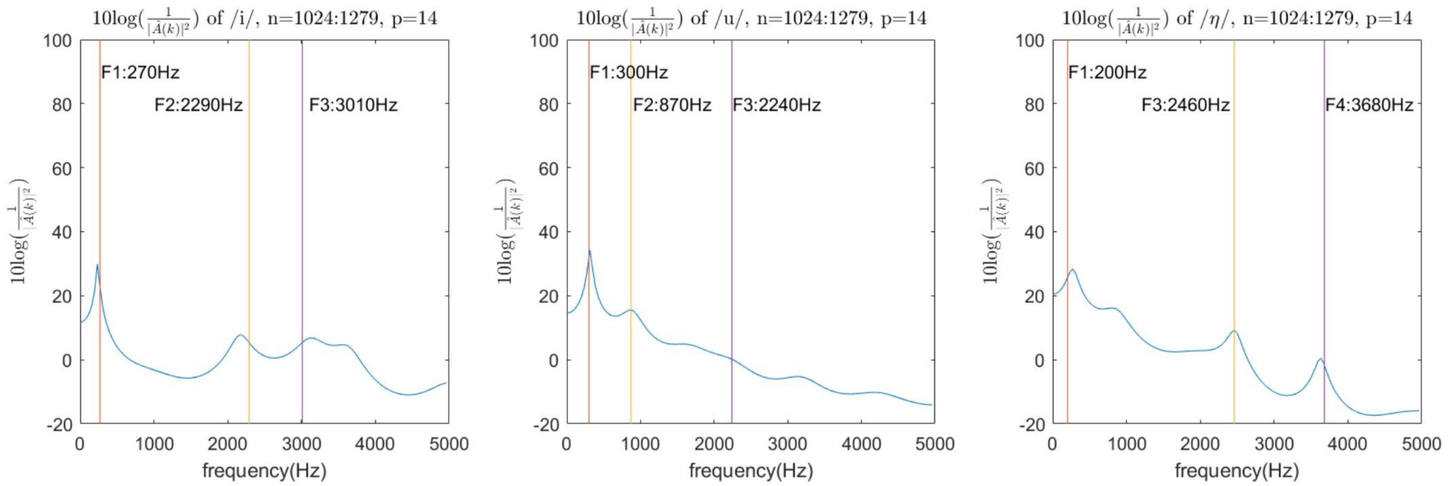

Figure 5: Spectrum of $10\log(\frac{1}{|\hat{A}(k)|^2})$, n=1024:1279

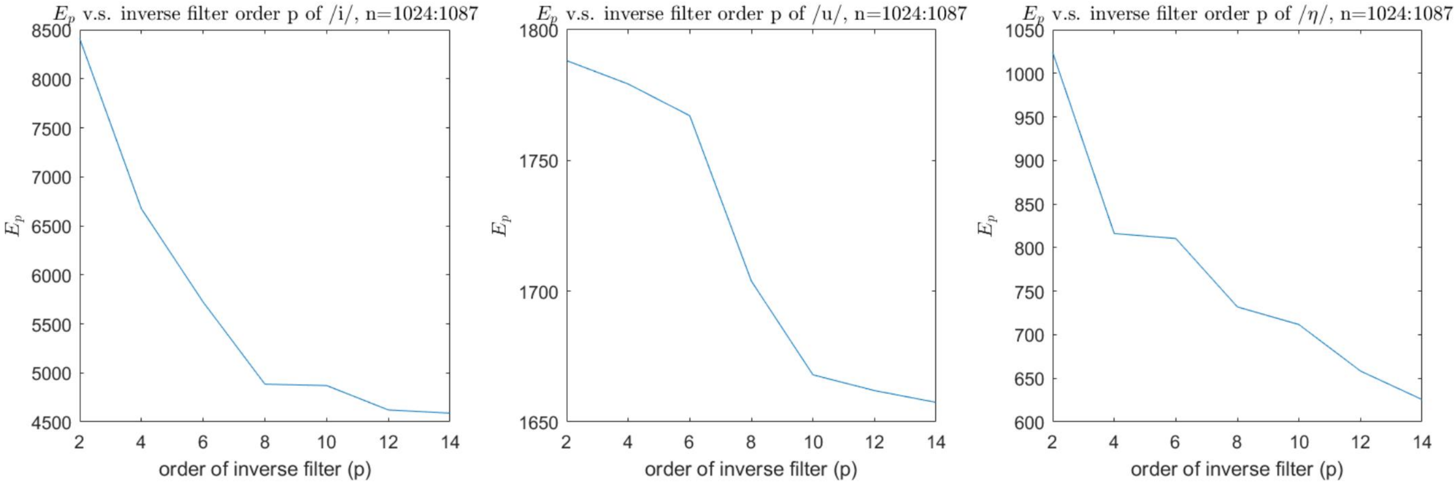

Figure 6: E_p v.s. Inverse Filter Order p, n=1024:1087

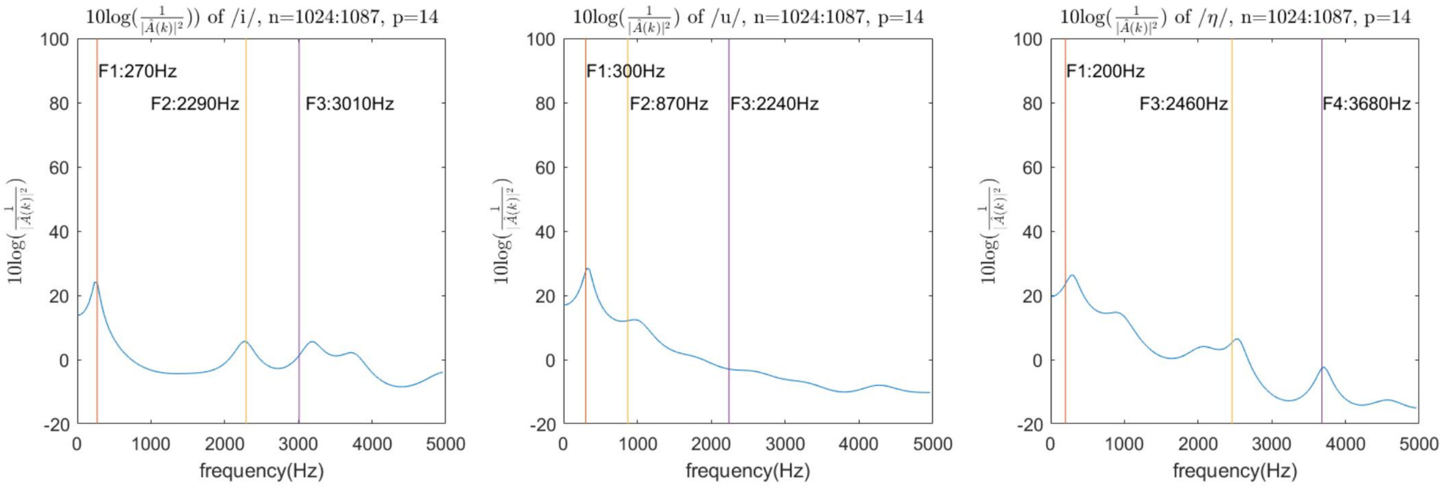

Figure 7: Spectrum of $10\log(\frac{1}{|\hat{A}(k)|^2})$, n=1024:1087

3 Appendix

```
% Homework 8
% A. Conventional Spectral Analysis
I = xlsread('phonemes.xlsx','A1025:A1280');
U = xlsread('phonemes.xlsx','B1025:B1280');
NG = xlsread('phonemes.xlsx','C1025:C1280');
n = 1024:1279;
figure(1);
subplot(3,1,1);
plot(n,I);
xlabel('time sequence n=1024:1279');
ylabel('impulse response');
title('impulse response of time series of phoneme /i/');
subplot(3,1,2);
plot(n,U);
xlabel('time sequence n=1024:1279');
ylabel('impulse response');
title('impulse response of time series of phoneme /u/');
subplot(3,1,3);
plot(n,NG);
xlabel('time sequence n=1024:1279');
ylabel('impulse response');
title('impulse response of time series of phoneme /\eta/');

H_I = fft(I,256);
H_U = fft(U,256);
H_NG = fft(NG,256);
L_I = zeros(1,128);
L_U = zeros(1,128);
L_NG = zeros(1,128);
for i = 1:128
    L_I(i) = mag2db(abs(H_I(i)));
    L_U(i) = mag2db(abs(H_U(i)));
    L_NG(i) = mag2db(abs(H_NG(i)));
end
N = 0:127;
Nf = N*10000/256;
figure(2)
subplot(1,3,1);
plot(Nf,L_I,[270,270],[0,100],[2290,2290],[0,100],[3010,3010],[0
,100]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(280,90,'F1:270Hz');
text(1000,80,'F2:2290Hz');
text(3100,80,'F3:3010Hz');
```

```
title('spectral magnitude of phoneme /i/, n=1024:1279');
subplot(1,3,2);
plot(Nf,L_U,[300,300],[0,100],[870,870],[0,100],[2240,2240],[0,1
00]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(310,90,'F1:300Hz');
text(900,80,'F2:870Hz');
text(2250,80,'F3:2240Hz');
title('spectral magnitude of phoneme /u/, n=1024:1279');
subplot(1,3,3);
plot(Nf,L_NG,[200,200],[0,100],[2460,2460],[0,100],[3680,3680],[
0,100]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(210,90,'F1:200Hz');
text(1200,80,'F3:2460Hz');
text(3690,80,'F4:3680Hz');
title('spectral magnitude of phoneme /\eta/, n=1024:1279');

I(65:256) = zeros(192,1);
U(65:256) = zeros(192,1);
NG(65:256) = zeros(192,1);
H_I = fft(I,256);
H_U = fft(U,256);
H_NG = fft(NG,256);
L_I = zeros(1,128);
L_U = zeros(1,128);
L_NG = zeros(1,128);
for i = 1:128
    L_I(i) = mag2db(abs(H_I(i)));
    L_U(i) = mag2db(abs(H_U(i)));
    L_NG(i) = mag2db(abs(H_NG(i)));
end
N = 0:127;
Nf = N*10000/256;
figure(3)
subplot(1,3,1);
plot(Nf,L_I,[270,270],[0,100],[2290,2290],[0,100],[3010,3010],[0
,100]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(280,90,'F1:270Hz');
text(1000,80,'F2:2290Hz');
text(3100,80,'F3:3010Hz');
title('spectral magnitude of phoneme /i/, n=1024:1087');
subplot(1,3,2);
```

```
plot(Nf,L_U,[300,300],[0,100],[870,870],[0,100],[2240,2240],[0,1
00]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(310,90,'F1:300Hz');
text(900,80,'F2:870Hz');
text(2250,80,'F3:2240Hz');
title('spectral magnitude of phoneme /u/, n=1024:1087');
subplot(1,3,3);
plot(Nf,L_NG,[200,200],[0,100],[2460,2460],[0,100],[3680,3680],[
0,100]);
xlabel('frequency(Hz)');
ylabel('magnitude(dB)');
text(210,90,'F1:200Hz');
text(1200,80,'F3:2460Hz');
text(3690,80,'F4:3680Hz');
title('spectral magnitude of phoneme /\eta/, n=1024:1087');

%B. Autocorrelation Method of Linear Prediction
I = xlsread('phonemes.xlsx','A1025:A1280');
U = xlsread('phonemes.xlsx','B1025:B1280');
NG = xlsread('phonemes.xlsx','C1025:C1280');
I = I.*kaiser(256,1.5);
U = U.*kaiser(256,1.5);
NG = NG.*kaiser(256,1.5);

[~,g2] = lpc(I,2);
[~,g4] = lpc(I,4);
[~,g6] = lpc(I,6);
[~,g8] = lpc(I,8);
[~,g10] = lpc(I,10);
[~,g12] = lpc(I,12);
[~,g14] = lpc(I,14);
p = [2,4,6,8,10,12,14];
figure(4)
subplot(1,3,1);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
title('$E_p$ v.s. inverse filter order p of /i/,
n=1024:1279','Interpreter','latex');

[~,g2] = lpc(U,2);
[~,g4] = lpc(U,4);
[~,g6] = lpc(U,6);
[~,g8] = lpc(U,8);
[~,g10] = lpc(U,10);
```

```
[~,g12] = lpc(U,12);
[~,g14] = lpc(U,14);
p = [2,4,6,8,10,12,14];
subplot(1,3,2);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
title('$E_p$ v.s. inverse filter order p of /u/,
n=1024:1279','Interpreter','latex');

[~,g2] = lpc(NG,2);
[~,g4] = lpc(NG,4);
[~,g6] = lpc(NG,6);
[~,g8] = lpc(NG,8);
[~,g10] = lpc(NG,10);
[~,g12] = lpc(NG,12);
[~,g14] = lpc(NG,14);
p = [2,4,6,8,10,12,14];
subplot(1,3,3);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
title('$E_p$ v.s. inverse filter order p of /$\eta$/,
n=1024:1279','Interpreter','latex');

I14 = zeros(1,256);
[I14(1:15),~] = lpc(I,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(I14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(5)
subplot(1,3,1)
plot(Nf,L,[270,270],[-20,100],[2290,2290],[-
20,100],[3010,3010],[-20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(280,90,'F1:270Hz');
text(1000,80,'F2:2290Hz');
text(3100,80,'F3:3010Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$) of /i/, n=1024:1279,
p=14','Interpreter','latex');
```

```
U14 = zeros(1,256);
[U14(1:15),~] = lpc(U,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(U14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(5)
subplot(1,3,2)
plot(Nf,L,[300,300],[-20,100],[870,870],[-20,100],[2240,2240],[-
20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(310,90,'F1:300Hz');
text(900,80,'F2:870Hz');
text(2250,80,'F3:2240Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$) of /u/, n=1024:1279,
p=14','Interpreter','latex');

NG14 = zeros(1,256);
[NG14(1:15),~] = lpc(NG,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(NG14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(5)
subplot(1,3,3)
plot(Nf,L,[200,200],[-20,100],[2460,2460],[-
20,100],[3680,3680],[-20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(210,90,'F1:200Hz');
text(1200,80,'F3:2460Hz');
text(3690,80,'F4:3680Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$) of /$\eta$/,
n=1024:1279, p=14','Interpreter','latex');

I = xlsread('phonemes.xlsx','A1025:A1088');
U = xlsread('phonemes.xlsx','B1025:B1088');
NG = xlsread('phonemes.xlsx','C1025:C1088');
```

```
I = I.*kaiser(64,1.5);
U = U.*kaiser(64,1.5);
NG = NG.*kaiser(64,1.5);

[~,g2] = lpc(I,2);
[~,g4] = lpc(I,4);
[~,g6] = lpc(I,6);
[~,g8] = lpc(I,8);
[~,g10] = lpc(I,10);
[~,g12] = lpc(I,12);
[~,g14] = lpc(I,14);
p = [2,4,6,8,10,12,14];
figure(6)
subplot(1,3,1);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
title('$E_p$ v.s. inverse filter order p of /i/,
n=1024:1087','Interpreter','latex');

[~,g2] = lpc(U,2);
[~,g4] = lpc(U,4);
[~,g6] = lpc(U,6);
[~,g8] = lpc(U,8);
[~,g10] = lpc(U,10);
[~,g12] = lpc(U,12);
[~,g14] = lpc(U,14);
p = [2,4,6,8,10,12,14];
subplot(1,3,2);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
title('$E_p$ v.s. inverse filter order p of /u/,
n=1024:1087','Interpreter','latex');

[~,g2] = lpc(NG,2);
[~,g4] = lpc(NG,4);
[~,g6] = lpc(NG,6);
[~,g8] = lpc(NG,8);
[~,g10] = lpc(NG,10);
[~,g12] = lpc(NG,12);
[~,g14] = lpc(NG,14);
p = [2,4,6,8,10,12,14];
subplot(1,3,3);
plot(p,[g2,g4,g6,g8,g10,g12,g14]);
xlabel('order of inverse filter (p)');
ylabel('$E_p$','Interpreter','latex','FontSize',11);
```

```
title('$E_p$ v.s. inverse filter order p of /$\eta$/,
n=1024:1087','Interpreter','latex');

I14 = zeros(1,256);
[I14(1:15),~] = lpc(I,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(I14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(7)
subplot(1,3,1)
plot(Nf,L,[270,270],[-20,100],[2290,2290],[-
20,100],[3010,3010],[-20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(280,90,'F1:270Hz');
text(1000,80,'F2:2290Hz');
text(3100,80,'F3:3010Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$)) of /i/, n=1024:1087,
p=14','Interpreter','latex');

U14 = zeros(1,256);
[U14(1:15),~] = lpc(U,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(U14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(7)
subplot(1,3,2)
plot(Nf,L,[300,300],[-20,100],[870,870],[-20,100],[2240,2240],[-
20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(310,90,'F1:300Hz');
text(900,80,'F2:870Hz');
text(2250,80,'F3:2240Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$) of /u/, n=1024:1087,
p=14','Interpreter','latex');
```

```
NG14 = zeros(1,256);
[NG14(1:15),~] = lpc(NG,14);
N = 0:127;
Nf = N*10000/256;
A14 = fft(NG14,256);
L = zeros(1,128);
for i = 1:128
    L(i) = 0.5*mag2db(1/((abs(A14(i)))^2));
end
figure(7)
subplot(1,3,3)
plot(Nf,L,[200,200],[-20,100],[2460,2460],[-
20,100],[3680,3680],[-20,100]);
xlabel('frequency(Hz)');
ylabel('10log($\frac{1}{|\hat{A}(k)|^2}$)','Interpreter','latex'
,'FontSize',11);
text(210,90,'F1:200Hz');
text(1200,80,'F3:2460Hz');
text(3690,80,'F4:3680Hz');
title('10log($\frac{1}{|\hat{A}(k)|^2}$) of /$\eta$/,
n=1024:1087, p=14','Interpreter','latex');
```

1 HW8 **100 / 100**

- ✓ **- 0 pts** **Correct**
- **- 5 pts** Part B1: Plot(s) incorrect/ missing
- **- 5 pts** Part B2: Plot(s) incorrect/ missing
- **- 5 pts** Results not commented properly
- **- 5 pts** Frequency axis incorrect
- **- 10 pts** Late submission