THE HONG KONG POLYTECHNIC UNIVERSITY
Department of Electronic and Information Engineering

Title: Speech Signal Analysis

Objectives: To understand LP, cepstral and short-time Fourier analysis of speech.

Equipment:
1. A PC with Windows 95/98/NT/2000/XP
2. Speech analysis package V3.0 (spana.exe)
3. Sound card, speakers, and microphone

Procedures

1) Download "SpanaV3.zip" and all wavefiles from http://www.eie.polyu.edu.hk/~mwmak/Download.htm
2) Unzip the file "SpanaV3.zip" and copy all .wav files to the C:\temp\<your_name> directory of your local PC.
3) Double click the file "Spana.exe" in your local PC.
4) Click **File**=>**Open file**, then select a .wav file, then press **Open.**
5) A dialog box will pop-up. Don't modify the default setting. Click **OK**.
6) Click **Plot**=>**Speech Signal** to display the speech signal in the time domain.
7) Click **Plot**=>**Windowed Signal** and select a frame number (e.g., 12) to display windowed speech so that a voiced portion of the speech waveform is displayed. The frame numbers are displayed on the bottom of the speech signal. Note that you could also click on any positions of the speech signal to view the corresponding windowed speech and spectral envelope. A window will pop up automatically. You could also change the type of plot (prediction error, spectral envelop, and cepstrum) in the pop-up window by selecting **Option=>Configuration** in the menu bar.
8) Click **Plot**=>**Prediction Error** and select a frame number (e.g., 12) to display the prediction error of the selected frame.

Question 1: Identify the similarity between the speech signal and the error signal? Why does such similarity exist?

9) Repeat Steps 6 and 7 with an unvoiced frame.

Question 2: What are the differences between the prediction error of voiced frames and that of unvoiced frames?

10) Click **Plot**=>**Spectral Envelope** and input a frame number to display the spectral envelope.

Question 3: Describe a method to obtain the spectral envelope from the LP coefficients. What is the relationship between the spectral envelope and the spectrum of the windowed speech signal? Why?

11) Click **Plot**=>**Cepstral Plot** and input a frame number to display a cepstral plot.

Question 4: Explain your observation. How to determine whether the signal inside the window is voiced or unvoiced using the cepstral plot? How to determine the pitch period for voiced sound from the plot?

Question 5: Describe the procedure of producing a smoothed cepstral envelope from the cepstral plot obtained in Step 10. What would be the differences between the spectral envelope obtained by cepstral smoothing (The DFT of cepstrum after low-pass liftering) and that obtained from LP coefficients?

12) Click **Plot**=>**Power Spectrum of Windowed data.**
13) Click **Option**=>**Configuration** and change the **Order of Prediction** setting to 18. Display the spectral envelope by clicking **Plot**=>**Spectral Envelope**.

Question 6: What is the difference between the spectral envelopes obtained in Step 10 and Step 13? Why does such difference exist? What would you obtain when the prediction order becomes infinitely large? Hence, explain how to choose the prediction order.

14) Run Matlab and input the following matlab code

```
clear all;
close all;
t = 1:0.01:4*pi;
N = 512;
t = t(1:N);
A1 = 1;
A2 = 0.5;
delta = 10;

xr  = A1*cos(1000*t)+A2*cos((1000+delta)*t);
xh = hamming(N).*xr';
xk = kaiser(N,6).*xr';
plot(xr);
figure;
plot(xh);
Xr = abs(fft(xr,N));
```

```
Xh = abs(fft(xh,N));
Xk = abs(fft(xk,N));

figure;
semilogy(Xr(1:N/2),'b');
hold on;
semilogy(Xh(1:N/2),'r');
%semilogy(Xk(1:N/2),'k');

figure;
wh = hamming(50);
wr = ones(50,1);
[Wh,f] = freqz(wh/sum(wh),1,N,2);
[Wr,f] = freqz(wr/sum(wr),1,N,2);

plot(f,20*log10(abs(Wh)),'r');grid on;
hold on;
plot(f,20*log10(abs(Wr)),'b');
legend('Hamming','Rectangular');
```

Progressively reduce the value of `delta` and observe the effect on windowing.

Restore the change made in Step 13. Click **Plot**=>**Spectral Envelope** and input a frame number. Click **Tools**=>**Windowing Type Control** and select **non-overlapping display** mode. Select different window types to compare their spectral envelope. Select **overlapping display** mode to observe the overlapped spectral envelopes.

Question 7: Based on the above Matlab code and Spana, compare the effect of the Hamming window on the speech spectrum with that of the rectangular window.

15) Click **Plot**=>**Spectral Envelope** and input a frame number. Click **Tools**=>**Pole Control** and drag the poles on the dialog box. These are the poles of the LP transfer function.

Question 8: What is the relationship between the pole positions on the z-plane and the peak positions of the spectral envelope? Why does such relationship exist? What will happen if all the poles are on the unit circle? You may find that some of the poles are on the real axis. Why do such poles exist? You may also find that the number of peaks in the spectral envelope is less than the number of pole-pairs. Why is it the case?

16) Click **Plot**=>**Spectral Envelope** and input a frame number. Click **Tools**=>**LP Control** and select different LP and PARCOR (reflection coefficients) parameters. Change their values by moving the slider. Observe the changes in the spectral envelope.

17) Click **Audio**=>**Record** to start recording for about 10 seconds and you should speak to the microphone some voiced and unvoiced sounds. Click **Audio**=>**Stop** to stop

recording. You could select different display modes by clicking **Audio=>Display**. Note that this is a direct-to-disk function. Therefore, do not use this function to record sound indefinitely; otherwise, it will not only blow up your hard disk but also take a long time to process the recorded speech. Click **Plot**=>**Cepstral plot** and select different frame numbers to observe the results.

Question 9: How do you determine whether a speech frame is voiced or unvoiced?

- END -