New Mexico State University
Klipsch School of Electrical Engineering

EE589 Digital Speech Processing
Spring 2011 – Project #2
Due: 8.55am Thu. Mar. 3

Name: ________________________

Grade: ________________________

Project

The goal of this project is to utilize signal processing software to analyze speech signals using autocorrelation, linear prediction, and cepstrum. Some of the software will be developed by students in MATLAB based on the *DSP Software Toolkit* which can be found at

`http://www.ece.nmsu.edu/~pdeleon/Teaching/EE589/DSP_Toolkit.pdf`

In addition, we will utilize several popular MATLAB toolboxes such as the Signal Processing Toolbox and Statistics Toolbox as well as third-party toolboxes such as Voicebox

`http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html`

You are free (unless otherwise noted) to use other toolboxes or develop your own functions.

Report

Please submit a printed report with your results including commentary and plots. Please `mailto: pdeleon@nmsu.edu` a zip file containing your report and all MATLAB code to recreate your results.

Notes

Unless otherwise noted, all speech signals should be from the student's own recordings. Students are encouraged to discuss detailed, technical aspects with each other and Prof. De Leon. Students are encouraged to utilize existing MATLAB functions in this project. However, students must write all other required codes on an *individual* basis.

Some problems ask for comments. As is always the case, your analysis, comparison to expected theory or model, and resulting comments are far more important than the code or plot. Therefore, please pay special attention to your comments.

1 Software Tools

Please write the following function from the *DSP Software Toolkit*: **ar_synthesizer.m**, **correlation.m**. Although not required, you may wish to also implement **levinson2.m**. Be sure to use the appropriate function header described on p. 5 and available at

`http://www.ece.nmsu.edu/~pdeleon/Teaching/EE589/funchead.m`

In addition, you may wish to use tools from the MATLAB Signal Processing Toolbox and the Voicebox Toolbox in this project.

2 Problems

Autocorrelation Analysis of Speech

1. Plot the first 10 ms (160 lags) of autocorrelations from your entire SPEECH.WAV. You should remove any DC component from your signal and normalize for unit variance, i.e. $r[0] = 1$ before you compute correlations:

   ```
   x = x  mean(x);  % zero mean
   x = x ./ sqrt(cov(x));  % scale for unit variance
   ```

   Compare your autocorrelation(s) to other students' autocorrelations. Comment on the general trend amongst the autocorrelation sequences noting the duration of initial strong correlation.

2. Redo Example 5.2 in Quatieri, i.e. compute the autocorrelation sequence using your recordings of (i) vowel /aa/, (ii) unvoiced stop /k/, (iii) unvoiced fricative /f/, and (iv) voiced stop /g/. Since the autocorrelation is symmetric, you need only plot the positive lags. Comment on whether the autocorrelations are noise-like and/or periodic-like.

3. Textbook problem 5.2.

Linear Prediction Analysis of Speech

4. Filter white noise (1 s @ $f_s = 16$ kHz) through an IIR filter with peaks at 500, 1000, 2000, 4000 Hz as in Lecture 8 (Demo LPA), i.e. synthesize an AR process. Compute the power spectrum and LPCs for model orders $p = 4$ (undermodeled), $p = 8$ (critically modeled), and $p = 18$ (overmodeled).

   Plot the power spectrum of the process with the magnitude response of the model (defined by the LPCs) overlayed as in Quatieri Fig. 5.13 or the last plot in Lecture 8 (Demo LPA). Comment on what you observe.

5. Compute the power spectrum and LPCs (you decide the model order p) for the following phonemes: /aa/, /iy/, and /v/. For each phoneme, plot the power spectrum of the phoneme with the magnitude response of the model (defined by the LPCs) overlayed.

6. Redo Figure 5.11 using the equivalents from your recorded phonemes. Be sure to indicate which vowel, unvoiced plosive, and unvoiced fricative you use. Comment on your results after reading p. 212.

Cepstral Analysis of Speech

7. Create a 256-sample periodic pulse train with a fundamental $f_0 = 123$ Hz (assume $f_s = 8000$ Hz). Plot the pulse train waveform, log-magnitude spectrum, and cepstrum. Plot units should be t (s), f (Hz), or t (quefrency). Create a transfer function $H(z)$ with poles at 768, 1333, and 2522 Hz (assume 0.95 pole magnitude). Plot the log-magnitude response and cepstrum of $H(z)$. Add the two cepstra together to simulate the cepstrum of the phoneme /a/ for the average male. Comment.

8. Plot the real cepstrum for the following phonemes: /aa/, /iy/, and /v/. Comment on the vocal tract information and pitch information you see in the plots.

9. Compute the sequence of mel-frequency cepstra for the following phonemes: /aa/, /iy/, and /v/ using a 256-point Hamming window with 50% overlap. Use imagesc to visualize the sequence of mel-frequency cepstra where the frame number is on the horizontal axis and mel-frequency cepstral coefficient is on the vertical axis.