



New Mexico State University
Klipsch School of Electrical Engineering

EE589 Digital Speech Processing
Spring 2011 – Project #1
Due: 8.55am Thu. Feb. 17

Name: _____

Grade: _____

Project

The goal of this project is to utilize signal processing software to analyze speech signals in different ways. 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*: **plotcsig2.m**, **periorogram_plot.m**, **olaistft.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>

The following tools from the MATLAB Signal Processing Toolbox may also be used: `spectrogram.m`, `periodogram.m`, and `pwelch.m`. Keep in mind that for a real signal, `pwelch.m` returns the one-sided PSD but `periodogram.m` assumes the input PSD is two-sided. You may need to make small accommodations to use `pwelch.m` in conjunction with `periorogram_plot.m`.

The following tools from the Voicebox Toolbox may also be used: `enframe.m`, `overlapadd.m`, and `spgrambw.m`.

You should read the help file associated with each of these functions.

2 Problems

Speech Signal Analysis

1. Plot the waveform from NAME.WAV and hand-label the individual phonemes (phonetic labeling). Assume the 40 or so American English phonemes can each be represented by a unique 6 bit codeword, determine the information rate in bits/s of your utterance.
2. Plot the waveform and Power Spectral Density (PSD) using Welch's method for the first 10–20 s of SPEECH.WAV.

Signal Analysis of Phonemes

3. Plot the waveform, PSD (pwelch), and spectrogram of the **vowels** IY (front) and AA (mid). Use a 256-point Hamming window with 50% overlap. Estimate the fundamental frequency, F_0 . Compare to the information in the text and lecture notes and please give your comments.
4. Plot the waveform, PSD (pwelch), and spectrogram of the **nasals** M and N. Note the “hole” or “dip” in the spectrogram (if any). Compare to the information in the text and lecture notes and please give your comments.
5. Plot the waveform, PSD (pwelch), and spectrogram of the **unvoiced fricative** F. Compare to the information in the text and lecture notes and please give your comments.
6. Plot the waveform, PSD (pwelch), and spectrogram of the **voiced fricative** V. Compare to the information in the text and lecture notes and please give your comments.
7. Plot the waveform, PSD (pwelch), and spectrogram of the **voiced plosive** (stop) G. Note the “voice bar” in the spectrogram (if any). Compare to the information in the text and lecture notes and please give your comments.
8. Plot the waveform, PSD (pwelch), and spectrogram of the **unvoiced plosive** (stop) K. Compare to the information in the text and lecture notes and please give your comments.
9. Plot the waveform and spectrogram of the **diphthong** OY. Note the formant transition in the spectrum. Compare to the information in the text and lecture notes and please give your comments.
10. Plot the waveform and spectrogram of the **semivowels** Y (glide) and L (liquid). Compare to the information in the text and lecture notes and please give your comments.
11. From PSD data (no need to print PSDs), create a table with estimates of the first three formant frequencies F_1 , F_2 , and F_3 of the vowels IY, IH, EH, AE (front); AA, ER, AH, AO (mid); UW, UH (back). On a single figure, plot F_1 vs. F_2 for each vowel (data points). Sketch the “vowel triangle”, i.e. triangle which encompasses data points. You may wish to use typical or average formant frequency values (available in the text or other online resource) to help you estimate your values.

STFT and Inversion of STFT

12. (Direct Inversion) Compute the STFT (Hamming window with 50% overlap) of NAME.WAV and then directly invert this STFT to recover the waveform. Compute the absolute difference between the before- and after- waveforms

$$\sum_{n=0}^M |x_{\text{before}}[n] - x_{\text{after}}[n]|. \quad (1)$$

Please use the following steps: 1) compute the inverse DFT for each windowed frame, 2) divide out the window values to obtain the frame samples, and 3) reconstruct the signal by choosing each sample from either of the two recovered frames. You may write your own tools to do this (including the `stft.m`) or you may wish to piece available tools together to accomplish this task.

13. (Overlap and Add ISTFT) Repeat the above problem but use the `olaistft.m` tool. Show that your code also works with 25% and 75% overlap.