EEL 6586: HW#3

Assignment is due Monday, Feb 20, 2004 in class. Late homework loses $e^{\# of \ days \ late} - 1$ percentage points. See the current late penalty at http://www.cnel.ufl.edu/hybrid/harris/latepoints.html

PART A: Textbook problems (Use Matlab only to optionally check your work)

A1 Compute the autocorrelation function r[k], of the following discretetime signal:

$$x[n] = \sin(\omega n)$$

A2 A train of impulses is fed through a simple all-pole model to produce a stationary voiced phoneme. The resulting sequence is windowed by a rectangular window and its Fourier magnitude is plotted as shown below. The sampling rate is 10KHz. Answer the following questions.



- a) Estimate the frequency of the first two formants (F1 and F2) in Hz.
- b) Estimate the pitch of the voice in Hz.
- c) Estimate the size of the rectangular window (in ms) used in processing the speech. Explain your answer.

J.G. Harris

A3 Assume that an infinite impulse train

$$\sum_k \delta[n+kP]$$

is filtered by a vocal-tract model given by $H(z) = 1/(1 + .9z^{-1} + .81z^{-2})$ to produce a speech signal s[n]. For simplicity, assume that the excitation (the impulse train) is uncorrelated and answer the following:

- a) Derive the difference equation for s[n].
- b) Compute the autocorrelation function r[0] for the signal s[n].
- c) Compute the autocorrelation function r[1] for the signal s[n].
- d) Compute the single LPC coefficient (p = 1) for this system.
- e) How does this coefficient compare to the first coefficient when p = 2? Explain.
- A4 Assume that white noise excitation w[n] is filtered by an *all-pole* vocaltract model $H(z) = 1/(1 + .25z^{-2})$ to produce a speech signal s[n]. w[n] is defined:

$$E\{w[n]w[m]\} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

In this problem you will use LPC to derive an all-pole approximation to H(z).

- a) Derive the difference equation for s[n].
- b) Compute the autocorrelation function r[0] for the speech signal s[n].
- c) Compute the autocorrelation function r[1] and r[2] for the speech signal s[n].
- d) Compute the first two LPC coefficients (p = 2).
- e) Derive $\hat{H}(z)$, the all-pole approximation to H(z). Does your answer make sense?
- A5 (for extra credit) Prove that the 3db bandwidth of a formant caused by a single dominant pole can be approximated by

$$bw \approx -\ln(r)f_s/\pi$$

where r is the distance of the pole to the origin and f_s is the sampling frequency in Hz.

J.G. Harris

PART B: Short Answer (No more than a few sentences each)

- B1 Suppose that r(1) = r(2) = r(3) = 0 for some stochastic signal. Does this necessarily imply that r(4) = 0? Explain.
- B2 Explain what happens to your speech when you breath helium into your vocal tract and try to speak.
- B3 A common algorithm for pitch determination is to perform autocorrelation on the LPC residual (error) and look for peaks. Why does this algorithm work better than performing autocorrelation on the original speech signal and looking for peaks?
- B4 Explain why humans have no problem determining the pitch of voices through the telephone, even though the cutoff frequency (about 300Hz) is larger than typical pitch frequencies.
- B5 A single word was recorded and shown below in a time domain plot. Identify the word in the list of the following. Explain your reasoning.
 - a) fail
 - b) test
 - c) obey
 - d) encyclopedia
 - e) none of the above



PART C: Computer Analysis of Speech In this part you will write a program for automatic formant analysis. You will run your program on the recorded sentence at

http://www.cnel.ufl.edu/hybrid/courses/EEL6586/sentence.wav Through the following steps, you will develop a formant detection algorithm that processes each window of speech using LPC analysis:

- C1 Break the sentence into overlapping windows. Describe how you choose the window type, length and overlap for this pitch estimation algorithm. Plot one of the windowed signals in the time domain. Also plot the power spectrum of the same windowed signal.
- C2 Write a routine to compute the LPC coefficients for each window. Show a typical example of the windowed signal (the same one in [C1]) along with the superimposed smooth envelope derived from the computed LPC coefficients. Use the same LPC order for all of Part C.
- C3 Write a matlab routine that returns the first two formants of speech given the LPC coefficients. Describe the strategy you used to write this algorithm. Show a graphical example of a smooth spectral envelope from [C2] with the two computed formant frequencies indicated. How did you choose the best order for the LPC computation.
- C4 Put all of the pieces together and write an algorithm to compute the first two formants for each window (assuming the formants exists) of a full sentence. Turn in a final plot of F1 and F2 (in Hz) vs. time (in seconds) for the given sentence. You may need to add an additional filtering step to smooth out the values. Indicate unvoiced regions and silence with formants of zero value.
- C5 Indicate at least three different vowel phonemes on the plot and indicate the numerical values of their formant frequencies. Argue that these formant frequencies are reasonable given the table of standard phonemes in the book.

As always, hand in all of your matlab code as the appendix of your homework. Discuss your algorithms in detail and comment on the accuracy of your algorithms.