ECE 769 – Special Topics in Signal Processing: Speech and Audio Processing

Assignment #1: All-pole analysis and synthesis of speech

The Due date: Report and MATLAB code to be submitted (preferably via email) before class on Thursday March 14.

The Assignment: Complete Exercises 5.23 and 5.24 at the end of Quatieri Chapter 5.
These exercises are reproduced on the back of this sheet for those who don't have a copy of the text book.
I will make the MATLAB speech files from the textbook's web site also available for downloading from the ECE 769 course web site.

The report: Should include plots and descriptions of results and answers to specific questions.

The code: The MATLAB M-files generating the results in your report should be submitted.

The rules: Each student should work on their own code and report. If you get stuck, you can ask for advice from another student. If you get really stuck, you can ask me.

Dr. Ian C. Bruce February 27, 2003

5.23 (MATLAB) In this exercise, use the voiced speech signal *speech1_10k* (at 10000 samples/s) in the workspace *ex5M1.mat* located in companion website directory Chap_exercises/chapter5. This problem illustrates the autocorrelation method of linear prediction.

(a) Window *speech1_10k* with a 25-ms Hamming window. Compute the autocorrelation of the resulting windowed signal and plot.

(b) Assume that two resonances represent the signal and model the vocal tract with 4 poles. Set up the autocorrelation matrix R_n, using your result from part (a). The autocorrelation matrix is of dimension 4×4.

(c) Solve for the linear predictor coefficients by matrix inversion.

(d) Plot the log-magnitude of the resulting frequency response:

$$H(\omega) = \frac{A}{1 - \sum_{k=1}^{p} \alpha_k e^{-j\omega k}}$$

where the gain is given by Equation (5.30). Compare your result with the log-magnitude of the Fourier transform of the windowed signal. What similarities and differences do you observe?

(e) Using your estimates of the predictor coefficients from part (c), compute the prediction error sequence associated with *speech1_10k* and plot. From the prediction error sequence, what conclusions might one draw about the model (i.e., all-pole/impulse-train-driven) and estimation accuracy?

5.24 (MATLAB) In this problem you will use your results from Exercise 5.23 to perform speech synthesis of the speech waveform *speech1_10k* in the workspace *ex5M1.mat* located in companion website directory Chap_exercises/chapter5.

(a) Using your estimates of the predictor coefficients from Exercise 5.23, compute an estimate of the vocal tract impulse response.

(b) Using the prediction error sequence you computed in Exercise 5.23, estimate an average pitch period of the voiced sequence *speech1_10k*.

(c) Using your results from parts (a) and (b), synthesize an estimate of *speech1_10k*. How does your waveform estimate differ from the original? Consider the minimum-phase nature of the impulse response estimate.

(d) Using the MATLAB number generator *randm.m*, synthesize the "whispered" counterpart to your voiced synthesis of part (c). Using the MATLAB *sound.m* function, listen to your two estimates and compare to the original.