Homework 3 (ECE6255 Spring 2010) Grade=4/100

1. Define the following three functions for a random or periodic signal:

Long-time autorrelation function: $\phi(k) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} x(m)x(m+k)$

Short-time autorrelation function: $R_n(k) = \sum_{m=0}^{N-|k|-1} x(n+m)w'(m)x(n+m+k)w'(m+k)$

Modified short-time autorrelation function: $\hat{R}_n(k) = \sum_{m=0}^{N-1} x(n+m)x(n+m+k)$

Show whether or not the following statements are true or false:

(a) If $x(n) = x(n+P),\ -\infty < n < \infty,$ then

(i) $\phi(k) = \phi(k+P),\ -\infty < k < \infty;$

(ii) $R_n(k) = R_n(k+P),\ -(N-1) < k < (N-1);$

(iii) $\hat{R}_n(k) = \hat{R}_n(k+P),\ -(N-1) < k < (N-1);$

(b)

(i) $\phi(-k) = \phi(k),\ -\infty < k < \infty;$

(ii) $R_n(-k) = R_n(k),\ -(N-1) < k < (N-1);$

(iii) $\hat{R}_n(-k) = \hat{R}_n(k),\ -(N-1) < k < (N-1);$

(c)

(i) $\phi(k) \le \phi(0),\ -\infty < k < \infty;$

(ii) $R_n(k) \le R_n(0),\ -(N-1) < k < (N-1);$

(iii) $\hat{R}_n(k) \le \hat{R}_n(0),\ -(N-1) < k < (N-1);$

(d)

(i) $\phi(0)$ is equal to the power in the signal;

(ii) $R_n(0)$ is the short-time energy;

(iii) $\hat{R}_n(0)$ is the short-time energy.

2. Here is an exercise to test your knowledge about English:
   a) In American English, which phoneme can never occur in word initial position (hint: only one and it exists in other languages)?
   b) There are many word initial consonant clusters of length two, such as speak, drank, plead, press etc. For word initial consonant clusters of length three there are only six in American English, give one word example for each of the six three-

sound consonant clusters. What general rule describes the sounds in each of the three positions?

c) A nasal consonant can be combined with a stop consonant (e.g., camp, tend) in a limited number of ways. What general rule do such combinations obey? There are several notable exceptions to this general rule. Can you give a few exceptions? What kind of speaking irregularity often results from these exceptions?

3. The following MATLAB exercise gives you a bit of experience with processing speech. Using the speech file 1.wav and the corresponding phone transcription file, do the following (you can do the same with the other 7 utterances):
   a) Compute the average energy of each speech segment, corresponding to each phone label, and compare the energy level for the five major manner classes;
   b) Zero out all the vowel regions of the utterance (replace the samples with zero-valued samples in the vowel regions) and listen to the sentence;
   c) Zero out all the consonant regions of the sentence (replace the samples with zero-valued samples during in consonant regions) and listen to the sentence;
   d) Determine which of the two modified waveforms (i.e., missing the vowels or missing the consonants) is the most 'intelligible'. See if someone who had never heard the sentence is able to correctly identify all of the words from either version of the modified waveform.

4. Write a MATLAB program to analyze a speech file and simultaneously, on one page, plot the following measurements:

   [1] the entire speech waveform
   [2] the short-time energy, E_n
   [3] the short-time magnitude, M_n
   [4] the short-time zero-crossing, Z_n

   Use the speech waveforms in the files 1.wav and 5.wav to test your program. Choose appropriate window sizes, window shifts, and window for the analysis. Explain your choice of these parameters. (Don't forget to normalize your analysis depending on the sampling rate of the speech signal in each file).