EE 627 - Speech Signal Processing

Assignment # 6

1. a) Describe local path and global path constraints in word spotting using dynamic time warping (Hint: Refer chapter 11 of deller's book).

   b) Give a stepwise description of the (i) K-means algorithm and (ii) LBG algorithm, for delivering VQ codebooks in speech recognition.

2. a) Derive the expression for the maximum likelihood estimates for $\hat{\mu}$ and $\hat{\sigma}^2$ for a univariate Gaussian pdf and multivariate case.

   b) Read Appendix 14.A (page 752-754) of Thomas Quatieri book and list your thoughts.

   c) (MATLAB) In this problem, you investigate the time resolution properties of the mel-scale and subband filter output energy representations. You will use the speech signal *speech1_10k* in workspace *exl4MI.mat* and functions *make_mel_filters.m* and *make_sub_filters.m*.

   i. Argue that the subband filters, particularly for high frequencies, are capable of greater temporal resolution of speech energy fluctuations within auditory critical bands than are the mel-scale filters. Consider the ability of the energy functions to reflect speech transitions, periodicity, and short -time events such as plosives in different spectral regions. Assume that the analysis window duration used in the STFT of the mel-scale filter bank configuration is 20 ms and is about equal to the length of the filters in the low 1000-Hz band of the subband filter bank. What constrains the temporal resolution for each filter bank?

   ii. Write a MATLAB routine to compute the mel-scale filter and subband filter energies. In computing the mel-scale filter energies, use a 20-ms Hamming analysis window and the 24- component mel-scale filter bank from function *make_mel_filters.m*, assuming a 4000-Hz bandwidth. In computing the subband-filter energies, use complex zero-phase subband filters from function *make_sub_filters.m*. For each filter bank, plot different low- and high-frequency filter-bank energies in time for the voiced speech signal *speech1_10k* in workspace *ex14MI.mat*. For the subband filter bank, investigate different energy smoothing filters p[n] and discuss the resulting temporal resolution differences with respect to the mel-scale filter analysis.