Digital Speech Processing Projects

Algorithmic Projects:

1. Speech Detector to separate speech from background sounds, noise, etc.
   a. Based on log energy and zero crossing rate statistics
   b. Need training set for setting statistical thresholds, testing set to
determine algorithm performance

2. Voice/Unvoiced/Silence Detector
   a. Based on statistical distributions of speech parameters measured from
      a training set—e.g., published set of 5 parameters
   b. Need to select appropriate speech features and measure their
distributions for all three categories of signal
   c. Need to measure performance on an independent test set

3. Pitch Detection
   a. Parallel time domain processor described in class
   b. Autocorrelation (full band, 900 Hz filtered)
   c. Average Magnitude Difference Function
   d. Cepstrum
   e. Harmonic spectrum
      i. Need a training set of utterances with known pitch contours to
tune algorithm parameters
      ii. Need a testing set of utterances with known pitch contours to
test algorithm performance

4. Formant Estimation
   a. LPC polynomial root solving—issues of deciding which roots are valid
      formants
   b. Spectrum peak picking
   c. Homomorphic formant estimation
      i. Need a training set of utterances to tune algorithm parameters
      ii. Need a testing set of utterances to test algorithm performance

Speech Systems

1. Speech Coders
   a. ADPCM or higher complexity waveform coder
   b. Subband coder using 2 or more bands
   c. LPC coder using pitch or full excitation
   d. Homomorphic coder
   e. Multi-Pulse LPC (MPLPC) coder
   f. Code-Excited LP (CELP) coder
   g. Channel Coder
i. Need to consider quantization and sampling scheme for coder parameters

2. Speech Synthesizer
   a. Concatenated whole word synthesizer
   b. Number string synthesizer
   c. Units, features, prosody issues to be resolved

3. Speech Recognizer
   a. Isolated words, speaker trained system
   b. HMM or template representations
   c. Viterbi decoding or dynamic time warping (DTW) decoding
   d. Vocabulary for recognition
   e. Feature set for recognition
   f. Training set for building models
   g. Testing set for evaluating system performance

4. Speaker Verification System
   a. Based on number strings
   b. Training set of real customer strings, imposter strings
   c. Testing set of real customer strings, imposter strings
   d. Setting algorithmic thresholds

5. Audio Coder
   a. MP3 encoder/decoder
   b. AAC encoder/decoder
   c. Use of range of bit rates from 192 Kbps to 128 Kbps to 96 Kbps to 64 Kbps
   d. Measurement of audio quality