CCE5223 Speech Processing and Coding

Assignment 2013 A

Linear Prediction Coding – LPC Analysis and Synthesis using Multipulse Excitation

Build a Multipulse Excitation Based Linear Predictive Coding System based on speech sampled at 16 KHz.

Analysis

This requires the following components:

Section A – Getting the analysis LPC based parameters

(i) separation of the input to frames. Use frames of 20 ms with an overlap of 10 ms

(ii) for each frame calculate a value for energy (note tha this can be R0 in the autocorrelation sequence)

(iii) build a voiced/unvoiced frame decision based on energy and zero crossing

(iv) build a pitch estimation based on autocorrelation techniques

(v) use a hamming window and preemphasis, estimate for each frame a 10-bit LPC using the autocorrelation method and the Durbin-Levinson algorithm. But transform your basic alpha parameters to log area ratios.

(There are standard software routines for this transformation – see Orfanidis or web.)

So this should give for each frame a value for pitch (equals 0 for unvoiced),an energy value for R0 and ten LAR coefficients. Note (iii) and (iv) are not necessary if section B is done.

Section B – Use analysis by synthesis similar to Fig 12.23 of Quatieri to set up a set of pulses associated with a time frame of the LPC frame to reduce the error

(i) For current time frame subtract effect from past impulses (nothing for first frame)

(ii) find location and amplitude of one impulse that reduces the mean squared error for the frame

(iii) subtract the effect of considering this pulse by obtaining the synthesis and finding the new error between result from (i) and current synthesised output

(iv) repeat (ii) and (iii) obtaining 8 impulses per frame.

(v) Include perceptual weighting filter to reduce further e(n). (In Atal and Remde's original paper the parameter used is 0.8.)
(see Atal and Remde – ICASSP 1982)

Section C – Synthesis this is tied to Section B since

Hence build a synthesis system that considers one frame after another and their interrelationships as the effect of pulses run over into the next frame.

The duration of a frame, based on 16,000 samples per second rate and a duration of 20ms is (160 x 20) 320 time points. Each output frame must be overlapped with a previous frame by 10 ms , before obtaining the continuous output speech waveform.

References Deller Proakis and Hansen, Chapters 1, 4,5 and 7

Quatieri Chapter 12
Orfanidis (for progammes and algorithms)
B.S. Atal and J.R.Remde, "A new model of LPC excitation for producing natural sounding speech at low bit rates", ICASSP 1982

Use as speech source one of the files in the Annotation section of my web page eg go into FMM001.zip and then use fmm1m_2.wav.
You can use MATLAB or C or C++ for your software programmes.