Proceedings of The Institute of Acoustics

COMPARATIVE EVALUATION OF LP ANALYSIS METHODS FOR MULTI-PULSE EXCITED LP CODING OF SPEECH

T. Svendsen and K.K. Paliwal

Division of Telecommunications
University of Trondheim, Trondheim-NTH, Norway

Abstract:

Multi-pulse excited LP (linear prediction) coding has been recently suggested as an alternative technique of producing natural sounding speech at medium bit-rates. However, in this coding technique the speech quality may depend on the type of LP analysis method used. In the present paper, we have studied standard methods of LP analysis, namely, the autocorrelation method, the covariance method, the modified autocorrelation method [2] and the Burg method. In terms of segmental signal-to-noise ratio, we have found the autocorrelation method and Burg's method to be the best.

Simulations:

The multi-pulse coder used in the simulations is based upon the original proposal by Atal and Remde [1], its block diagram shown in fig. 1. The mode of operation was the following:

The LPC-analysis was performed over blocks of 20 ms of speech, and the analysis parameters were updated every 10 ms. The overlap percentage was 50% in both forward and backward direction and the order of the LPC-analysis was 10. The block-length for the multi-pulse error minimization was 10 ms and the search for optimum pulse amplitudes and positions were performed over the entire 10 ms block. Each block was searched for 8 pulses, yielding a pulse/sample ratio of 1/10. The weighting filter was of the form

$$w(z) = \frac{1 - A(z)}{1 - A(\gamma z)}$$

where $1 - A(z)$ is the LPC prediction filter and $\gamma = 0.8$ is a weighting coefficient.

The input speech was two English sentences, one spoken by a man, the other spoken by a woman. Total length of the test sequence was 5.3 s. The sampling frequency was 8kHz.

Simulation results for natural speech are shown in Table I

LPC-algorithm	SNR [dB]	Segmental SNR [dB]
Mod. covariance	10.69	10.44
Covariance	5.72	10.51
Burg	10.75	10.86
Autocorrelation	10.63	10.98

Table 1. SNR-performance for the multi-pulse coder for different LPC-algorithms. Natural speech.

The cause for the low SNR value for the covariance method is instability in the LPC-filters.

Simulations were also carried out for telephone (IRS) filtered speech, and the results are shown in Table 2.

LPC-algorithm	SNR [dB]	Segmental SNR [dB]
Mod. covariance	9.65	7.34
Covariance	8.46	7.28
Burg	9.68	7.38
Autocorrelation	9.79	7.64

Table 2. SNR-performance for the multi-pulse coder for different LPC-algorithms. Telephone speech.

The drop in SNR-values are due to the multi-pulse algorithm's ability to better trace low-frequency components containing high energy than high-frequency components.

We also simulated the Burg algorithm when performing a joint optimizatien of the pulse amplitudes after each new pulse was found [3]. For natural speech the simulation resulted in SNR = 11.28 dB, segmented SNR = 11.08 dB. The corresponding figures for telephone speech were 10.07 dB and 7.57 dB respectively.