EVALUATION OF LINEAR PREDICTION SCHEMESFOR USE IN MOBILE SATELLITE COMMUNICATIONS

J. Kostogiannis, A. Perkis

Department of Electrical and Computer Engineering
University of Wollongong

ABSTRACT - This paper evaluates and compares several analysis methods and quantization schemes applied to parts of a Linear Predictive coder, considering error free conditions and in the presence of random errors. Subjective and objective measures indicate that all the spectral analysis methods perform comparably, while the quantization schemes are shown to have a great impact on the degradation in speech quality at high bit error rates (BER). The sensitivity of the spectral information is reduced by the implementation of a "smart" filter stability correction algorithm, based on Line Spectrum Pairs (LSP).

INTRODUCTION

The economic and technical advantages of digital communications coupled by satellite systems' capacity to simultaneously link all users on a global and regional scale, appeal to the lucrative and rapidly expanding mobile communications industry. The difficulties facing a merger between the two , are the inherent high bit error rates (BER) on a satellite link and the need for low bit rate speech coding to accommodate more users on a link that has power limitations. A low bit rate speech coder, based on linear prediction, is a suitable candidate for such a system.

An integral feature of the coder and by far the most sensitive at high BER, is the spectral information represented by the linear prediction coefficients (LPC coefficients). High BER and to a lesser extent quantization distort the spectral information, causing irritable bangs and squeaks in the synthesized speech (Perkis & Ribbum 1990).

This paper presents several spectral analysis methods and quantization schemes and evaluates their performance considering error free conditions and in the presence of random errors.

The spectral analysis methods presented are ;

(i) Burg's algorithm (Rabiner & Schafer 1978).
(ii) The autocorrelation method, solved using Levinson's recursion (Markel & Gray 1978).
(iii) The covariance method, solved using the Gram-Schmidt orthogonalization procedure (Markel & Gray 1978).
(iv) A robust linear prediction algorithm based on minimizing the sum of appropriately weighted residuals (Lee 1988).

The LPC coefficients from the spectral analysis methods are quantized using the following schemes;

(i) Linear quantization of Log Area Ratios (LAR) (Perkis 1990).
(ii) Scalar quantization of Line Spectrum Pairs (LSP) (US DoD 1989).
(iii) Scalar quantization of Line spectrum Pair Differences (LSPD) (Soong & Juang 1984).

Performance evaluation of these schemes are carried out on a simple model based on an analysis filter cascaded with a synthesis filter. The test utterance used is by an Australian male of four seconds duration and sampled at 8kHz. The model parameters are the time varying LPC coefficients determined and quantized by the above mentioned schemes. The quantized LPC coefficients are used in the analysis filter, while the quantized LPC coefficients with random errors are used in the synthesis filter. To determine which is the most robust scheme segmental signal to noise ratio (SNRseg), log RMS spectral distortion and paired comparison tests are used (Schuyler 1988).

The use of Line Spectrum Pairs (LSP) to represent the spectral information introduces the possibility of checking for instability in the synthesis filter. This happens predominantly at high BER, which significantly affects the speech quality. The performance of a "smart" filter stability correction algorithm, which minimizes this degradation is presented.

Section 2 gives a brief outline of the losses in a satellite channel, while section 3 and 4 deal with the performance of the spectral analysis methods and quantization schemes respectively.

LOSSES IN A SATELLITE CHANNEL

A satellite link has an inherent high BER. This is due to transmission losses and noise, which characterize the channel. Transmission losses, due to the vast distances the transmitted signal has to travel, are a major problem considering the power available in a satellite system is limited. Absorption losses and noise are noticeable when there is rain, fog, cloud, snow and hail about. Another factor that can be linked to high BER is frequency interference from other terrestrial communication systems, especially in the 6/4GHz link.

Types of errors

The satellite channel is characterized by the occurrence of both random and burst errors. Burst errors can occur from lightning, man-made disturbances, shadowing of the line-of-sight and fading. The effects of burst errors on the spectral information will not be treated in this paper. A binary symmetric channel model is implemented into the simulation model for the introduction of random errors in the spectral information.

SPECTRAL ANALYSIS METHODS

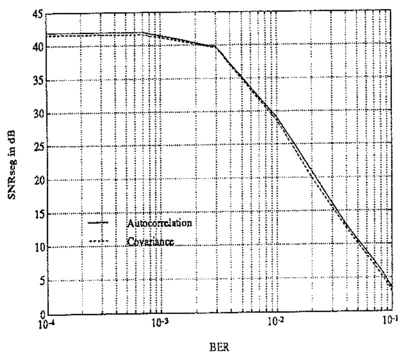
Spectral analysis methods address the problem of determining the LPC coefficients directly from the speech samples, to obtain an accurate estimate of the spectral envelope. The autocorrelation and covariance methods determine the LPC coefficients by minimizing the short term average prediction error, over the interval [0, N-1] and [-p, N-1] respectively,where p, the order of the prediction filter, is equal to 10. The sets of equations are solved using Levinson's recursion algorithm (which takes advantage of the toeplitz nature of the matrix) and by the Gram-Schmidt orthogonalization procedure respectively (Markel & Gray 1978).

Burg's algorithm determines the LPC coefficients by minimizing the forward and backward prediction errors simultaneously (Rabiner & Schafer 1978). The robust linear prediction algorithm as stated previously, minimizes the sum of appropriately weighted residuals. The weight is a function of the prediction residual (Lee 1988). Huber's loss function is selected to give more weight to the smaller residuals and less to the small portion of large residuals (the nature of excitations for voiced speech). It is reported that this formulation gives a more efficient and less biased estimate for the LPC coefficients. The system of equations require iterative methods and a preliminary estimate is necessary to determine the prediction residuals, which increases the complexity. Newton's algorithm and Cholesky's decomposition are implemented to solve these equations. The preliminary estimate is determined by the autocorrelation method.

Performance of the spectral analysis methods

All spectral analysis methods perform comparably for each quantization scheme for various BER ranging from 10^{-4} to 10^{-1}. To avoid repetition the objective measures for speech quality are illustrated in Fig. 1 & 2, for the LSP quantization scheme.

All the methods degrade in speech quality at the same rate (as the BER is increased). At a BER of 10^{-3}, there is no noticeable degradation in the speech. A log RMS spectral distortion of 1.6 dB at a BER of 10^{-2} indicates tolerable speech quality, but the emergence of bangs and squeaks in the speech are irritating to the ear. At a BER of 10^{-1}, the speech quality has degraded to an unacceptable level with bangs, squeaks and bird song like background noise predominant.



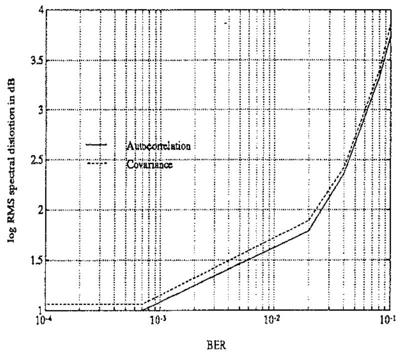


Figure 1. SNRseg as a function of BER for the autocorrelation and covariance methods respectively.

Figure 2. Log RMS spectral distortion as a function of BER for the autocorrelation and covariance method respectively.

The autocorrelation method attains the most favourable results, but not significant enough to distinguish a difference among the rest of the spectral analysis methods. This is reinforced with a paired comparison test. Here, pairs of utterances were presented to ten listeners asked to choose their preference in each pair. The pairs consisted of combinations among utterances using the spectral analysis methods at a BER of 10^{-1} and the original utterance. Each utterance was represented four times. The test indicates that the autocorrelation method is preferred among the spectral analysis methods, with an average score of 2.3, but listeners complained that in most cases there was no perceivable difference.

QUANTIZATION SHEMES

Direct quantization of the LPC coefficients is wasteful. Thus for transmission purposes at low bit rates, other forms representing the spectral information must be used . We will consider LAR (Perkis 1990) and LSP (Kang & Fransen 1979).

Quantization of log area ratios (LAR)

Quantization of the spectral information using reflection coefficients, k_i, provides a filter stability check, where stability is assured if $k_i < 1$. Due to the non-uniform spectral sensitivity of the reflection coefficients, especially near unity magnitude, linear quantization of the LAR are more suitable.

As no training data base of any statistical significance is available, the LAR quantizer was optimized using a male utterance sampled at 8kHz and for each analysis scheme using a 30 msec. frame length. A number of bit allocations were tested and the least spectral distortion is attained using a 5 5 3 3 3 3 3 3 2 2 bit allocation for a 32 bit quantizer.

Quantization of the line spectrum pairs and their differences (LSP & LSPD)

It has been reported that the LSP are capable of reducing the bit rate by 25% (compared to LAR) for quantizing the spectral information without any degradation to the speech. This can be attributed to the fact that an error in a LSP frequency affects the prediction filter's spectrum near that frequency and not the rest of the spectrum. The quantization scheme for the LSP, developed by Bells Labs and the US Department of Defence (US DoD 1989), has the following bit allocations for the ten LSP frequencies 3 4 4 4 4 3 3 3 3 3. The scheme has a bandwidth expansion of 10 Hz.

A more desirable transformation of the LSP frequencies that is less susceptible to speaker characteristics, as well as recording and analysis conditions, is the differences between adjacent frequencies (LSPD) (Soong & Juang 1983). For scalar quantization of the LSPD, optimum Max quantizers for (0, 1, 2, 3, 4, 5 bits) have been designed for each LSP frequency difference (Perkis 1990). The bit allocation for the scheme is 3 3 4 4 3 3 3 3 3 3.

Performance of the quantization schemes

All spectral analysis methods are applied to all the quantization schemes. To avoid repetition the schemes' performances are evaluated using the autocorrelation method. The quantizers' performances are deemed acceptable in error free conditions. Average log RMS distortion values of 0.665 dB, 1.01 dB and 0.778 dB are exhibited by the LAR, LSP and LSPD quantizers respectively. These measures are evaluated using a male utterance of four seconds duration. The objective measures in Figure 3 & 4 indicate that the LAR scheme performs best at high BER. This is supported by a paired comparison test of these quantization schemes using ten listeners.

As the BER is increased, the speech quality degrades for all the quantization schemes with the LSPD scheme the most affected. For the LAR scheme, bangs and squeaks can be heard, while in the LSP scheme, bird song like background noise is also predominant. In the LSPD scheme the bangs and squeaks are louder and there is evidence of high pitched resonances (due to the shifting of the speech spectrum to the right).

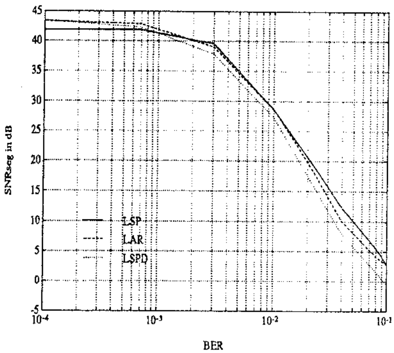


Figure 3. SNRseg as a function of BER for the quantization schemes using the autocorrelation method.

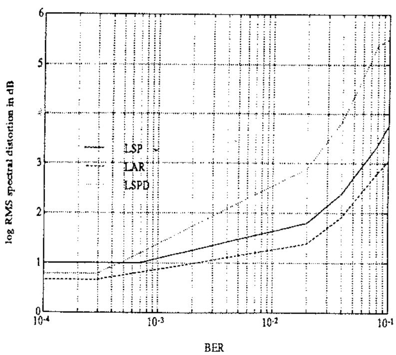


Figure 4. Log RMS spectral distortion as a function of BER for the quantization schemes using the autocorrelation method.

Correction of unstable filters

In the case of the LAR, there is no indication of unstable filters throughout all the BER. In the LSPD scheme any unstable filter is replaced by the previous stable one. In the LSP scheme, conserving filter stability after quantization and at high BER is possible by observing that the LSP frequencies are interleaved. At high BER, closely spaced LSP frequencies may change order, resulting in an unstable synthesis filter. A similar problem occurs even if the order is not changed. Frequencies may move closer together resulting in large spectral distortions in the synthesized speech.

A "smart" filter stability correction algorithm is proposed which involves sorting the LSP frequencies in ascending order and then checking for closeness. The close LSP frequencies are moved further apart and then sorted. Close LSP frequencies are moved apart by 40Hz for the first three LSP frequencies and 200hz for the rest. This takes advantage of the short term spectra of predominantly voiced speech where it is generally considered that the energy levels towards the higher frequency scale is relatively low.

The SNRseg measure indicates that it performs better than the original LSP correcting algorithm of just sorting the LSP frequencies in ascending order. Subjectively the irritable cracks' intensities are greatly reduced and some are totally eliminated. A paired comparison test using ten listeners indicated that this correction algorithm based on LSP is preferred to LAR. This "smart" filter stability algorithm is implemented to the LSPD scheme with favourable results.

CONCLUSION

In this paper we have compared four standard LPC analysis schemes and shown that they all perform comparably, both in error free conditions and in the presence of random errors. Both objective and subjective measures indicate that quantization schemes have a great impact on speech quality at high BER. Subjective listening indicated that filter stability is perceptually important. For LSP, a "smart" filter stability correction algorithm involving adaptive bandwidth expansion is presented. It minimizes the speech degradation and is subjectively preferred to both LAR and LSPD.

REFERENCES

Lee, C. (1988) ,*On robust linear prediction of speech* Trans. , ASSP, Vol 36, No. 5, May.

Kang, G. S and Fransen, L. J. (1984). *Application of Line-Spectrum Pairs to Low-Bit-Rate Speech Encoders*, Naval Research Laboratory Report.

Markel, D. H., Gray, A. H. (1976), *Linear Prediction of Speech* , (Springer-Verlag, Berlin.)

Perkis, A. (1990), Personal conversations.

Perkis, A., Ribbum, B. (1990) ,*Application of stochastic coding schemes in satellite communications* Advances in speech coding, (Kluwer Academic publishers , New York)

Proposed Federal Standard 1010, *Telecommunications: Analog to digital conversion of radio voice by 4800 bps Code Excited Linear Prediction Vocoder (CELP)*, First draft Sept. 1, 1989, U.S Department of Defence.

Rabbiner ,L. R. and Schafer ,R. W. (1978), *Digital processing of speech signals* , (Prentice Hall, New Jersey)

Soong, F. K. and Juang, B. H. (1984), *Line spectrum pairs (LSP) and speech data compression*, Proc. ICASSP

Schuyler, R. Q., Barnwell, P. T. and Clements M. A. (1988), *Objective measures for speech quality* , (Prentice Hall, New Jersey)