SIGNAL PROCESSING III: Theories and Applications
I.T. Young et al. (editors)
Elsevier Science Publishers B.V. (North-Holland)
© EURASIP, 1986

ROBUST LP ANALYSIS METHOD BASED ON PITCH INFORMATION FOR NOISY SPEECH

K.K. PALIWAL

Computer Systems and Communications Group
Tata Institute of Fundamental research
Homi Bhabha Road, Bombay 400005, India

In this paper, we propose a robust linear prediction (LP) analysis method for noisy speech which exploits the periodic nature of the voiced speech signal. This method first computes the pitch period P and then computes the M LP coefficients by matching the actual and the model autocorrelation coefficients from i=P+1 to P+q, where q is greater than M. This method is studied here for the voiced speech signals corrupted by the white Gaussian noise, though the method can, in principle, be applied to the coloured noise situations as well. It is found that this method performs better than the conventional autocorrelation method of LP analysis for noisy voiced speech.

1. INTRODUCTION

In linear prediction (LP) analysis, the speech signal is assumed to be the output of an Mth order all-pole filter (where M is about 8-10 for 8 kHz sampling rate). In the autocorrelation method of LP analysis, the parameters of the all-pole filter or the LP coefficients are estimated by matching the first M actual autocorrelation coefficients (computed from the speech signal) with the corresponding model autocorrelation coefficients [1]. When the speech signal is relatively noise-free, the matching of the first M autocorrelation coefficients also matches the remaining higher autocorrelation coefficients fairly well. Because of this, the autocorrelation method performs satisfactorily for the noise-free speech signals.

When the speech signal is corrupted by the additive white noise, performance of the autocorrelation method deteriorates severely [2,3]. This is because the addition of white noise changes the value of the zeroth autocorrelation coefficient. Since the matching of the first M autocorrelation coefficients makes use of the zeroth autocorrelation coefficient, this does not result in a proper matching of the higher autocorrelation coefficients. For the noisy signals having signal-to-noise ratios (SNRs) less than 20 dB, this fit between higher actual and model autocorrelation coefficients is not good and, hence, the autocorrelation method results in poor performance for the noisy signals. This poor performance is manifested in the power spectrum in the form of broadening of formant peaks and disappearance of some of the higher formant peaks.

Some LP analysis methods have been reported in the literature [4-8] which employ the matching of higher autocorrelation coefficients and, thus, avoid the use of the zeroth autocorrealtion coefficient. These methods perform better than the autocorrelation method. But, their performance is not very satisfactory because of the following two reasons: Firstly, the higher autocorrelation coefficients do not contain as much information about the all-pole system as the low autocorrelation coefficients [9]. Secondly, estimation of higher autocorrelation coefficients is statistically less reliable [6].

Recently, two methods have been reported in the literature [10,11] which employ matching of the low as well as the high autocorrelation coefficients for computing the AR parameters. However, before matching, these methods compensate (either explicitly or implicitly) the zeroth autocorrelation coefficient for the additive white noise.

In the present paper, we propose an alternate method for LP analysis of noisy speech which also uses long correlation matching (i.e., matching of both low and high autocorrelation coefficients) for computing LP parameters. This method does not compensate zeroth autocorrelation coefficient for the additive white noise. Instead, it exploits the periodic nature of the voiced speech signal whose autocorrelation function is also periodic. This can be seen from Fig. 1 where we show the autocorrelation function of the clean and corrupted speech signals. (Here, SNR=0 dB, number of samples N=200 and sampling rate=8 kHz.) The speech signal here corresponds to vowel /a/ having a pitch period P=19 samples. It may be noted here that the autocorrelation functions of the clean and corrupted speech signals are almost similar except for the zeroth autocorrelation coefficient.

The AR spectral estimation methods mentioned earlier [4-8,10,11] employ matching of the autocorrelation coefficients from the first period of the autocorrelation function. We have already discussed the problems associated with these methods. The robust LP analysis method proposed in this paper uses the matching of low as well as high autocorrelation coefficients from the the second period of the autocorrelation function. Since this does not involve the use of the zeroth autocorrelation coefficient, the AR parameters estimated by this method are not affected by the additive white noise. In addition, the present method has the advantage that it can be used for the coloured noise situations as well because the autocorrelation coefficients of the coloured noise, due to their exponential decay, are expected to be negligibly small during the second period.

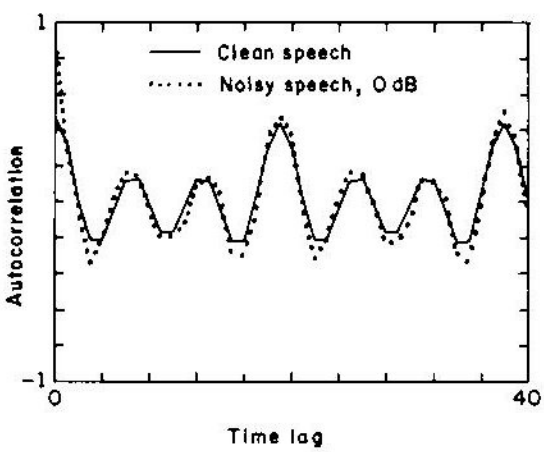

Fig. 1. Autocorrelation functions of the clean and the noisy speech signals. Pitch period P=19 samples.

The organization of this paper is as follows. We describe in Section 2 the robust LP analysis method. Section 3 presents the results and the conclusions are described in Section 4.

2. THE ROBUST LP ANALYSIS METHOD

In this section, we describe the robust LP analysis method proposed in the present paper for noisy speech. Let the observed (corrupted) speech signal be

$$y_n = x_n + w_n, \quad n=1,2,---,N,$$

where $\{x_n\}$ is the uncontaiminated (clean) speech signal and $\{w_n\}$ the noise signal. The aim here is to estimate the parameters $\{a_i\}$ of the Mth order all-pole system from the corrupted signal $\{y_n\}$. This is done here in the following steps:

Step 1. Compute the pitch period P from the corrupted signal $\{y_n\}$ using any of the pitch estimation method reported in the literature [12].

Step 2. Compute the autocorrelation coefficients $\{R(i)\}$ from the corrupted signal as follows:

$$R(i) = \frac{1}{N-i} \sum_{n=1}^{N-i} y_n y_{n+i} ,$$

for $i=P-M+1,---,P+q$. Here, $q>M$.

Step 3. Use a least-squares procedure to solve the following overdetermined set of q Yule-Walker equations for the M LP coefficients:

$$\sum_{i=1}^{M} a_i R(j-i) = -R(i), \quad j=P+1,---,P+q.$$

3. RESULTS AND DISCUSSION

In this section, we study the robust LP analysis method proposed in the present paper for the voiced speech signal corrupted by the white Gaussian noise. In order to put this method in a proper perspective, we compare here its performance with that of the autocorrelation method of LP analysis.

As an illustration, we select here a speech segment of vowel /a/ from the male speech digitized at 8 kHz sampling rate. The duration of this segment is 25 ms and the speech signal here has pitch period=61 samples. We add the simulated white Gaussian noise to it to make its SNR equal to 10 dB. The corrupted speech signal is analyzed using the autocorrelation method and the present method. The pitch period is estimated here from the corrupted signal using the autocorrelation method of pitch estimation [13]. We show in Fig. 2 the power spectral estimates from the autocorrelation method and the present method. This figure also shows the spectrum of clean speech which act as a reference. We can see from this figure that the formant peaks in the spectrum resulting from the autocorrelation method are broader than those in the actual spectrum. The present method rectifies this problem and results in better estimate of the power spectrum.

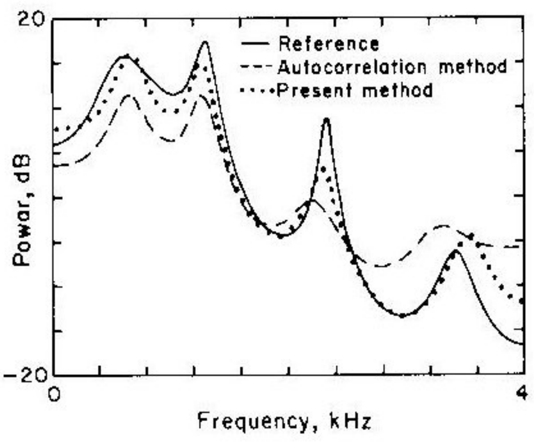

Fig. 2. Estimated power spectra of vowel /a/. (N=200, SNR=10 dB, M=10, q=60 and P=61 samples.)

Thus, we have seen that the present method results in better performance for the noisy voiced speech than the autocorrelation method of LP analysis. We have also noted earlier that this method can also be used for speech signals corrupted by the coloured noise, though we have not explored these signals in the present study. At this point, we want to note the two limitations of the present method. Firstly, this method requires pitch estimation from the corrupted signal. Since estimation of pitch is a complex problem even for the relatively less noisy speech signals [12], this method can not be used for very low SNR values. Secondly, this method uses the matching of the higher autocorrelation coefficients (from i=P+1 to P+q). Since estimation of the higher autocorrelation coefficients from the corrupted signal is statistically less reliable [6], it might result in higher variance spectral estimates.

4. CONCLUSIONS

In this paper, we have presentd a robust LP analysis method for noisy speech. This method exploits the periodicity of the voiced speech signal and employes the long correlation matching for the estimation of the LP coefficients. This method is studied here for the voiced speech signal corrupted by the additive white Gaussian noise. Its performance has been found to be better than the autocorrelation method of LP analysis. This method has an additional advantage over the earlier robust methods reported in the literature [4-8,10,11] that it can, in principle, be used for speech corrupted by the coloured noise. This aspect of the method will be studied in future.

REFERENCES

[1] J. Makhoul, "Linear prediction: A tutorial review", Proc. IEEE, Vol. 63, Apr. 1975, pp. 561-580.

[2] M.R. Sambur and N.S. Jayant, "LPC analysis/synthesis from speech inputs containing quantizing noise or additive white noise", IEEE Trans. ASSP-24, Dec. 1976, pp. 488-494.

[3] J. Tierney, "A study of LPC analysis of speech in additive noise", IEEE Trans. ASSP-28, Aug. 1980, pp. 389-397.

[4] W. Gersch, "Estimation of the autoregressive parameters of a mixed autoregressive moving-average time series", IEEE Trans. AC-15, Oct. 1970, pp. 583-588.

[5] Y.T. Chan and R.P. Langford, "Spectral estimation via the high-order Yule-Walker equations", IEEE Trans. ASSP-30, Oct. 1982, pp. 689-698.

[6] H. Ogura and Y. Yoshida, "Spectral analysis and subtraction of noise in radar signals", IEEE Trans. AES-17, Jan. 1981, pp. 67-71.

[7] J.A. Cadzow, "Spectral estimation: An overdetermined rational model equation approach", Proc. IEEE, Vol. 70, Sept. 1982, pp. 907-939.

[8] K.K. Paliwal, "A constrained forward-backward correlation prediction method for AR spectral estimation of noisy signals", this volume.

[9] S.P. Bruzzone and M. Kaveh, "Statistical efficiency of the sample autocorrelation function in ARMA parameter estimation", Proc. ICASSP, Paris, May 1982, pp. 240-243.

[10] V.K. Jain and B.S. Atal, "Robust LPC analysis of speech by extended correlation matching", Proc. ICASSP, Tampa, Mar. 1985, pp. 473-476.

[11] K.K. Paliwal, "A noise-compensated long correlation matching method for AR spectral estimation of noisy signals", Proc. ICASSP, Tokyo, Apr. 1986, pp. 27.7.1-27.7.4.

[12] W.J. Hess, Pitch Determination of Speech Signals, Springer-Verlag, Berlin, 1983.

[13] K.K. Paliwal, "Effect of spectral flattening on the pitch estimation performance of the autocorrelation method for noisy speech", Acoustics Letters, Vol. 7, Nov. 1983, pp. 74-76.