SPEECH MODELING BY CORRELATION FITTING

Jianguo Huang

Northwestern Polytechnical University

Xian, P.R. China

Leland B. Jackson

University of Rhode Island

Kingston, RI 02881, U.S.A.

Abstract--We apply a previously proposed technique for AR modeling by fitting windowed correlation data to speech modeling in additive white noise. Using the correlation fitting (CF) method, the estimated speech autocorrelation function can be approximated over many more than the minimum number of correlation lags to produce a much more accurate fit to the corresponding power density spectrum, reducing the effect of the noise. Simulation shows the improved results.

Keywords: AR modeling of speech, Correlation fitting mehtod, Linear prediction coding, Yule-Walker equations.

1. Introduction

The method for autoregressive (AR) and autoregressive-moving-average (ARMA) modeling of stationary stochastic signal have previously been proposed based upon fitting the model autocorrelation function to the estimated (and biased) autocorrelation in the least-square sense over many more than the minimum number of autocorrelation value (1,2). Here we apply the correlation fitting (CF) method to the case of speech modeling. For speech analysis the method of linear prediction coding (LPC) is one of the most powerful technique (3). The CF method has the same results as LPC when only the minimum number of autocorrelation lags are used.

But in the case of speech contaminated by white noise the LPC analysis technique degrades rapidly. For high signal-noise-ratio (SNR), greater than 30db, the additive noise only covers very small rigions of the spectrum as well as autocorrelation. Speech signal still is approximately a pth-order all-pole signal. The fitting of the first p+1 autocorrelation ensures fitting the correspoding power density spectrum. When the sources of noise become significant (SNR<30db) the corrupted speech signal is no longer autoregressive. It is autoregressive-moving-average. The spectral noise zeros affect the true AR poles to move towards origin, thus producing a flat spectrum (5). So LPC analysis does not preserve the formant structure of the speech spectrum.

This point is shown in Fig. 1 where the autocorrelation and spectrum of speech-like signal segmant corrupted by additive white noise is shown by a solid curve. The order of AR model is p=2. The autocorrelation and spectrum of all-pole model is plotted with the dotted curve. LPC matches the first two lags exactly, but mismatches higher lags and degrades quickly. The corresponding LPC spectrum becomes flat as shown in Fig. 1(b).

AR modeling in the presence of noise has been a subject of much research (6-10) in recent years, and various modification to LPC have been proposed to retain high performance. Some of them need more computation to estimate the variance of noise, or use the high-order Yule-Walker equations (HOYWE) which delete the first p+1 equations, i.e. all of the original Yule-Walker equations.

By use of the correlation fitting method for AR modeling of speech we get improved results, especially in the case of speech contaminated by noise. The correlation fitting algorithm is derived in section 2. The robustness of CF is described in section 3. Finally, section 4 shows the simulation results and section 5 gives conclusions.

2. AR Modeling by the Correlation Fitting

The AR correlation fitting algorithm (1,2) is derived in this section.

Assume $x(n)$ is a stationary autoregressive stochastic signal. It can be well represented by an AR model after solving the famous Yule-Walker equations

$$\hat{R}_p \underline{a}_p = \sigma^2 \underline{\delta}_p \tag{1}$$

where $\underline{a}_p$ is the coefficient vector of a pth-order AR model, and $\hat{R}_p$ is a Hermitian Toeplitz matrix comprised by the p+1 estimated correlation values of $x(n)$, i.e. $\hat{r}(0), \hat{r}(1), \ldots, \hat{r}(p)$ which are matched exactly by the resulting AR model (4).

In order to improve the performance we use the fitting error criterion

$$e(n) = \hat{r}(n)u(n) - w_T(n)r(n)u(n) \tag{2}$$

with

$$min \sum_{n=0}^{L} e^2(n) \tag{3}$$