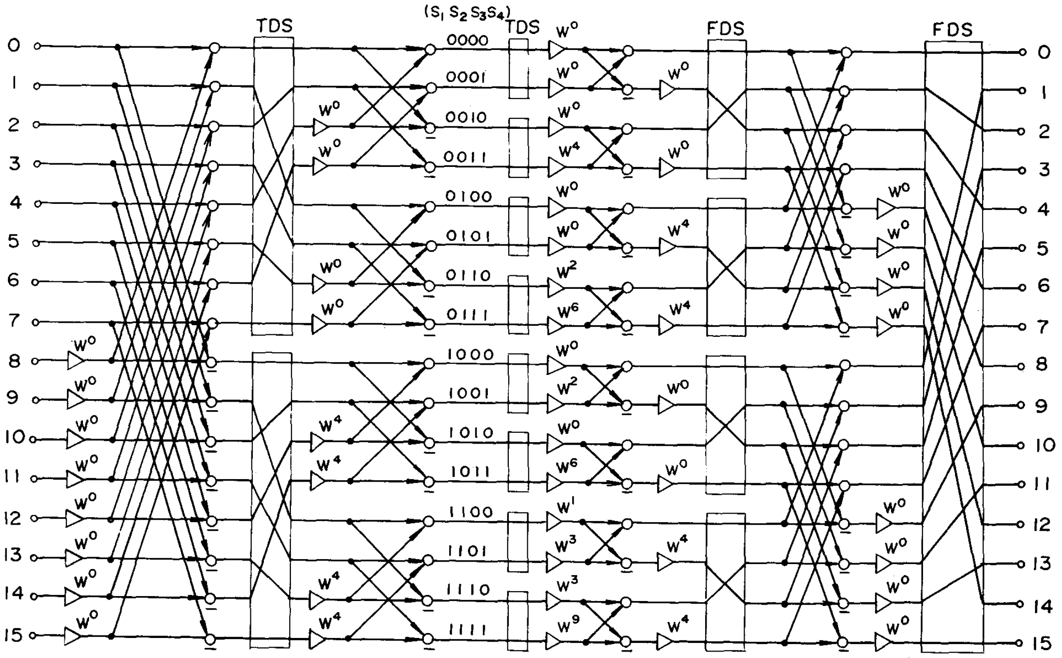

Fig. 1. Block diagram for 16 point radix 2 alternate MDFFT, where $w = \exp(-j2\pi/16)$.

under the same condition. From (27) and (28), MDFFT can save multiplications by 15–20 percent from the FFT algorithm. Furthermore, additions are slightly reduced. These reductions are also obtained for a radix 4 structure.

V. Conclusion

General formulas for submatrices and multiplicands appearing in MDFFT have been presented. Computational efficiency has been briefly discussed.

References

[1] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, pp. 297–301, Apr. 1965.
[2] R. C. Singleton, "An algorithm for computing the mixed radix fast Fourier transform," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 93–103, June 1969.
[3] C. M. Rader and N. M. Brenner, "A new principle for fast Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 264–265, Apr. 1976.
[4] S. Winograd, "On computing the discrete Fourier transform," *Proc. Nat. Acad. Sci. U.S.*, pp. 1005–1006, Apr. 1976.
[5] D. P. Kolba and T. W. Parks, "A prime factor FFT algorithm using high speed convolution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 281–294, Aug. 1977.
[6] K. Nakayama, "Fast Fourier transform using mixed frequency and time decimations (in Japanese)," *IECE Japan*, Rep. Tech. Meet. on Circuits Syst., vol. CAS79-94, pp. 49–54, Oct. 1979.
[7] C. Caraiscos and B. Liu, "Two dimensional DFT using mixed time and frequency decimations," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, May 1982, pp. 24–27.

Estimation of Noise Variance from the Noisy AR Signal and Its Application in Speech Enhancement

K. K. PALIWAL

***Abstract*—In a number of applications involving the processing of noisy signals, it is desirable to know *a priori* the noise variance. We propose here a method of estimating the noise variance from the autoregressive (AR) signal corrupted by the additive white noise. This method first estimates the AR parameters from the high-order Yule–Walker equations, and then uses these AR parameters to estimate the noise variance from the low-order Yule–Walker equations. The method is used in a speech enhancement application where its performance is studied for stationary as well as nonstationary noise conditions. The results are found to be encouraging.**

I. Introduction

In a number of signal processing applications, the signal available for analysis is noisy; i.e., it is corrupted by the addition of background noise which can be assumed to be white in nature for many practical problems of interest. In these applications, it is desirable to know the variance of the additive white noise process. For example, for speech enhancement there is a class of methods (based on short-time Fourier transform magnitude) for which it is

Manuscript received July 16, 1986; revised August 20, 1987.
The author is with the Computer Systems and Communications Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400005, India.
IEEE Log Number 8718020.

necessary to know *a priori* the white noise variance [1]. Similarly, some of the spectral estimation methods for noisy signals [2] require *a priori* knowledge of the noise variance.

In the speech enhancement literature [1], the white noise variance is estimated from the preceding silent portions of speech. However, this method which will be referred hereafter as the conventional method has two major drawbacks. First, detection of silent segments of speech is a difficult problem. When speech is noisy with signal-to-noise ratio (SNR) as low as 0 dB or lower, it becomes almost impossible to detect these silent segments. Second, the noise variance estimated from the preceding silent segments can be used in the processing of the present speech segment only when the noise is stationary in nature. However, in practice, it might be varying with time and the noise-variance estimate from the preceding silent segments may not give good results for the present segment. In this correspondence, we propose a method for estimating the noise variance. This method estimates the noise variance from the same segment of the noisy signal which is available for processing.

II. Noise-Variance Estimation Method

In this section, we describe the method of estimating noise variance from the given segment of the noisy signal. Let the noisy signal to be analyzed be

$$y(n) = x(n) + v(n), \qquad n = 1, 2, \cdots, N, \tag{1}$$

where $\{x(n)\}$ is the uncontaminated signal and $\{v(n)\}$ is the zero-mean white noise of variance σ_v^2. The aim here is to estimate the noise variance σ_v^2 from the observed noisy signal $\{y(n)\}$.

In order to solve this problem, we assume that the uncontaminated signal $\{x(n)\}$ follows the pth-order AR model

$$H(z) = 1 \bigg/ \left[1 + \sum_{i=1}^{p} a_i z^{-i}\right], \tag{2}$$

whose parameters $\{a_i\}$ satisfy the following set of Yule–Walker equations:

$$\sum_{k=1}^{p} a_k R_x(|i-k|) = -R_x(i), \qquad i > 0, \tag{3}$$

where $\{R_x(i)\}$ are the autocorrelation coefficients of the uncontaminated signal $\{x(n)\}$.

Since the additive noise $\{v(n)\}$ is white, the autocorrelation coefficients $\{R_x(i)\}$ of the uncontaminated signal $\{x(n)\}$ are related to the autocorrelation coefficients $\{R_y(i)\}$ of the noisy signal $\{y(n)\}$ as follows:

$$R_x(0) = R_y(0) - \sigma_v^2, \tag{4}$$

and

$$R_x(i) = R_y(i), \qquad |i| > 0. \tag{5}$$

With these preliminaries, we now present a three-step procedure for estimating the noise variance σ_v^2. These steps are outlined below.

Step 1: From the observed (noisy) signal $\{y(n)\}$, compute the unbiased estimates of the autocorrelation coefficients $\hat{R}_y(i)$, $i = 0, 1, \cdots, p + q$, where $q > p$.

Step 2: Compute the least-squares estimate of AR coefficients by using the Cadzow's method [3] from the $q(>p)$ high-order Yule–Walker equations [$i = p + 1, p + 2, \cdots, p + q$ in (3)].

Step 3: Use the AR coefficients obtained from Step 2 and compute the least-squares estimate of the noise variance from the overdetermined set of p low-order Yule–Walker equations [$i = 1, 2, \cdots, p$ in (3)]. This is given by

$$\hat{\sigma}_v^2 = \left[\sum_{i=1}^{p} a_i \left\{\hat{R}_y(i) + \sum_{k=1}^{p} a_k \hat{R}_y(|i-k|)\right\}\right] \bigg/ \sum_{i=1}^{p} a_i^2. \tag{6}$$

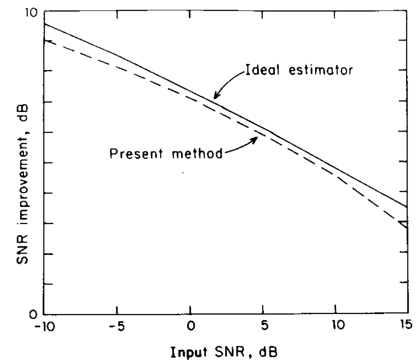

Fig. 1. SNR improvement in the output (enhanced) speech as a function of input SNR.

III. Results

The noise-variance estimation method has been tried out on a number of examples of noisy AR signals, and its performance is found to be close to Cramer–Rao lower bound for signal-to-noise ratios (SNR's) as low as 0 dB. Here, we apply the noise-variance estimation method for speech enhancement and study its performance for stationary as well as nonstationary noise conditions. The spectrum subtraction method [4] is used in the present study for speech enhancement.

A. Speech Enhancement Results for Stationary White Noise

Here, we study the speech enhancement performance of the present noise-variance estimation method for the speech signal corrupted by the addition of stationary white Gaussian noise. We compare the speech enhancement results of this method to those obtained using the ideal noise variance estimates. Continuous speech data of 15-s duration (consisting of 4 English sentences of 4 male and female speakers) is used in the present study.

Speech enhancement performance of the noise-variance estimation method is studied for input SNR values ranging from -10 dB to 15 dB with segment duration = 32 ms, segment overlap = 16 ms, $p = 10$, and $q = 15$. SNR improvement results are shown in Fig. 1. It can be seen from this figure that the present method of noise-variance estimation results in an improvement in SNR for all the input SNR values. For example, for noisy speech with input SNR = 0 dB, this method results in an SNR improvement of 7.4 dB. It can also be seen from Fig. 1 that the results from the present method are inferior to those obtained with the ideal noise-variance estimates by only 0.4 dB. Subjective listening of the output speech has also shown that the present method improves the quality of speech for all input SNR values and there is no audible difference between the output speech obtained through the present noise-variance estimation method and that obtained with the ideal noise variance estimates.

B. Speech Enhancement Results for Nonstationary White Noise

Here, we study the speech enhancement performance of the present noise-variance estimation method for the speech signal corrupted by the addition of nonstationary white Gaussian noise. We compare its performance to that of the conventional method [1] which computes the noise variance from the silent intervals preceding the speech segments.

In order to generate noisy speech with nonstationary white Gaussian noise, the noise variance σ_v^2 is first calculated to get the

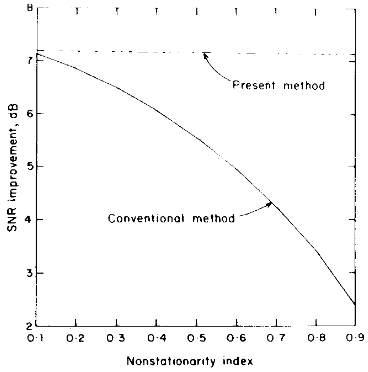

Fig. 2. SNR improvement in the output (enhanced) speech as a function of nonstationarity index.

desired input SNR. The speech data are then corrupted by the addition of zero-mean white Gaussian noise whose variance varies linearly with time with the variance being $(1 - g)\sigma_v^2$ at the beginning and $(1 + g)\sigma_v^2$ at the end of the speech data. The parameter g defines here the measure of nonstationarity of the white noise process and will be referred to as the "nonstationarity index."

Here, we study the noise-variance estimation method for input SNR = 0 dB and evaluate its speech enhancement performance as a function of nonstationarity index. The SNR results for this method are shown in Fig. 2. It can be seen from this figure that as the nonstationarity of the additive white noise process increases, performance of the conventional method deteriorates; while that of the present method remains steady throughout. For the nonstationarity index g equal to 0.9, the present method of noise-variance estimation gives an advantage of about 5 dB in output SNR. Subjective listening of the output speech also confirms these observations.

IV. Conclusion

In this correspondence, a noise-variance estimation method is proposed. This method is applied to the problem of speech enhancement where the spectrum subtraction scheme [4] is used. Speech enhancement performance of this method is studied for speech corrupted by stationary as well as nonstationary additive white Gaussian noise processes and encouraging results are obtained. This method has also been applied to the problem of spectral estimation for noisy AR signals and the results are reported elsewhere [5].

References

[1] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol. 67, pp. 1586-1604, Dec. 1979.

[2] S. M. Kay, "Noise compensation for autoregressive spectral estimates," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 292-303, June 1980.

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

[4] J. S. Lim, "Evaluation of a correlation subtraction method for enhancing speech degraded by additive white noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 471-472, Oct. 1978.

[5] K. K. Paliwal, "A noise-compensated long correlation matching method for AR spectral estimation of noisy signals," in *Proc. ICASSP*, Apr. 1986, pp. 1369-1372.

An Iterative Algorithm for Power Spectrum Estimation in the Maximum Entropy Method

NAILONG WU

***Abstract*—The maximum entropy method (MEM) is a nonlinear method of power spectrum estimation (PSE). It is well known that two forms of entropy are widely in use:**

$$H1 = \int \log S(f)\, df \quad \text{(continuous)}$$

or

$$H1 = \sum \log S(j) \quad \text{(discrete)} \tag{1}$$

(for the details of derivation, see [1] and [2]), and

$$H2 = -\int S(f) \log S(f)\, df$$

or

$$H2 = -\sum S(j) \log S(j) \tag{2}$$

(for details, see [3]), where S is the estimated spectrum, f is the frequency normalized to $[-\frac{1}{2}, \frac{1}{2}]$, and the operator log stands for the natural logarithm.

For MEM1 (MEM with $H1$), an explicit solution exists only in the one-dimensional (1-D) case [1], [2]. Lim and Malik [4] presented an iterative algorithm which can be used in the two-dimensional (2-D) case as well as in the 1-D case. The situation in MEM2 (MEM with $H2$) is much more difficult. Even in the 1-D case, an explicit solution can be determined only after imposing strong conditions (real, causal, and minimum-phase) on the signal [5]. It is otherwise inevitable to use implicit solutions. In this correspondence we present an iterative algorithm in MEM2, which does not require conditions such as causality, minimum-phase, etc., and can be used for any dimensionality.

I. Statement of the MEM2 Problem

The PSE problem in MEM2, with the uniformly sampled autocorrelation function (ACF) being known, can be stated as follows [6].

Given the ACF $R(n)$, $n \in D$, estimate the spectrum $S(j)$ such that

1) the constraints

$$R(n) = IFT[S(j)] \qquad \text{for } n \in D \tag{3}$$

are satisfied, where IFT stands for the inverse Fourier transform, and

2) the Lagrange undetermined multiplier (LUM)

$$\lambda_n^* \begin{cases} \text{to be determined}, & n \in D; \\ = 0, & \text{otherwise}, \end{cases} \tag{4a}$$

where

$$\begin{aligned} \lambda_n^* = \lambda_{-n} &= \delta_n + IFT[\log S(j)] \\ &= \delta_n + IFT[\log FT[R(n)]] \end{aligned} \tag{4b}$$

(* denotes complex conjugate). The spectrum

$$S(j) = \exp\{-1 + FT[\lambda_n^*]\}. \tag{5}$$

Manuscript received October 4, 1985; revised September 4, 1987.
The author is with Peking Astronomical Observatory, Academia Sinica, Peking, China.
IEEE Log Number 8718026.