A SINGLE MICROPHONE NOISE CANCELLER BASED ON ADAPTIVE KALMAN FILTER

M. Gabrea, E. Mandridake and M. Najim
Equipe Signal et Image, ENSERB and GDR-134, CNRS
BP 99, 33 402 Talence, FRANCE
email: najim@goelette.tsi.u-bordeaux.fr

ABSTRACT

This paper deals with the problem of Adaptive Noise Cancellation (ANC) when only corrupted speech signal with an additive Gaussian white noise is available for processing. We propose a new method based on adaptive Kalman filtering. All the approaches based on the Kalman filter proposed in the past, in this context, operate in two steps: they first estimate the noise variance and the parameters of the signal model and secondly estimate the speech signal. The approach presented in this paper gives an alternative to these approaches since it does not require the estimation of the noise variance. The noise variance estimation is a part of the Kalman gain calculation. For optimizing the Kalman gain we have reformulated and adapted, to the single-microphone ANC problem, the approach proposed in control by R. K. Mehra.

1 INTRODUCTION

Speech enhancement using a single microphone system has become an active research area for audio signal enhancement. The aim is to retrieve the desired speech signal from the noisy observations. These problems occur, for example, in hands free mobile phones and teleconferencing.

In the standard ANC systems one uses at least two microphones. One microphone to capt the observation signal and another one to serve as the noise reference.

Many approaches for speech enhancement based on the Kalman filtering [1-4] have been reported in the literature. These approaches differ essentially one from the other by the algorithm used to estimate the parameters of such a model.

A time-adaptive algorithm is used in [1] to adaptively estimate the speech model parameters and the noise variance. In [2] the ideal values of the parameters have been used and a delayed-Kalman filter is proposed. In [3] the speech signal is considered as an output of an ARMA process and an adaptive Kalman filter is used to estimate the speech signal. The estimation method of the speech model parameters used in [4] is a suboptimal solution that can be considered as a version of the Estimate-Maximize (EM) algorithm based on the maximum likelihood argument.

In this paper we reformulate, for the speech enhancement, the approach proposed by R. K. Mehra in the field of control [5]. In this approach, signal and noise variances estimation are handled by the optimization gain procedure. So, this new approach looks very attractive in comparison to the ones where these heavy tasks must also be made.

This paper is organized as follows. We present in section 2 the state-space model representation of the noisy speech observation and Kalman filtering. The section 3 is concerned with the presentation of the estimation of the AR parameters and the different steps of Mehra algorithm for the Kalman gain optimization. The section 4 presents our single-microphone ANC system based on the Kalman gain optimization. In the last section we provide experimental results and evaluate the performance of the proposed system.

2 NOISY SPEECH MODEL AND KALMAN FILTERING

Let us consider the speech signal modelled as a p order AR process:

$$s(n) = \sum_{i=1}^{p} a_i s(n-i) + u(n) \tag{1a}$$

$$z(n) = s(n) + v(n) \tag{1b}$$

This system can be represented by the following state-space model:

$$\mathbf{x}(n+1) = \mathbf{\Phi x}(n) + \Gamma u(n) \tag{2}$$

$$z(n) = \mathbf{Hx}(n) + v(n) \tag{3}$$

where:

$$\mathbf{x}(n+1) = [s(n-p+1),\ \cdots,\ s(n)]^T \tag{4}$$

is the state vector,

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ a_p & a_{p-1} & \cdots & a_1 \end{bmatrix} \quad (5)$$

is the state-transition matrix, $\mathbf{\Gamma}$ and $\mathbf{H}$ are the input and the output matrices defined by

$$\mathbf{\Gamma} = \mathbf{H}^T = [0,0,\cdots,0,1]^T \quad (6)$$

u(n) and v(n) are uncorrelated Gaussian white noises with means and covariances given below:

$$E\{u(n)\}=0;\ E\{u(n)u(n-m)\} = \sigma_u^2\delta(n-m) \quad (7)$$

$$E\{v(n)\}=0;\ E\{v(n)v(n-m)\} = \sigma_v^2\delta(n-m) \quad (8)$$

$$E\{u(n)v(m)\} = 0 \quad \forall n,m \quad (9)$$

The standard Kalman filter provides the following updating state-vector estimation [7]:

$$\hat{\mathbf{x}}(n+1/n)=\mathbf{\Phi}\hat{\mathbf{x}}(n/n-1)+\mathbf{\Phi}\mathbf{K}e(n) \quad (10)$$

$$z(n) = \mathbf{H}\hat{\mathbf{x}}(n/n-1) + e(n) \quad (11)$$

where $e(n)$ is the innovation sequence.

Here $\mathbf{\Phi}$ and $\mathbf{K}$ are unknown and hence must be estimated. Then, the updating state vector estimation becomes:

$$\hat{\mathbf{x}}(n+1/n) = \hat{\mathbf{\Phi}}\hat{\mathbf{x}}(n/n-1) + \hat{\mathbf{\Phi}}\hat{\mathbf{K}}e(n) \quad (12)$$

where $\hat{\mathbf{\Phi}}$ and $\hat{\mathbf{K}}$ are respectively the estimated state-transition matrix and the Kalman gain. $\hat{\mathbf{\Phi}}$ is in fact including the parameters to be estimated.

The estimation of $\mathbf{\Phi}$ and $\mathbf{K}$ is the object the following section.

3. PARAMETER ESTIMATIONS

3.1 Transition matrix estimation

The estimation of $\mathbf{\Phi}$ is achieved this way: one needs to estimate first the autocorrelation of the observation z(n) as follows:

$$\hat{c}(k) = \frac{1}{n}\sum_{i=k}^{n} z(i)z(i-k) \quad (13)$$

and, in the second step, the parameters of the AR process using (13) and the Cayley-Hamilton theorem applied for the matrix $\mathbf{\Phi}$.

$$\begin{bmatrix} \hat{a}_p \\ \vdots \\ \hat{a}_1 \end{bmatrix} = \begin{bmatrix} \hat{c}(1) & \cdots & \hat{c}(p) \\ \vdots & \ddots & \vdots \\ \hat{c}(p) & \cdots & \hat{c}(2p-1) \end{bmatrix}^{-1} \begin{bmatrix} \hat{c}(p+1) \\ \vdots \\ \hat{c}(2p) \end{bmatrix} \quad (14)$$

The estimation of the unknowns parameters in $\hat{\mathbf{\Phi}}$ are obtained from a set of algebraic equations used to estimate the AR parameters [6].

Using these estimated AR parameters one can obtain the estimation of the state-transition matrix:

$$\hat{\mathbf{\Phi}} = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \hat{a}_p & \hat{a}_{p-1} & \cdots & \hat{a}_1 \end{bmatrix} \quad (15)$$

The canonical representation (2) and (3) is complete if we can estimate the variances of the additive noise v(n) and the process noise u(n).

The estimation of additive noise variance is straightforward from the estimated AR parameters and the estimated autocorrelation of the observation z(n) [6].

The direct estimation of the variance of the process noise can be avoided. To avoid a direct calculation of the noise variance we adopt the Mehra approach which permits, by providing an estimated Kalman gain, to retrieve the speech signal by Kalman filtering.

3.2 Kalman gain estimation

In fact, our aim is to retrieve the estimated signal from the observation sequence using equations (10) and (11). So, this can also be done by estimating the Kalman gain instead of the estimation of the noise variance.

The estimation of K proposed by Mehra is based on the statistical test of the whiteness of the innovation sequence.

For an optimal Kalman filter where $\mathbf{\Phi}$ and $\mathbf{K}$ are known, this innovation sequence is a Gaussian white noise sequence [8]. We are not in the presence of an optimal Kalman filter since $\mathbf{\Phi}$ and $\mathbf{K}$ must be estimated. However, the innovation sequence is assumed to be a stationary Gaussian random sequence and an iterative procedure can be conducted to derive an asymptotical optimal Kalman filter.

We give in the following the outlines of the gain optimization algorithm, for more details see [5].

Let us call $\hat{\mathbf{K}}(0)$ the initial Kalman gain value.

The gain $\hat{\mathbf{K}}$ is estimated by the iterative algorithm:

$$\hat{\mathbf{K}}(i) = \hat{\mathbf{K}}(i-1) + \begin{bmatrix} \mathbf{H}\hat{\mathbf{\Phi}} \\ \mathbf{H}\hat{\mathbf{\Phi}}[\mathbf{I}-\hat{\mathbf{K}}(i-1)\mathbf{H}]\hat{\mathbf{\Phi}} \\ \vdots \\ \mathbf{H}\left[\hat{\mathbf{\Phi}}[\mathbf{I}-\hat{\mathbf{K}}(i-1)\mathbf{H}]\right]^{p-1}\hat{\mathbf{\Phi}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\gamma}(1) \\ \hat{\gamma}(2) \\ \vdots \\ \hat{\gamma}(p) \end{bmatrix} / \hat{\gamma}(0) \quad (16)$$

where $\hat{\gamma}(k)$ is the autocorrelation of the innovation sequence at the lag k:

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{i=k}^{n} e(i)e(i-k) \quad (17)$$

A statistical test on the whiteness of the innovation is used to check whether the Kalman filter is working optimally or not. If it is not, one estimates the autocorrelation of the new innovation sequence to adapt the Kalman gain based on the equation (16), and so on until the optimal Kalman gain is reached. Hence, the estimated state-vector $\hat{\mathbf{x}}$ is updated using the new value of $\hat{\mathbf{K}}$ according to the above test. Finally, the estimated speech signal can be retrieved from the equation:

$$\hat{s}(n) = \mathbf{H}\hat{\mathbf{x}}(n/n) = \mathbf{H}\hat{\mathbf{\Phi}}\hat{\mathbf{x}}(n/n-1) \quad (18)$$

In the next section, we present the application of this approach to achieve noise cancellation system when only a single observation containing the speech signal and the unwanted noise is available.

4. NOISE CANCELLER BASED ADAPTIVE KALMAN FILTER

Our proposed single microphone ANC system is sketched in the figure 1. The observation sequence z(n) contains the speech component and the unwanted noise v(n). The AR parameters estimation of the speech model are included in the transition matrix $\mathbf{\Phi}$ estimation using the procedure described in 3.1. The innovation sequence drives the iterative algorithm for optimizing the Kalman gain.

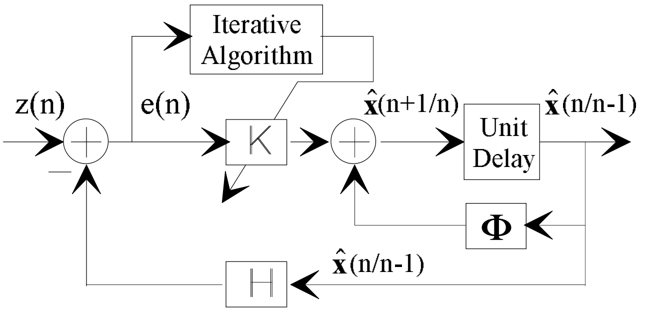

Fig. 1: Single microphone ANC system based on an adaptive gain Kalman filter

As far as we know, in all the ANC techniques, one estimates the noise variance during silence period. In this paper the noise variance estimation is a part of Kalman gain estimation. Actually, since Kalman filter requires the AR parameters and the noise variances, the noise variance estimation does not appear explicitly in the iterative algorithm of the gain optimization. This algorithm can be considered, in a sense, as operating in a global way in comparison to other approaches which can be considered as operating in multistage procedure.
The Kalman gain, and consequently the noise variance, is estimated using the autocorrelation of the innovation sequence instead of the autocorrelation of the observation sequence. It is an advantage in running the iterative procedure since the Kalman gain are adapted with the innovation sequence less correlated than the observation sequence. This fact has been confirmed in the simulation results where the iterative algorithm converges at most in 4 iterations as precisely stated by Mehra [5].

5 SIMULATIONS AND RESULTS

The effectiveness of the method is tested using natural speech signal corrupted by a Gaussian white noise. The order p of the AR process of speech signal has been fixed to 5. Using AR process order higher than 5 does not improve the SNR.
An example of speech enhancement results is reported in the Table 1. A SNR improvement from 1.57 dB to 10.48 dB has been obtained. For these results the iterative algorithm converges in 2 iterations. Figures 2, 3 and 4 represent, respectively, the cepstrogram followed by the time signal of the free-noise speech, the noisy speech and the enhanced speech. For this example, the SNR of the noisy speech signal is 0 dB.

Input SNR (dB)	-10	-5	0	5	10
Gain SNR (dB)	10.48	8.13	5.78	3.50	1.57

Table 1: Gain SNR for different input SNR

ACKNOWLEDGEMENTS

We would like to thank the ″Ministère des Affaires Etrangères (France) ″ for supporting one of the authors (M. Gabrea). We would also like to acknowledge MATRA Communication Company (Paris) who has gently provided the recorded speech signals.

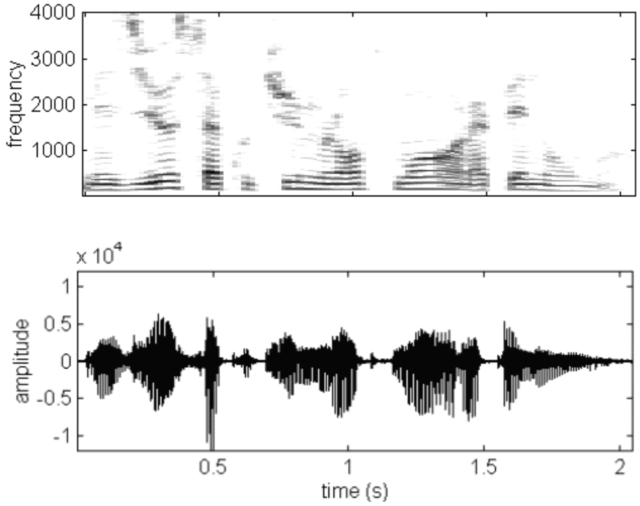

Fig. 2: Noise-free speech signal

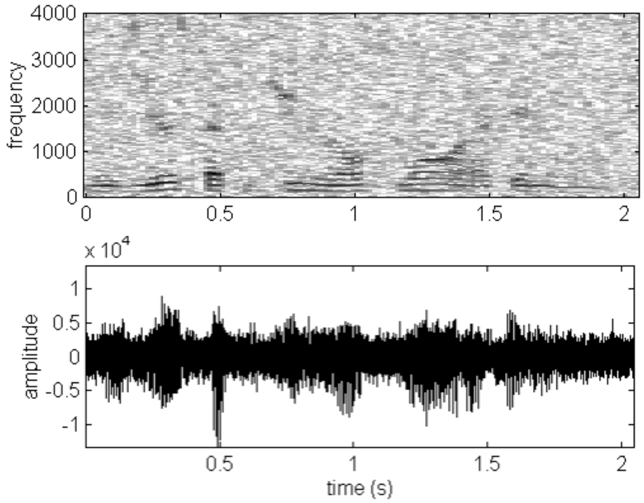

Fig. 3: Noisy speech signal

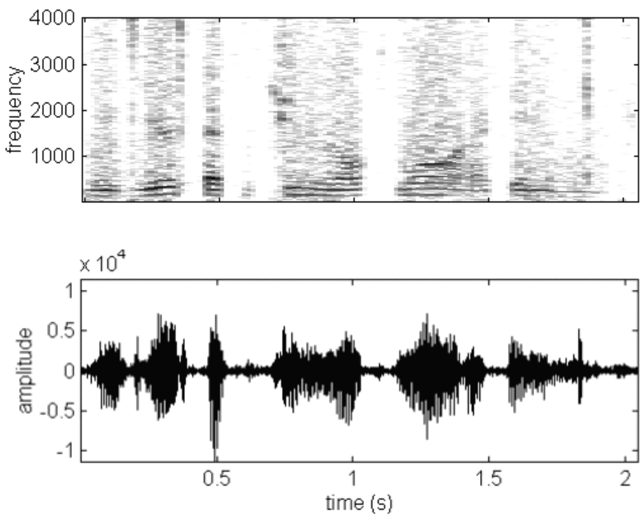

Fig. 4: Enhanced speech signal

REFERENCES

[1] A. V. Oppenheim, E.Weinstein, K.C.Zangi, M.Feder and D.Gauger, "Single-Sensor Active Noise Cancellation," IEEE Trans. on Speech and Audio Processing, Vol. 2,No. 2, pp. 285-290, April 1994.

[2] K. K. Paliwal and A.Basu, "A Speech Enhancement Method Based on Kalman Filtering," ICASSP 87, pp. 177-180.

[3] H. Morikawa and H. Fujisaki, "Noise Reduction of Speech Signal by Adaptive Kalman Filtering," Special issue in Signal Processing APII-AFCET-Edited by M.Najim, Vol. 22, No. 1, pp. 53-68, 1988.

[4] J. D. Gibson, B. Koo and S. D. Gray, " Filtering of Colored Noise for Speech Enhancement and Coding," IEEE Trans. on Signal Processing, Vol. 39, No.8, pp. 1732-1742, August 1991.

[5] R. K. Mehra, "On the Identification of Variances and Adaptive Kalman Filtering," IEEE Trans. on Automatic Control. Vol. AC-15, No. 2, pp. 175-184, April 1970.

[6] R. K. Mehra, "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," IEEE Trans. on Automatic Control. Vol. AC-16, No. 1, pp. 12-21, February 1971.

[7] M. Najim, *Modelization and Identifications in Signal Processing*, Masson, Paris 1988 (in French).

[8] T. Kailath, " An innovations approach to least-squares estimation, part I: linear filtering in additive white noise," IEEE Trans. Automatic Control, Vol. AC-13, pp. 646-655, December 1968.