Adaptive crosstalk-resistant noise-cancellation using Kalman filters.

T.J.Moir— School of Engineering, AUT University, Auckland, New-Zealand Tom.Moir@aut.ac.nz

Introduction

The problem of suppression of a non-stationary noise from a non-stationary signal has been studied for decades with varying degrees of success. For example, the simplest method of minimizing mean-square error for stationary signals and noise was solved by Wiener[1] (continuous time) and Kolmogorov[2] (discrete-time). The non-stationary case was first solved by Kalman[3]. These approaches all required an accurate model of the signal and noise generating processes. This can be quite appropriate for problems in say rocket research, but in acoustic problems where the signal and noise are both non-stationary speech, it is quite impractical. Instead, other approaches were taken much later using the least-mean-squares (LMS) algorithm as an estimator of the statistical models of the signals and noise. This work had its foundations with the work of Widrow et al [4]. The idea was to use more than one sensor. One sensor could measure the noise in an area which was signal-free. This “noise-alone” measurement could then be phase and amplitude shifted using an adaptive LMS algorithm (as part of an adaptive filter) and subsequently an estimate of the noise obtained which could be subtracted from the signal plus noise mixture. The phase and amplitude shifting is done with respect to signal plus noise so as to minimize mean-square error. Such an approach requires a degree of isolation so that the noise measurement is free from signal, otherwise both signal and noise will be cancelled or at least the performance will be severely impaired. In the case of stationary signal and noise, the converged filter becomes the optimal Wiener filter as approximated by a finite-impulse response (FIR) adaptive filter. However, making one sensor signal-free necessitates moving the sensors far apart, which in turn leads to large filter orders.

As an alternative, a few authors use a similar approach which puts both sensors close together rather than far apart.[5, 6] By switching two LMS algorithms in turn, one during signal and noise and one during noise-alone, the signal can again be estimated that minimizes mean-square error. Although this idea has elegance, it requires an extra algorithm to detect when noise is present without the signal. When dealing with speech signals such an algorithm has become known as a voice-activity detector (VAD). The estimate of the noise on its own via the VAD determines the performance of the overall adaptive filter. It fact it is a hard problem to discriminate what is noise and what is crosstalk signal when both are non-stationary, as with the case of two speech signals. A good approach which is immune to absolute threshold values of energy and instead works on the geometry of the positioning of desired speech is given in[7]. Generally speaking though, even the best VAD is not good enough to distinguish one signal from another and any errors in distinguishing the noise (or undesired speech) will result in cancellation of the desired speech, as is case with the original Widrow approach. A method therefore had to be found which was immune to crosstalk and eliminate VAD’s altogether.

Such an approach was used in[8, 9], and involved a cross-coupling of the error signals in two parallel running LMS identification algorithms

This method was much later extended to microphone arrays[10]. The technique does not guarantee convergence to the optimal values, but is nevertheless is a successful practical method in that the microphones (or sensors) can be placed closely together and no VAD is required. Although other approaches exist (for example Independent Component Analysis (ICA)[11]) they are in general far more computationally demanding. The newer methods have become known as “Blind-Source” separation (BSS) since no *apriori* knowledge is assumed about the noise or signal statistics other than that they are statistically independent.

Methodology

The method below is known in the literature as backwards separation[8, 13, 14]. This method is more of a decorrelation method than a direct signal estimation method. However, there remains the problem of estimating the unknown coupling transfer functions .

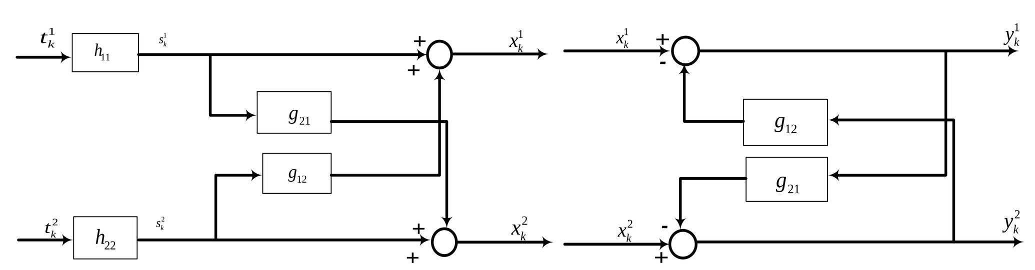

Fig 1.Signal Generation by mixing (left) and the separation process (right)

We can write two cross-coupled Kalman filters that can perform blind-source separation of two random signals. We use the model of Figure 1 as the mixing and un-mixing process.

Parameter Updates:

$$\hat{\boldsymbol{\theta}}_{k+1}^1 = \hat{\boldsymbol{\theta}}_k^1 + \boldsymbol{K}_k^1 \boldsymbol{U}_k^1 \varepsilon_k^1$$

$$\hat{\boldsymbol{\theta}}_{k+1}^2 = \hat{\boldsymbol{\theta}}_k^2 + \boldsymbol{K}_k^2 \boldsymbol{U}_k^2 \varepsilon_k^2$$

Innovations updates:

$$\varepsilon_k^1 = x_k^1 - \left[\boldsymbol{U}_k^1\right]^T \hat{\boldsymbol{\theta}}_k^1$$

$$\varepsilon_k^2 = x_k^2 - \left[\boldsymbol{U}_k^2\right]^T \hat{\boldsymbol{\theta}}_k^2$$

Each of the regressor vectors have values determined by the others innovations process

$$\boldsymbol{U}_k^1 = [\varepsilon_{k-1}^2, \varepsilon_{k-2}^2 \ldots \varepsilon_{k-n}^2]^T$$

and

$$\boldsymbol{U}_k^2 = [\varepsilon_{k-1}^1, \varepsilon_{k-2}^1 \ldots \varepsilon_{k-n}^1]^T$$

The n-square positive definite error-covariance matrices are found iteratively from

$$\boldsymbol{P}_k^1 = \boldsymbol{P}_{k-1}^1 - \frac{\boldsymbol{P}_{k-1}^1 \boldsymbol{U}_k^1 \left[\boldsymbol{U}_k^1\right]^T \boldsymbol{P}_{k-1}^1}{\sigma_v^2 + \left[\boldsymbol{U}_k^1\right]^T \boldsymbol{P}_{k-1}^1 \boldsymbol{U}_k^1} + \sigma_\xi^2 \boldsymbol{I}_n$$

$$\boldsymbol{P}_k^2 = \boldsymbol{P}_{k-1}^2 - \frac{\boldsymbol{P}_{k-1}^2 \boldsymbol{U}_k^2 \left[\boldsymbol{U}_k^2\right]^T \boldsymbol{P}_{k-1}^2}{\sigma_v^2 + \left[\boldsymbol{U}_k^2\right]^T \boldsymbol{P}_{k-1}^2 \boldsymbol{U}_k^2} + \sigma_\xi^2 \boldsymbol{I}_n$$

Finally, the Kalman gain vectors are updated according to

$$\boldsymbol{K}_k^1 = \frac{\boldsymbol{P}_{k-1}^1 \boldsymbol{U}_k^1}{\sigma_v^2 + \left[\boldsymbol{U}_k^1\right]^T \boldsymbol{P}_{k-1}^1 \boldsymbol{U}_k^1}$$

$$\boldsymbol{K}_k^2 = \frac{\boldsymbol{P}_{k-1}^2 \boldsymbol{U}_k^2}{\sigma_v^2 + \left[\boldsymbol{U}_k^2\right]^T \boldsymbol{P}_{k-1}^2 \boldsymbol{U}_k^2}$$

Example: Consider the mixture of helicopter noise and pure speech signal through an AR-3 vector process given by:

$$\hat{\boldsymbol{H}}(z^{-1}) = \boldsymbol{A}(z^{-1})^{-1} \boldsymbol{C}(z^{-1})$$

where

$$\boldsymbol{A}(z^{-1}) = \begin{bmatrix} 1 & -0.4 \\ 0.5 & 1 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.8 \\ 0.2 & -0.4 \end{bmatrix} z^{-1} + \begin{bmatrix} 0.6 & 0.5 \\ 0.3 & 0.4 \end{bmatrix} z^{-2} + \begin{bmatrix} 0.2 & 0.3 \\ 0.4 & 0.2 \end{bmatrix} z^{-3}$$

and

$$\boldsymbol{C}(z^{-1}) = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

The two channel outputs are shown in Figures 2 and 3 respectively.

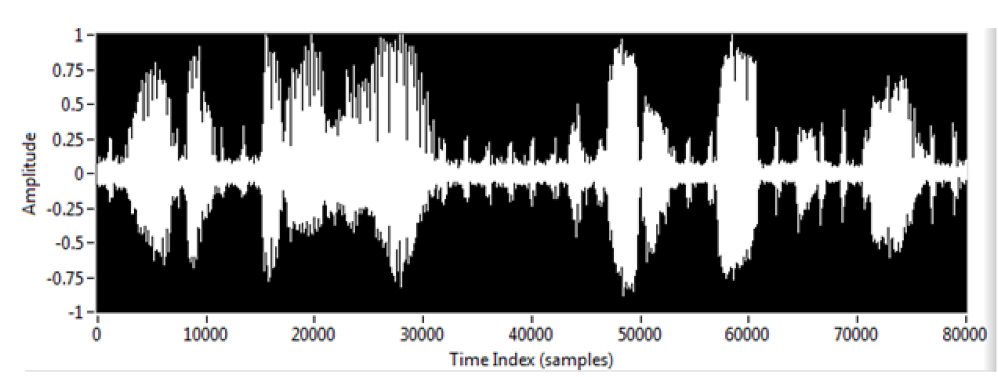

Figure 2 Channel 1 convolutive mixture

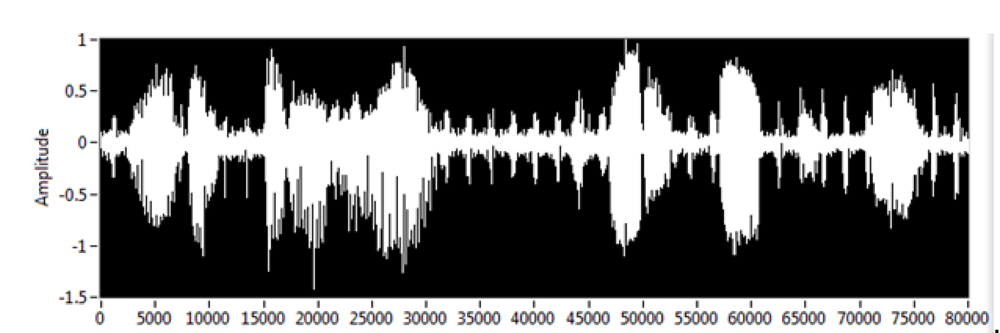

Figure3 Channel 2 convolutive mixture

Table 1. Comparison of NRR’s for various methods.

Method	LMS	RLS	Kalman	Nat Grad
NRR(dB)	4.7dB	6dB	11dB	11dB

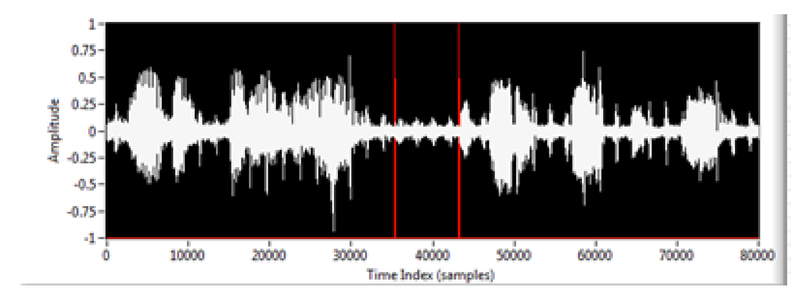

Figure 4 Signal estimate using cross-coupled LMS

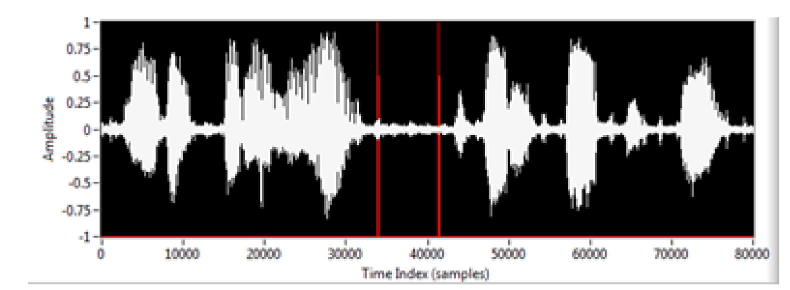

Figure 5 Signal estimate using cross-coupled Kalman filters

Conclusion

An extension to the cross-coupled LMS approach to adaptive filtering has been shown. The new method which uses cross-coupled Kalman filters, has a similar performance to the Natural Gradient method and outperforms LMS and RLS.

References

[1] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*: New York,Wiley, 1949.

[2] A. N. Kolmogorov, "Stationary sequences in Hilbert Space (English Trans In Kailath,T (ed) Linear Least Squares Estimation, pp.66-89, Dowden, Hutchinson & Ross, Pennsylvania 1977)," *Bolletin Moskovskogo Gosudarstvenogo Universiteta.Matematika,* vol. 2, pp. 1-40, 1941.

[3] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Trans ASME,Journal of Basic Engineering,* vol. 82 Series D, pp. 35-45, 1960.

[4] B. Widrow, J. R. Glover JR, J. M. Mc Cool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong JR, and R. C. Goodlin, "Adaptive noise cancelling:Principles and applications.," *IEEE Proceedings,* vol. 63, pp. 1692-1716, 1975.

[5] D. V. Compernolle, Leuven, K.U., "Switching adaptive filters for enhancing noisy and reverberant speech from microphone array recordings," in *Proceedings of the IEEE international conference on acoustics, speech and signal processing*, Albuquerque, 1990, pp. 833-836.

[6] J. Vanden Berghe and J. Wouters, "An adaptive noise canceller for hearing aids using two nearby microphones," *Journal of the Acoustical Society of America,* vol. 103, pp. 3621-3626, Jun 1998.

[7] H. Agaiby and T. J. Moir, "A robust word boundary detection algorithm with application to speech recognition," in *Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on*, 1997, pp. 753-755 vol.2.

[8] R. Zinser, Jr., G. Mirchandani, and J. Evans, "Some experimental and theoretical results using a new adaptive filter structure for noise cancellation in the presence of crosstalk," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '85.*, 1985, pp. 1253-1256.

[9] G. Mirchandani, R. L. Zinser, Jr., and J. B. Evans, "A new adaptive noise cancellation scheme in the presence of crosstalk [speech signals]," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on,* vol. 39, pp. 681-694, 1992.

[10] T. Mei and F. Yin, "Blind separation of convolutive mixtures by decorrelation," *Signal Processing,* vol. 84, pp. 2297-2313, 2004.

[11] S. Choi, A. Cichocki, H. Park, and S. Lee, "Blind Source Separation and Independent Component Analysis: A Review.," *Neural Information Processing (Letters and Reviews),* vol. 6, pp. 1-57, 2005.

[12] L. Lepauloux, P. Scalart, and C. Marro, "A Efficient Low-Complexity Algorithm for Crosstalk-Resistant Adaptive Noise-Canceller," in *17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow,Scotland, August 24-28, 2009, pp. 204-208.

[13] S. Van Gerven and D. Van Compernolle, "Signal separation by symmetric adaptive decorrelation: stability, convergence, and uniqueness," *Signal Processing, IEEE Transactions on,* vol. 43, pp. 1602-1612, 1995.

[14] E. Weinstein, M. Feder, and A. V. Oppenheim, "Multi-channel signal separation by decorrelation," *Speech and Audio Processing, IEEE Transactions on,* vol. 1, pp. 405-413, 1993.