Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 53(67), Fascicola 2, 2008

Two Microphones Speech Enhancement System Based on Instrumental Variable Algorithms

Marcel Gabrea[1]

Abstract – In this paper a symmetric feedback implementation scheme of a two microphones speech enhancement is presented. This approach can be extended for a subclass of signal separations where the direct link is stronger than the interference link in the both channels. We consider the coupling systems modeled as a linear time-invariant Finite Impulse Response (FIR) filters and propose new instrumental variable-based adaptive filters solution to enhance the noisy speech. The optimum filter weight adaptation is based on two instrumental variable algorithms: the generalized least mean square (GLMS) algorithms and the overdetermined recursive instrumental variable (ORIV) algorithms. A comparative study with other adaptive algorithms is presented.

Keywords: speech enhancement, adaptive filters, instrumental variable, generalized least mean square, overdetermined recursive instrumental variable.

I. INTRODUCTION

Let us consider the system modeled by the diagram represented in the Fig. 1. The purpose is to recover the free noise speech signal s(n) from the two available observations $p_1(n)$ and $p_2(n)$ in the presence of the noise signal $b(n)$.

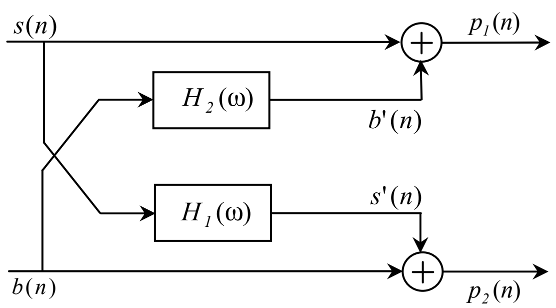

Fig. 1. Signal model for noise cancellation

The primary input source $p_1(n)$ is assumed to contain the speech signal $s(n)$ plus an additive noise component $b'(n)$, and the secondary or the reference is assumed to contain the noise $b(n)$ plus a speech component $s'(n)$. $b(n)$ is correlated with $b'(n)$ but not with the speech signal $s(n)$. The basic scheme of adaptive noise canceller given in [1] uses an adaptive filter based on the Least Mean Squares (LMS) algorithm for estimating the additive noise $b'(n)$, which is then subtracted from the primary input. One problem with the adaptive noise canceling algorithm is the need for the reference microphone to be well separated from the primary microphone, so that it picks up as little speech as possible. If the microphones are too close to one another, cross talk occurs and a typical adaptive filter will thereby suppress a portion of the input speech characteristics. One means of addressing this problem is to place a second adaptive filter in the feedback loop.

In the simplified case where the filters $H_1(\omega)$ and $H_2(\omega)$ are assumed to be single tap another system called Symmetric Adaptive Decorrelation (SAD) using two adaptive filters, as an extension of the classical LMS acoustic noise canceller, has been presented in [2]. This result has been later generalized to a convolutive mixtures modeled by two FIR filters $H_1(\omega)$ and $H_2(\omega)$ [3].

The feedback implementation of an adaptive noise canceller (see Fig. 2) has been proposed in [4] using Double Least Mean Squares (DLMS) algorithm. Other noise cancellers using two adaptive filters: feedforward and feedback symmetric adaptive noise canceller have been described in [5][6][7][8].

In this paper we present a new feedback implementation of a noise canceller based on the two instrumental variable algorithms: the generalized least mean square (GLMS) algorithms and the overdetermined recursive instrumental variable (ORIV) algorithms. We only suppose that the speech signal and the noise are statistically independents and we consider the coupling systems being FIR filters. These algorithms can also be used for a subclass of signal separations where the direct link must be stronger than the interference link in the both channels. A comparative performance study is presented in the framework of noise cancellation.

[1] École de technologie supérieure, Département de génie électrique, 1200 Notre-Dame Ouest, H3C 1K3 Montréal, e-mail marcel.gabrea@etsmtl.ca

II. PROPOSED ALGORITHMS

A. *Introduction*

Fig. 2 shows the feedback implementation of the noise canceller. $W_1(\omega)$ and $W_2(\omega)$ are two adaptive filters. Each one has as input the output error signal of the other filter. $W_1(\omega)$ is an adaptive filter which has an input signal $s_1(n)$, a desired signal $p_2(n)$ and an error signal $s_2(n)$. $W_2(\omega)$ is an adaptive filter which has an input $s_2(n)$ and an error signal $s_1(n)$.

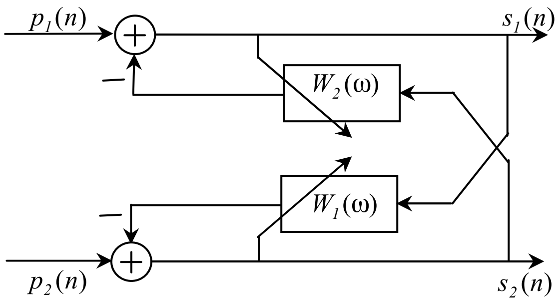

Fig. 2. Feedback implementation of the noise canceller

The optimum values in the Wiener sense, in the case of wide sense stationary processes and in term of the power density spectrum, of the filters $W_1(\omega)$ and $W_2(\omega)$ are given by [9]:

$$W_1(\omega) = \frac{S_{p_2 s_1}(\omega)}{S_{s_1 s_1}(\omega)} \quad (1)$$

$$W_2(\omega) = \frac{S_{p_1 s_2}(\omega)}{S_{s_2 s_2}(\omega)} \quad (2)$$

and if we suppose that the speech signal $s(n)$ and the noise $b(n)$ are two uncorrelated processes we can rewrite (1) and (2) for $i = 1,2$ as follows [7]:

$$S_{ss}(\omega)\big(H_1(\omega) - W_1(\omega)\big)\big(1 - H_1^*(\omega)W_2^*(\omega)\big)/D_i(\omega) +$$
$$S_{bb}(\omega)\big(H_2^*(\omega) - W_2^*(\omega)\big)\big(1 - W_1(\omega)H_2(\omega)\big)/D_i(\omega) = 0 \quad (3)$$
$$i = 1,2$$

where:

$$D_1(\omega) = S_{ss}(\omega)\left|1 - H_1(\omega)W_2(\omega)\right|^2 + S_{bb}(\omega)\left|H_2(\omega) - W_2(\omega)\right|^2 \quad (4)$$

$$D_2(\omega) = S_{ss}(\omega)\left|H_1(\omega) - W_1(\omega)\right|^2 + S_{bb}(\omega)\left|1 - H_2(\omega)W_1(\omega)\right|^2 \quad (5)$$

We can see that the equations (3) provide multiple solutions. Among all these solutions we can find the "*desired solution*" $W_i(\omega) = H_i(\omega)$, $i = 1,2$. In this case it is easy to verify that $s_1(n) = s(n)$ and $s_2(n) = b(n)$ and it is possible to recover the signals that would have been measured at each microphone in the absence of the other source signal.
If for each generating filter:

$$\sum h_i^2(n) < 1,\; i = 1,2 \quad (6)$$

then the filters $W_i(\omega)$ $(i = 1,2)$ converge to the desired solutions.
These desired solutions can be reached using weight adaptive filters.

B. *The overdetermined recursive instrumental variable (ORIV) algorithm*

The ORIV algorithm originally proposed in [10] tries to solve the following matrix equation:

$$\mathbf{Rw} = \mathbf{r} \quad (7)$$

where $\mathbf{R} = E\{\tilde{\mathbf{x}}(n)\mathbf{x}^T(n)\}$, $\mathbf{r} = E\{\tilde{\mathbf{x}}(n)d(n)\}$, and $\mathbf{w}$ is the unknown solution of the adaptive filter. The vectors $\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-q+1)]^T$ and $\tilde{\mathbf{x}}(n) = [\tilde{x}(n), \tilde{x}(n-1), \cdots, \tilde{x}(n-l+1)]^T$ with $l > q$, while the random processes $\{x(n)\}, \{\tilde{x}(n)\}$, and $d(n)$ are input data, instrumental variable and desired response of the adaptive filter. The algorithm is defined as follows [10]:

- *Initial conditions*:

$$\mathbf{R}^T(0) = \beta[\mathbf{I}_{qxq} \,|\, \mathbf{0}_{qxl-q}],\; \beta = \textit{arbitrary scalar} \quad (8)$$

$$\Gamma^{-1}(0) = \frac{1}{\beta^2}\mathbf{I}_{qxq},\; \mathbf{w}(0) = \mathbf{0},\; \mathbf{r}(0) = \mathbf{0} \quad (9)$$

- *Recursive process*:

$$\mathbf{X}(n) = \left[\mathbf{R}^T(n-1)\tilde{\mathbf{x}}(n) \quad \mathbf{x}(n)\right] \quad (10)$$

$$\lambda^2\Lambda(n) = \begin{pmatrix} -\tilde{\mathbf{x}}^T(n)\tilde{\mathbf{x}}(n) & \lambda \\ \lambda & 0 \end{pmatrix} \quad (11)$$

$$\mathbf{K}(n) = \Gamma^{-1}(n-1)\mathbf{X}(n)\left[\lambda^2\Lambda(n) + \mathbf{X}^T(n)\Gamma^{-1}(n-1)\mathbf{X}(n)\right]^{-1} \quad (12)$$

$$\Gamma^{-1}(n) = \frac{1}{\lambda^2}\left[\Gamma^{-1}(n-1) - \mathbf{K}(n)\mathbf{X}^T(n)\Gamma^{-1}(n-1)\right] \quad (13)$$

$$\alpha(n) = \begin{pmatrix} \tilde{\mathbf{x}}^T(n)\mathbf{r}(n-1) \\ d(n) \end{pmatrix} - \mathbf{X}(n)\mathbf{w}(n-1) \quad (14)$$

$$\mathbf{r}(n) = \lambda\mathbf{r}(n-1) + \tilde{\mathbf{x}}(n)d(n) \quad (15)$$

$$\mathbf{R}(n) = \lambda\mathbf{R}(n-1) + \tilde{\mathbf{x}}(n)\mathbf{x}^T(n) \quad (16)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{K}(n)\alpha(n) \quad (17)$$

where the forgetting factor $0 < \lambda \le 1$ has been introduced for nonstationary statistics.

C. *The generalized least mean square (GLMS) algorithm*

The GLMS algorithm is an instrumental variable-based LMS algorithm and is robust versus additif noise. The algorithm is defined as follows [11]:

- *Initial conditions*:

$$\mathbf{w}(0) = \mathbf{0} \tag{18}$$

- *Recursive process*:

$$e(n) = d(n) - \mathbf{w}^T(n-1)\mathbf{x}(n) \tag{19}$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu\tilde{\mathbf{x}}(n)e(n) \tag{20}$$

where μ is the step size.

D. *The double GLMS (DGLMS) and the double ORIV (DORIV) algorithms*

The DGLMS algorithm is defined as follows:

- *Initialization*:

$$\mathbf{w}_1(0) = \mathbf{0}_{N_1+1} \tag{21}$$

$$\mathbf{w}_2(0) = \mathbf{0}_{N_2+1} \tag{22}$$

- *Estimation*:

$$tp_1(n) = p_1(n) - \sum_{k=1}^{N_2} w_{2k}(n-1)s_2(n-k) \tag{23}$$

$$tp_2(n) = p_2(n) - \sum_{k=1}^{N_2} w_{1k}(n-1)s_1(n-k) \tag{24}$$

$$s_1(n) = \frac{tp_1(n) - w_{20}(n-1)tp_1(n)}{1 - w_{10}(n-1)w_{20}(n-1)} \tag{25}$$

$$s_2(n) = \frac{tp_2(n) - w_{10}(n-1)tp_2(n)}{1 - w_{10}(n-1)w_{20}(n-1)} \tag{26}$$

- *Filters update*:

$$\mathbf{w}_1(n) = \mathbf{w}_1(n-1) + \mu_1\tilde{\mathbf{s}}_1(n)s_2(n) \tag{27}$$

$$\mathbf{w}_2(n) = \mathbf{w}_2(n-1) + \mu_2\tilde{\mathbf{s}}_2(n)s_1(n) \tag{28}$$

The DORIV algorithm is defined as follows:

- *Initialization*:

$$\begin{aligned}\mathbf{R}_1^T(0) &= \beta_1[\mathbf{I}_{N_1xN_1} \,|\, \mathbf{0}_{N_1xL-N_1}],\ \Gamma_1^{-1}(0) = \frac{1}{\beta_1^2}\mathbf{I}_{N_1xN_1} \\ \mathbf{R}_2^T(0) &= \beta_2[\mathbf{I}_{N_2xN_2} \,|\, \mathbf{0}_{N_2xL-N_2}],\ \Gamma_2^{-1}(0) = \frac{1}{\beta_2^2}\mathbf{I}_{N_2xN_2}\end{aligned} \tag{29}$$

$$\mathbf{w}_1(0) = \mathbf{0},\ \mathbf{r}_1(0) = \mathbf{0},\ \mathbf{w}_2(0) = \mathbf{0},\ \mathbf{r}_2(0) = \mathbf{0} \tag{30}$$

- *Estimation*:

$$tp_1(n) = p_1(n) - \sum_{k=1}^{N_2} w_{2k}(n-1)s_2(n-k) \tag{31}$$

$$tp_2(n) = p_2(n) - \sum_{k=1}^{N_2} w_{1k}(n-1)s_1(n-k) \tag{32}$$

$$s_1(n) = \frac{tp_1(n) - w_{20}(n-1)tp_1(n)}{1 - w_{10}(n-1)w_{20}(n-1)} \tag{33}$$

$$s_2(n) = \frac{tp_2(n) - w_{10}(n-1)tp_2(n)}{1 - w_{10}(n-1)w_{20}(n-1)} \tag{34}$$

$$\mathbf{S}_1(n) = \left[\mathbf{R}_1^T(n-1)\tilde{\mathbf{s}}_1(n) \quad \mathbf{s}_1(n)\right] \tag{35}$$

$$\mathbf{S}_2(n) = \left[\mathbf{R}_2^T(n-1)\tilde{\mathbf{s}}_2(n) \quad \mathbf{s}_2(n)\right] \tag{36}$$

$$\lambda_1^2\Lambda_1(n) = \begin{pmatrix} -\tilde{\mathbf{s}}_1^T(n)\tilde{\mathbf{s}}_1(n) & \lambda_1 \\ \lambda_1 & 0 \end{pmatrix} \tag{37}$$

$$\lambda_2^2\Lambda_2(n) = \begin{pmatrix} -\tilde{\mathbf{s}}_2^T(n)\tilde{\mathbf{s}}_2(n) & \lambda_2 \\ \lambda_2 & 0 \end{pmatrix} \tag{38}$$

$$\begin{aligned}\mathbf{K}_1(n) = {} & \Gamma_1^{-1}(n-1)\mathbf{S}_1(n)\times \\ & \left[\lambda_1^2\Lambda_1(n) + \mathbf{S}_1^T(n)\Gamma_1^{-1}(n-1)\mathbf{S}_1(n)\right]^{-1}\end{aligned} \tag{39}$$

$$\begin{aligned}\mathbf{K}_2(n) = {} & \Gamma_2^{-1}(n-1)\mathbf{S}_2(n)\times \\ & \left[\lambda_2^2\Lambda_2(n) + \mathbf{S}_2^T(n)\Gamma_2^{-1}(n-1)\mathbf{S}_2(n)\right]^{-1}\end{aligned} \tag{40}$$

$$\Gamma_1^{-1}(n) = \frac{1}{\lambda_1^2}\left[\Gamma_1^{-1}(n-1) - \mathbf{K}_1(n)\mathbf{S}_1^T(n)\Gamma_1^{-1}(n-1)\right] \tag{41}$$

$$\Gamma_2^{-1}(n) = \frac{1}{\lambda_2^2}\left[\Gamma_2^{-1}(n-1) - \mathbf{K}_2(n)\mathbf{S}_2^T(n)\Gamma_2^{-1}(n-1)\right] \tag{42}$$

$$\boldsymbol{\alpha}_1(n) = \begin{pmatrix} \tilde{\mathbf{s}}_1^T(n)\mathbf{r}_1(n-1) \\ p_2(n) \end{pmatrix} - \mathbf{S}_1(n)\mathbf{w}_1(n-1) \tag{43}$$

$$\boldsymbol{\alpha}_2(n) = \begin{pmatrix} \tilde{\mathbf{s}}_2^T(n)\mathbf{r}_2(n-1) \\ p_1(n) \end{pmatrix} - \mathbf{S}_2(n)\mathbf{w}_2(n-1) \tag{44}$$

$$\mathbf{r}_1(n) = \lambda_1\mathbf{r}_1(n-1) + \tilde{\mathbf{s}}_1(n)p_2(n) \tag{45}$$

$$\mathbf{r}_2(n) = \lambda_2\mathbf{r}_2(n-1) + \tilde{\mathbf{s}}_2(n)p_1(n) \tag{46}$$

$$\mathbf{R}_1(n) = \lambda_1\mathbf{R}_1(n-1) + \tilde{\mathbf{s}}_1(n)\mathbf{s}_1^T(n) \tag{47}$$

$$\mathbf{R}_2(n) = \lambda_2\mathbf{R}_2(n-1) + \tilde{\mathbf{s}}_2(n)\mathbf{s}_2^T(n) \tag{48}$$

- *Filters update*:

$$\mathbf{w}_1(n) = \mathbf{w}_1(n-1) + \mathbf{K}_1(n)\boldsymbol{\alpha}_1(n) \tag{49}$$

$$\mathbf{w}_2(n) = \mathbf{w}_2(n-1) + \mathbf{K}_2(n)\boldsymbol{\alpha}_2(n) \tag{50}$$

III. SIMULATION RESULTS

The noise has been separately recorded in a car moving in five different conditions, the microphone is placed in front of the driver and the noises have been artificially added to the noise-free speech so that one would master the SNR input. The coupling systems are 10 taps two FIR filters [8].

An example of one signal captured by the first microphone $p_1(n)$ and another by the second microphone $p_2(n)$ is respectively shown in Fig. 3 and Fig. 4. In this case the SNR of $p_1(n)$ and $p_2(n)$ are respectively 3.09 dB and 3.79 dB.

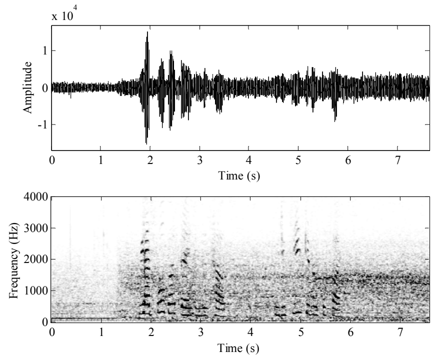

Fig. 3. The signal $p_1(n)$ captured by the first microphone and its spectrogram (SNR = 3.09 dB)

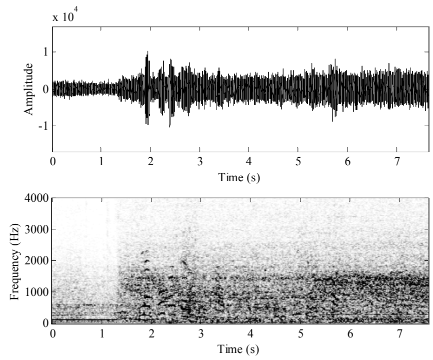

Fig. 4: The signal $p_2(n)$ captured by the second microphone and its spectrogram (SNR = 3.79 dB)

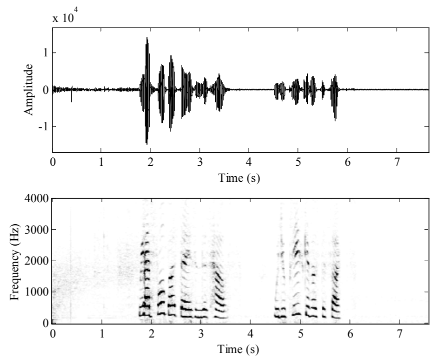

Fig. 5. Enhanced speech $s_1(n)$ obtained with the noise canceller system based DGLMS algorithm and its spectrogram (SNR = 16.29 dB)

The output signal of the noise canceller system using the DGLMS algorithm is shown in Fig. 5.

A comparative SNR output gain between the Extended LMS [5], the Double LMS [4], the DGLMS and the DORIV algorithms is provided in table 1. This table shows the superiority of the noise canceller instrumental variable DGLMS and DORIV based algorithms. This global performance behaviour is confirmed also by the frame by frame SNR output. The ELMS algorithms take more time before handling the noise field after which its segmental SNR behaviour is close to the segmental behaviour of the DGLMS algorithm.

Informal quality and intelligibility tests indicate also significant superiority of such algorithms to enhance speech signal.

Table 1: The SNR gain of $s_1(n)$ for different algorithms

Case	Input SNR(dB) $p_1(n)$	Gain SNR (dB) $s_1(n)$			
		ELMS [5]	DLMS [4]	DGLMS	DORIV
1	10.48	9.10	9.77	10.28	12.17
2	2.45	13.48	13.22	14.57	22.38
3	11.40	8.36	7.93	9.76	16.19
4	9.97	11.15	12.35	13.81	21.14
5	3.09	12.06	14.74	16.29	23.64

REFERENCES

[1] Widrow B. and al, “Adaptive Noise Canceling: Principles and Applications,” Proceedings of IEEE, vol. 63, no. 3, pp. 1692-1716, Dec. 1975.

[2] Compernolle D.V. and Gerven S.V. “Blind Separation of Sources: Signal Separation in a Symmetric Adaptive Noise Canceller by Output Decorrelation”, Proc. of ICASSP 1992, vol. IV, pp. 221-224.

[3] Gerven S. V. and Compernolle D. V. “Signal Separation by Symmetric Adaptive Decorrelation: Stability, Convergence, and Uniqueness”. *IEEE Transactions on Signal Processing*, vol. 43, no. 7, pp. 1602-1612, July 1995.

[4] Vallauri A. “L’étude et le développement de méthodes de reconnaissance de la parole et de réduction du bruit, et application”. Ph.D. thesis, Univ. Nice, France, December 1992 (in French).

[5] Weinstein E., Feder M. and Oppenheim A.V. “Multi-Channel Signal Separation by Decorrelation,” *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 4, pp. 405-413, October 1993.

[6] Gustafsson T., Lindgren U. and Sahlin H. “Statistical Analysis of a Signal Separation Method Based on Second-Order Statistics,” *IEEE Transactions on Signal Processing*, vol. 49, no.2, pp. 441-444 , February 2001.

[7] Gabrea M., “Rehaussement de la parole en ambience bruitée. Méthodes monovoie et bivoie”. Ph.D. thesis, University of Bordeaux 1, France (in French).

[8] Gabrea M., “Double affine projection algorithm-based speech enhancement algorithm”, Proc. of ICASSP 2003, vol. 1, pp. 904 - 907.

[9] Therrien C. W. *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, New Jersey 1992.

[10]Friedlander B. “The overdetermined recursive instrumental variable method”, *IEEE Transactions on Automatic Control*, vol. 29, no. 4, pp. 353 – 356.

[11]Aboutajdine D., Adib A. and Meziane A, “Fast adaptive algorithms for AR parameters estimation using higher order statistics” *IEEE Transactions on Signal Processing,* vol. 44, no. 8, pp. 1998 – 2009.