ADAPTIVE FILTERING ALGORITHMS FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION

J. BENESTY *, F. AMAND **, A. GILLOIRE **, Y. GRENIER *

* TELECOM PARIS Département Signal 46 rue Barrault 75634 PARIS CEDEX 13 - FRANCE
** CNET LAA/TSS/CMC BP 40 22301 LANNION CEDEX - FRANCE
E-mail : benesty@sig.enst.fr

ABSTRACT

It is likely that stereophonic (and more generally, multi-channel) sound pick-up, transmission and diffusion will be implemented in future teleconference systems to provide the users with enhanced quality. Therefore, adequate solutions must be found to solve the problem of stereophonic acoustic echo which will occur in such systems. We explain in this paper the difference between the mono and two-channel systems and the behavior of the two-channel classical adaptive algorithms in comparison with the same algorithms in the mono-channel case. Also, we outline a new NLMS-like algorithm derived from the two-channel RLS algorithm as a first member of a family of improved two-channel adaptive filters.

1. INTRODUCTION

Acoustic echo cancellers are necessary for communication systems such as teleconferencing in order to reduce echoes which impair the quality of communications.
Conceptually, stereophonic acoustic echo cancellation can be viewed as a straightforward generalization of the usual single channel acousic echo cancellation principle to the two channel case [1,2,3]. A basic scheme for stereophonic acoustic echo cancellation is sketched in figure 1. Note that only one microphone path is shown because the arrangement is fully symmetrical with respect to the two microphone signals in the local room. Clearly, according to this scheme, stereophonic acoustic echo cancellation fits within the framework of direct identification of a multi-input, unknown linear system, this latter one being formed by the parallel combination of the 2 acoustic paths (W_1, W_2) extending through the local room from the loudspeakers (HP1, HP2) to the microphone (M). The stereophonic acoustic echo canceller tries to modelize this unknown system by a pair of adaptive filters (H_1, H_2). The same model applies to the other microphone with the acoustic paths replaced by the ones appropriate to that microphone. Also, a similar canceller system would apply to the distant room.
In the following, we compare the mono-channel to the two-channel systems.

2. THE STEREOPHONIC ACOUSTIC ECHO CANCELLATION PROBLEM

We assume that the system (distant room) is stationary, linear and time invariant; we have the following relation [4]:

$$X_1^t(n)G_2 = X_2^t(n)G_1 \tag{1}$$

where G_1 and G_2 stand for the impulse responses of the source-to-microphone acoustic paths in the remote room as indicated in figure 1, and $X_1(n)$ and $X_2(n)$ stand for vectors of signal samples at the microphones outputs in the same room. $.^t$ denotes the transpose of a vector or a matrix and :

$$X_1(n) = [x1(n) \quad x1(n-1) \quad \cdots \quad x1(n-M+1)]^t$$
$$X_2(n) = [x2(n) \quad x2(n-1) \quad \cdots \quad x2(n-M+1)]^t$$
$$G_1 = [g_{11} \quad g_{12} \quad \cdots \quad g_{1M}]^t$$
$$G_2 = [g_{21} \quad g_{22} \quad \cdots \quad g_{2M}]^t$$

M being the size of the impulse responses.

Minimization of the following recursive least squares criterion (see figure 1 for notations) :

$$J(n) = \sum_{p=1}^{n} w^{n-p}\left[y(p) - H_1^t(n)X_1(p) - H_2^t(n)X_2(p)\right]^2 \tag{2}$$

leads to the equation :

$$R(n)\begin{bmatrix} H_1(n) \\ H_2(n) \end{bmatrix} = r(n) \tag{3}$$

where w $(0 < w \le 1)$ is the exponential forgetting factor, $R(n)$ the correlation matrix :

$$R(n) = \sum_{p=1}^{n} w^{n-p}\begin{bmatrix} X_1(p) \\ X_2(p) \end{bmatrix}\left[X_1^t(p) \quad X_2^t(p)\right] = \begin{bmatrix} R_{x1x1}(n) & R_{x1x2}(n) \\ R_{x2x1}(n) & R_{x2x2}(n) \end{bmatrix} \tag{4}$$

and $r(n)$ the correlation vector between the input signals and the output signal in the local room :

$$r(n) = \sum_{p=1}^{n} w^{n-p}\, y(p)\begin{bmatrix} X_1(p) \\ X_2(p) \end{bmatrix} \tag{5}$$

Our aim is to obtain the optimum filters from eq. (3). Now consider the vector :

$$U = \begin{bmatrix} G_2 \\ -G_1 \end{bmatrix} \tag{6}$$

it can be readily verified by using eq. (1) that :

$$R(n)U = 0 \tag{7}$$

0-7803-2431-5/95 $4.00 © 1995 IEEE

which means that the matrix $R(n)$ is not invertible. Therefore, there is no unique solution to the problem of minimizing (2), and the adaptive algorithm drives to any one of the possible solutions, which can be very different from the "true" expected solution $H_1 = W_1$, $H_2 = W_2$.

However, in practical situations there are - at least - two reasons that make this matrix invertible :

a) The signals $x1$ and $x2$ at the outputs of the distant room contain noisy components that are uncorrelated ;

b) The filters (H_1, H_2) that modelize the impulse responses of the local room are of finite length, so the size of X_1 and X_2 is much smaller than the length of G_1 and G_2, and the relation (1) is not satisfied.

Hence the matrix $R(n)$ becomes invertible (but it is ill-conditioned because the two input signals are strongly correlated) and the true solution $H_1 = W_1$, $H_2 = W_2$ can be found accordingly.

The main difference between the mono-channel case and the two-channel case lies in the eigenvalue spread $C(R)$: in the first case, $C(R)$ depends only on the nature of the input signal, whereas in the second case $C(R)$ is considerably increased due to the correlation between the two input signals whatever the input signals may be. This is a crucial point because the convergence rate of many adaptive algorithms depends on $C(R)$.

3. THE RECURSIVE LEAST SQUARES APPROACH

From eq. (3) and equations (4), (5) at time $n+1$ we can derive easily the two-channel RLS algorithm. The development of fast algorithms relies on certain update relations that arise in the characterization of adaptive forward and backward linear predictions that are optimized in the least-squares problem. The other update relations needed for the development of fast algorithms involve the Kalman gain vector. In this section we give a fast version of the two-channel RLS algorithm [5], in which we have implemented a numerical stabilization technique derived from [6,7] (this version is more simple than the one proposed in [7]) :

(8)

$$a)\ e_a(n+1) = \chi(n+1) - A'(n)X(n) \quad (2\times 1)$$

$$b)\ \alpha_1(n+1) = \alpha(n) + e_a^t(n+1)E_a^{-1}(n)e_a(n+1) \quad (1\times 1)$$

$$c)\ G_1'(n+1) = \begin{bmatrix} 0 \\ 0 \\ G'(n) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \multicolumn{2}{c}{-A(n)} \end{bmatrix} E_a^{-1}(n)e_a(n+1)$$

$$= \begin{bmatrix} M(n+1) \\ m(n+1) \end{bmatrix} \quad ((2L+2)\times 1)$$

$$d)\ A(n+1) = A(n) + \frac{G'(n)e_a^t(n+1)}{\alpha(n)} \quad (2L\times 2)$$

$$e)\ G'(n+1) = M(n+1) + B(n)m(n+1) \quad (2L\times 1)$$

$$f)\ e_b^1(n+1) = E_b(n)m(n+1) \quad (2\times 1)$$

$$g)\ e_b^2(n+1) = \chi(n+1-L) - B'(n)X(n+1) \quad (2\times 1)$$

$$h)\ e_b(n+1) = k\,e_b^2(n+1) + (1-k)e_b^1(n+1) \quad (2\times 1)$$

$$i)\ E_a(n+1) = (E_a(n) + \frac{e_a(n+1)e_a^t(n+1)}{\alpha(n)})w \quad (2\times 2)$$

$$j)\ \alpha(n+1) = \alpha_1(n+1) - (e_b^2(n+1))^t\, m(n+1) \quad (1\times 1)$$

$$k)\ E_b(n+1) = (E_b(n) + \frac{e_b^2(n+1)(e_b^2(n+1))^t}{\alpha(n+1)})w \quad (2\times 2)$$

$$l)\ B(n+1) = B(n) + \frac{G'(n+1)e_b^t(n+1)}{\alpha(n+1)} \quad (2L\times 2)$$

$$m)\ e(n+1) = y(n+1) - H'(n)X(n+1) \quad (1\times 1)$$

$$n)\ H(n+1) = H(n) + \frac{G'(n+1)e(n+1)}{\alpha(n+1)} \quad (2L\times 1)$$

where :

$$\chi'(n) = [x1(n) \quad x2(n)]$$

$$X'(n) = [\chi'(n) \quad \chi'(n-1) \quad \cdots \quad \chi'(n-L+1)]$$

$$H'(n) = [h_{11}(n) \quad h_{21}(n) \quad \cdots \quad h_{1L}(n) \quad h_{2L}(n)]$$

and the sizes of the vectors and matrices are indicated after each equation; L is the length of the filters H_1 and H_2, and k is the feedback constant for numerical stabilization : $1.5 \le k \le 2.5$. We have implemented in this algorithm additional controls which are necessary for continuous operation with non-stationary signals like speech [6,8]. This recursive least-squares solution gives good performance but it is very expensive in terms of number of operations : $28L$ multiplications and $28L$ additions (instead of $8L$ muliplications and $8L$ additions in the mono-channel case).

Comparison with the mono-channel RLS algorithm : unlike the LMS algorithm, the rate of convergence of the RLS is essentially insensitive to variations in the eigenvalue spread of the matrix $R(n)$. This property is preserved in the two-channel case for which the RLS algorithm has almost the same convergence as the mono-channel RLS algorithm whatever the input signals may be. On the contrary, the behavior of the LMS algorithm is severely degraded in the two-channel case as shown in the following section.

4. THE STOCHASTIC GRADIENT APPROACH

The two-channel LMS algorithm is given by :

$$e(n+1) = y(n+1) - \sum_{i=1}^{2} X_i^t(n+1)H_i(n) \qquad (9)$$

$$\begin{bmatrix} H_1(n+1) \\ H_2(n+1) \end{bmatrix} = \begin{bmatrix} H_1(n) \\ H_2(n) \end{bmatrix} + \mu \begin{bmatrix} X_1(n+1) \\ X_2(n+1) \end{bmatrix} e(n+1) \qquad (10)$$

where μ is a parameter that controls stability and rate of convergence. We can easily show that the stability condition is :

$$0 < \mu < \frac{2}{L(\sigma_{x1}^2 + \sigma_{x2}^2)} \qquad (11)$$

where σ_{x1}^2 and σ_{x2}^2 are the powers of the input signals. When this condition is satisfied, the weight vector converges to the optimal solution of Wiener-Hopf.
For nonstationary input signals, we use the two-channel normalized LMS where μ is replaced by :

$$\mu(n+1) = \frac{\alpha}{X_1^t(n+1)X_1(n+1) + X_2^t(n+1)X_2(n+1)} \quad (12)$$

with $0 < \alpha < 2$.
The arithmetic complexity of the two-channel LMS algorithm is $4L$ multiplications and $4L$ additions (instead of $2L$ multiplications and $2L$ additions for the mono-channel LMS algorithm).
A standard analysis in the mean shows that the convergence rate of this algorithm depends on the eigenvalue spread of the correlation matrix :

$$E\left\{\begin{bmatrix} X_1(n) \\ X_2(n) \end{bmatrix}\begin{bmatrix} X_1^t(n) & X_2^t(n) \end{bmatrix}\right\}$$

Comparison with the mono-channel LMS algorithm : in the mono-channel case, when the autocorrelation matrix is ill-conditioned (this is the case of speech input signal), the LMS converges very slowly and has almost the same convergence rate as the two-channel LMS algorithm. On the other hand, when the autocorrelation matrix is well-conditioned (USASI input signal for instance) the LMS has a good convergence rate but the two-channel LMS converges still very slowly.
This behavior is due to the fact that in the two-channel case the correlation matrix is generally ill-conditioned. In other words, the two-channel LMS will converge very slowly because it doesn't take into account the cross-correlation between the two input signals.

5. SIMULATIONS

This section compares by simulation the previous algorithms. The impulse responses (W_1, W_2) to be identified are truncated to 256 points. They were obtained from measurements in a real teleconference room like the two input signals. The length of the filters (H_1, H_2) is $L = 256$. All plots show the mean-squared modeling error versus the number of iterations. A white noise is added to the output (SNR=60 dB). Figure 2 shows the behavior of the two-channel fast RLS with USASI as input signals, we have verified that the filters converge to the true solution. Figure 3 shows the convergence curve of the same algorithm with speech as input signals. Figures 4 and 5 compare the mono and the two channel LMS algorithms with USASI and speech as input signals. We can point out that, as expected, there is a great difference between the mono-channel and two-channel LMS algorithms with USASI signals, which is not the case with speech signals.
The mono-channel performance was obtained by adapting separately each of the two filters (H_1, H_2) using the corresponding error then adding the two errors; thus, the total number of adapted coefficients is $2L$ like in the two-channel case.

6. CONCLUSION AND PROSPECTS

A very important conclusion of this study is that a two-channel adaptive filtering algorithm should take into account the cross-correlation between the two input signals to have a good convergence rate. A direct approximation of the two-channel RLS algorithm gives a first solution : The "Extended LMS" (ELMS) algorithm, which is :

$$\begin{bmatrix} H_1(n+1) \\ H_2(n+1) \end{bmatrix} = \begin{bmatrix} H_1(n) \\ H_2(n) \end{bmatrix} + \alpha_E M^{-1}(n+1)\begin{bmatrix} X_1(n+1) \\ X_2(n+1) \end{bmatrix} e(n+1) \quad (13)$$

where :

$$M(n+1) = \begin{bmatrix} p_{11}(n+1)I & \rho r_{12}(n+1)I \\ \rho r_{12}(n+1)I & p_{22}(n+1)I \end{bmatrix} \quad (14)$$

and :

$$p_{11}(n+1) = X_1^t(n+1)X_1(n+1)$$
$$p_{22}(n+1) = X_2^t(n+1)X_2(n+1)$$
$$r_{12}(n+1) = X_1^t(n+1)X_2(n+1)$$

The stability conditions of the proposed algorithm are :

$$0 < \alpha_E < 1$$
$$0 \le \rho < 1 \quad (15)$$

The ELMS has a better behavior than the two-channel LMS algorithm and its arithmetic complexity is $6L$ multiplications and $6L$ additions. This algorithm can be considered as a simple member of a general family of the two-channel adaptive filters that take into account cross-correlation of the input signals in a variable amount.

ACKNOWLEDGMENT

The authors would like to thank Pr P. Duhamel for useful discussions on the subject of this paper.

REFERENCES

[1] M. M. Sondhi, D. R. Morgan, "Acoustic Echo Cancellation for Stereophonic Teleconferencing", IEEE Workshop on Audio and Acoustics, Mohonk Mountain House, NY, 1991.
[2] A. Hirano, A. Sugiyama, "Convergence Characteristics of a Multi-Channel Echo Canceller with Strongly Cross-Correlated Input Signals", Proc. of 6th DSP symposium, Nov. 1991.
[3] Y. Mahieux, A. Gilloire, F. Khalil, "Annulation d'Echo en Téléconférence Stéréophonique", GRETSI 1993, FRANCE.
[4] J. Benesty, Y. Grenier, "Annulation d'Echo Acoustique Multi-voies", Telecom Paris, 1994.
[5] M. Bellanger, "Adaptive Digital Filters and Signal Analysis", M. Dekker Inc., NY, 1987.
[6] A. Benallal, "Etude des Algorithmes des Moindres Carrés Transversaux Rapides", Ph. D. Thesis, FRANCE, 1989.
[7] D. T. M. Slock, L. Chisci, H. Lev-Ari, T. Kailath, "Modular and Numerically Stable Fast Transversal Filters for Multichannel and Multiexperiment RLS", IEEE Trans. on SP, April 1992.
[8] J. Benesty, Y. Grenier, "Algorithmes Récursifs pour l'Annulation d'Echo Acoustique Stéréophonique", Telecom Paris, 1994.

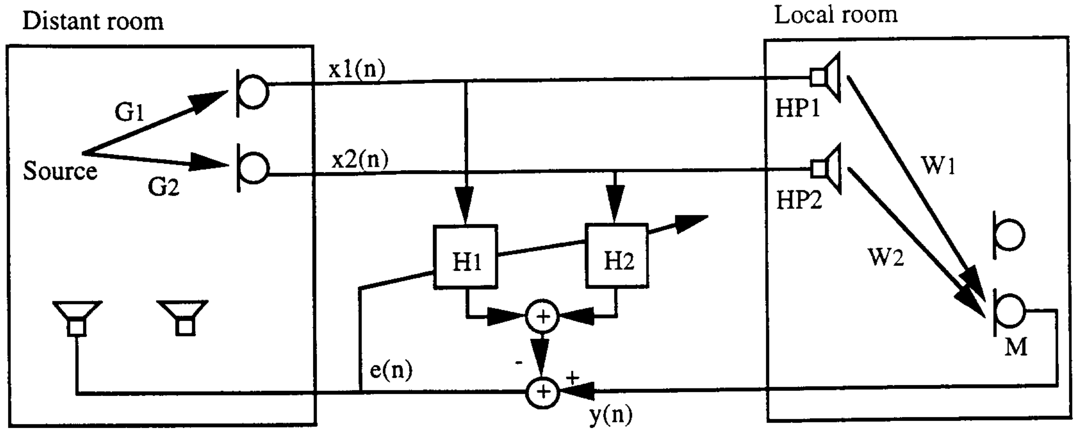

Figure 1 : Basic scheme for stereophonic acoustic echo cancellation

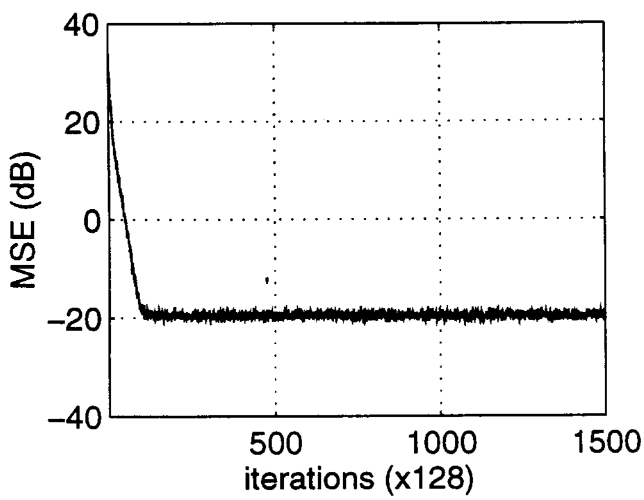

Fig. 2 : Convergence behavior of the two-channel FRLS algorithm with USASI as input signals

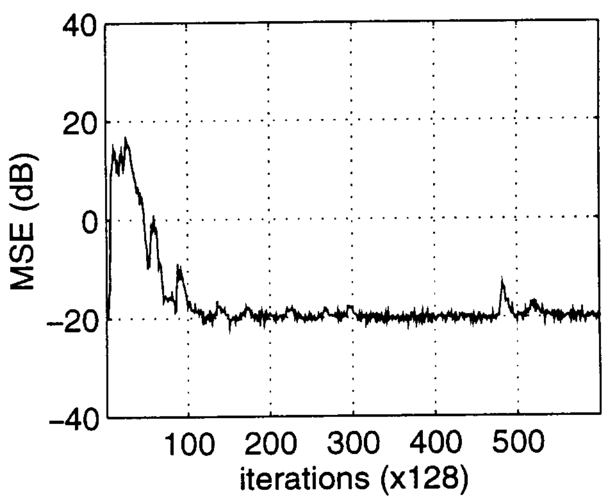

Fig. 3 : Convergence behavior of the two-channel FRLS algorithm with speech as input signals

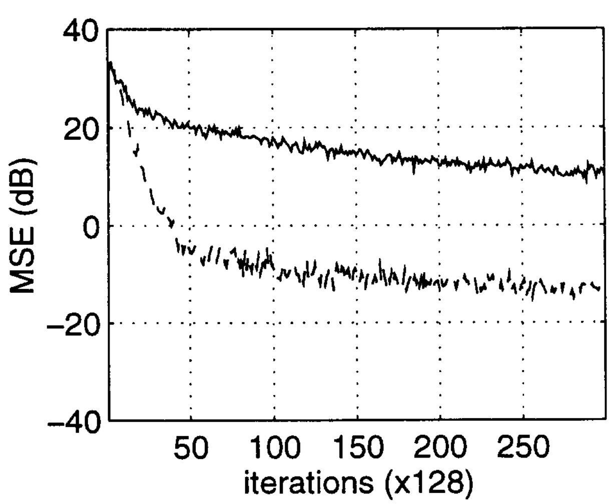

Fig. 4 : Comparison of the mono (--) and two-channel (-) LMS algorithms with USASI as input signals

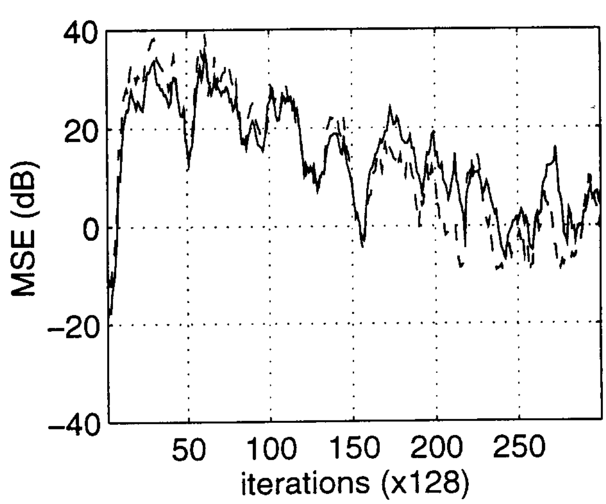

Fig. 5 : Comparison of the mono (--) and two-channel (-) LMS algorithms with speech as input signals