Multi-Channel Adaptive Filtering Applied to Multi-Channel Acoustic Echo Cancellation

Jacob Benesty[1], **Pierre Duhamel**[2], **Yves Grenier**[2]
[1]Lucent Technologies, Bell Labs Innovations, New Jersey, USA
[2]ENST Paris, Dept. Signal, France

ABSTRACT

This paper presents some new ways of deriving multi-channel (M-C) adaptive algorithms in the context of M-C acoustic echo cancellation (AEC).

1 Introduction

Multi-Channel (M-C) sound pick-up, transmission and diffusion is likely to be implemented in future teleconference systems to provide the users with enhanced quality. However, the echoes in such systems are even more important than in classical hand-free communication, so that acoustic echo cancellers are necessary.

This new topic requires the use of M-C adaptive algorithms which are not yet well understood in the context of M-C AEC even with the M-C RLS algorithm [1]. Few researchers have worked on the problems encountered in adaptive filtering M-C AEC and hardly nothing has been written on this subject. Nevertheless, we can point out the works of [2, 3, 4, 5].

M-C AEC can be viewed as a straighforward generalization of the usual single channel AEC [2, 3]. The M-C echo canceller consists of P^2 adaptive filters aiming at identifying P^2 echo paths from loudspeakers to microphones (in a P channel case). Each adaptive filter estimates the impulse response of the corresponding echo path.

In our application, the microphone signals in the distant room come from the same source, this gives rise to some identification problems of the impulse responses of the local room. Section 2 describes the origin of this problem.

2 Multi-channel identification. Specificities of M-C AEC

2.1 Ideal case

We assume that the system (echo path in the distant room) is stationary, linear and time invariant; since the source is unique and by writing the convolutions resulting from this unique source and each impulse response, we obtain the $(P(P-1)/2)$ following relations:

$$X_i^T(n)G_j = X_j^T(n)G_i \tag{1}$$

$i, j = 1, 2, \cdots, P$; $i \neq j$; where G_i stand for the impulse responses of the source-to-microphone acoustic paths in the remote room and $X_i(n)$ stand for vectors of signal samples at the microphone outputs in the same room. Let $.^T$ denote the transposition operator and M the size of the impulse responses, we have:

$$\begin{aligned} X_i(n) &= [x_i(n),\ x_i(n-1), \cdots, x_i(n-M+1)]^T \quad (2) \\ G_i &= [g_{1,i},\ g_{2,i}, \cdots, g_{M,i}]^T \quad (3) \end{aligned}$$

Minimization of the following recursive least squares criterion (4) leads to the normal equation (5):

$$\begin{aligned} J(n) &= \sum_{l=1}^{n} w^{n-l}[y(l) - H^T(n)X(l)]^2 \quad (4) \\ R(n)H(n) &= r(n) \quad (5) \end{aligned}$$

where w, $(0 < w < 1)$ is the exponential forgetting factor, $y(n)$ is the echo at time n, $H_i(n)$ the filters at time n with $H(n) = [H_1^T(n), H_2^T(n), \cdots, H_P^T(n)]^T$ and $X(n) = [X_1^T(n), X_2^T(n), \cdots, X_P^T(n)]^T$. $R(n)$ is the covariance matrix which can be decomposed in a block matrix, using the cross and auto correlation matrices $R_{x_ix_j}(n)$:

$$\begin{aligned} R(n) &= \sum_{l=1}^{n} w^{n-l} X(l)X^T(l) \quad (6) \\ &= \begin{pmatrix} R_{x_1x_1}(n) & R_{x_1x_2}(n) & \cdots & R_{x_1x_p}(n) \\ R_{x_2x_1}(n) & R_{x_2x_2}(n) & \cdots & R_{x_2x_p}(n) \\ \vdots & \vdots & & \vdots \\ R_{x_px_1}(n) & R_{x_px_2}(n) & \cdots & R_{x_px_p}(n) \end{pmatrix} \end{aligned}$$

with $R_{x_ix_j}(n) = \sum_{l=1}^{n} w^{n-l} X_i(l)X_j^T(l)$ and $r(n)$ is the correlation vector between the input signals and the output signal in the local room.

Our aim is to derive the optimum filters from (4). Now consider the vector:

$$U = [\sum_{i=2}^{P} G_i^T,\ -G_1^T, \cdots, -G_1^T]^T \tag{7}$$

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

$$R(n)U = 0 \tag{8}$$

which means that he matrix $R(n)$ is not full-rank. Therefore, there is no unique solution to the problem of minimizing (4), and the adaptive algorithm drives to any one of the possible solutions, which can be very different from the "true" expected solution $H_i(n) = W_i$ $(i = 1, 2, \cdots, P)$ where W_i are the impulse responses of the loudspeakers-to-microphone acoustic paths in the local room.

2.2 Real case

In practical situations there are - at least - two reasons that make the matrix $R(n)$ invertible :

- The signals $x_i(n)$ $(i = 1, 2, \cdots, P)$, measured at the microphones of the distant room contain noisy components that are uncorrelated; hence (1) does not exactly apply to the actual signals.
- The filters $H_i(n)$ that modelize the impulse responses of the local room are of finite length which is much smaller than the actual length of G_i, and (1) is not fully satisfied.

Matrix $R(n)$ would have a number of zero eigenvalues if (1) would hold. However, since this equation is only approximately met, these eigenvalues are not null, but very small. In other words, the input signals are strongly correlated. As a result, the adaptive filters can converge to the actual solution, but with a number of difficulties due to this very strong correlation. Thus, a good M-C adaptive algorithm (in terms of convergence rate and tracking abilities) should take into account the cross-correlation between the input signals. In the following, we assume that matrix $R(n)$ is always full-rank.

3 The M-C RLS algorithm

3.1 Classical recursions

From eq. (4) and equations (5), (6) at time n we can easily derive the M-C RLS algorithm. The adaptive filter equations are :

$$\begin{aligned} e(n) &= y(n) - H^T(n-1)X(n) && (9)\\ H(n) &= H(n-1) + R^{-1}(n)X(n)e(n) && (10) \end{aligned}$$

where $e(n)$ is the modelling error, H is the filter coeffients vector of size LP, L is the length of the filters H_i, $(i = 1, 2, \cdots, P)$: $H_i = [h_{1,i},\ h_{2,i}, \cdots, h_{L,i}]^T$. Of course, fast and stable versions of this algorithm can be derived but it is not the purpose of this paper.

3.2 Factorized recursions for $P = 2$

Another way to write the two-channel RLS algorithm is to first factorize the covariance matrix inverse. We get the factorization:

$$\begin{aligned} R^{-1}(n) &= \begin{bmatrix} R_1^{-1}(n) & O_{L\times L} \\ O_{L\times L} & R_2^{-1}(n) \end{bmatrix} \\ &\times \begin{bmatrix} I_{L\times L} & -R_{x_1x_2}(n)R^{-1}_{x_2x_2}(n)] \\ -R_{x_2x_1}(n)R^{-1}_{x_1x_1}(n) & I_{L\times L} \end{bmatrix} \end{aligned} \quad (11)$$

with :

$$\begin{aligned} R_1(n) &= R_{x_1x_1}(n) - R_{x_1x_2}(n)R^{-1}_{x_2x_2}(n)R_{x_2x_1}(n) \\ R_2(n) &= R_{x_2x_2}(n) - R_{x_2x_1}(n)R^{-1}_{x_1x_1}(n)R_{x_1x_2}(n) \end{aligned}$$

Using (11), we may rewrite the two-channel RLS algorithm as follows :

$$\begin{aligned} H_1(n) &= H_1(n-1) + R_1^{-1}(n)Z_1(n)e(n) && (12)\\ H_2(n) &= H_2(n-1) + R_2^{-1}(n)Z_2(n)e(n) && (13) \end{aligned}$$

where :

$$\begin{aligned} Z_1(n) &= X_1(n) - R_{x_1x_2}(n)R^{-1}_{x_2x_2}(n)X_2(n) && (14)\\ Z_2(n) &= X_2(n) - R_{x_2x_1}(n)R^{-1}_{x_1x_1}(n)X_1(n) && (15) \end{aligned}$$

Now a slight rearrangement of (5) for $P = 2$ yields :

$$\begin{aligned} H_1(n) &= R^{-1}_{x_1x_1}(n)r_{yx_1}(n) - R^{-1}_{x_1x_1}(n)R_{x_1x_2}(n)H_2(n) \\ H_2(n) &= R^{-1}_{x_2x_2}(n)r_{yx_2}(n) - R^{-1}_{x_2x_2}(n)R_{x_2x_1}(n)H_1(n) \end{aligned}$$

Hence, we deduct that:

$$\begin{aligned} \frac{\partial H_1(n)}{\partial H_2(n)} &= -R^{-1}_{x_1x_1}(n)R_{x_1x_2}(n) && (16)\\ \frac{\partial H_2(n)}{\partial H_1(n)} &= -R^{-1}_{x_2x_2}(n)R_{x_2x_1}(n) && (17) \end{aligned}$$

and expression (11) becomes :

$$R^{-1}(n) = \begin{bmatrix} R_1^{-1}(n) & O_{L\times L} \\ O_{L\times L} & R_2^{-1}(n) \end{bmatrix} \begin{bmatrix} I_{L\times L} & \frac{\partial H_1(n)}{\partial H_2(n)} \\ \frac{\partial H_2(n)}{\partial H_1(n)} & I_{L\times L} \end{bmatrix} \quad (18)$$

which is easily generalized to any number of channels (more than 2), as shown below.

3.3 Generalization to higher number of channels

The general factorization of the covariance matrix can be stated as follows:

Lemma 1 :

$$\begin{aligned} R^{-1}(n) &= \begin{bmatrix} R_1^{-1}(n) & O_{L\times L} & \cdots & O_{L\times L} \\ O_{L\times L} & R_2^{-1}(n) & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ O_{L\times L} & O_{L\times L} & \cdots & R_P^{-1}(n) \end{bmatrix} \\ &\times \begin{bmatrix} I_{L\times L} & \frac{\partial H_1(n)}{\partial H_2(n)} & \cdots & \frac{\partial H_1(n)}{\partial H_p(n)} \\ \frac{\partial H_2(n)}{\partial H_1(n)} & I_{L\times L} & \cdots & \frac{\partial H_2(n)}{\partial H_p(n)} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial H_P(n)}{\partial H_1(n)} & \frac{\partial H_P(n)}{\partial H_2(n)} & \cdots & I_{L\times L} \end{bmatrix} \end{aligned} \quad (19)$$

where:

$$R_i(n) = \sum_{j=1}^{P} \left(\frac{\partial H_j(n)}{\partial H_i(n)}\right)^T R_{x_jx_i}(n) \quad (20)$$

The proof is obvious. As a result, a general form of the M-C RLS algorithm is:

$$H_i(n) = H_i(n-1) + R_i^{-1}(n) Z_i(n) e(n) \qquad (21)$$

where :

$$Z_i(n) = \sum_{j=1}^{P} \left(\frac{\partial H_j(n)}{\partial H_i(n)}\right)^T X_j(n) \quad (i = 1, 2, \cdots, P) \qquad (22)$$

4 The M-C LMS algorithm

4.1 Classical derivation

The minimum mean-square error criterion is defined by:

$$J = \mathcal{E}\{(y(n) - H^T X(n))^2\} \qquad (23)$$

where $\mathcal{E}\{.\}$ is the expectation operator. Let $F(H)$ denote the value of the gradient vector. According to the steepest descent method, the updated value of H at iteration $n+1$ is computed by using the simple recursion:

$$H(n+1) = H(n) + \frac{\mu}{2}[-F(H(n))] \qquad (24)$$

where μ is a positive step-size constant. Differentiating (23) with respect to the filter, we get the following value for the gradient vector :

$$\begin{aligned} F(H) &= [F_1^T(H), F_2^T(H), \cdots, F_P^T(H)]^T &(25)\\ &= \frac{\partial J}{\partial H} = -2r + 2RH &(26) \end{aligned}$$

where $r = \mathcal{E}\{y(n)X(n)\}$ and $R = \mathcal{E}\{X(n)X^T(n)\}$. By taking $F(H) = 0$, we obtain the Wiener-Hopf equation:

$$RH = r \qquad (27)$$

The steepest-descent algorithm is now:

$$H_i(n+1) = H_i(n) + \mu \mathcal{E}\{X_i(n+1)e(n+1)\} \qquad (28)$$

and the classical stochastic approximation provides the M-C LMS algorithm:

$$H_i(n+1) = H_i(n) + \mu X_i(n+1)e(n+1) \qquad (29)$$

of which the classical stability condition under appropriate independence assumptions is:

$$0 < \mu < 2/\left(L \sum_{i=1}^{P} \sigma_{x_i}^2\right) \qquad (30)$$

where $\sigma_{x_i}^2$ are the variances of the input signals. Under this condition, the weight vector converges to the optimal Wiener-Hopf solution.

However, the gradient vector corresponding to filter number i is :

$$F_i(H) = -2\left(r_{yx_i} - \sum_{j=1}^{P} R_{x_i x_j} H_j\right) \qquad (31)$$

with $R_{x_i x_j} = \mathcal{E}\{X_i(n)X_j^T(n)\}$ and $r_{yx_i} = \mathcal{E}\{y(n)X_i(n)\}$, which clearly shows some dependency of F_i on the full vector H. In other words the filters H_j with $j \neq i$ should influence the gradient vector F_i when seeking the minimum [3].

4.2 Improved version for $P = 2$

We have seen in section 3.3 for the M-C RLS algorithm the dependency of each filter H_i on the other ones (Lemma 1). This information is now used to differentiate another way the criterion with respect to the tap-weight. This partial derivative is easily obtained by writing that H_1 (resp. H_2) depends on H_2 (resp. H_2). We get:

$$\begin{aligned} F_1(H_1) &= \frac{\partial J}{\partial H_1}(H_2) &(32)\\ &= -2\mathcal{E}\{\left(X_1(n) + \left(\frac{\partial H_2}{\partial H_1}\right)^T X_2(n)\right)\\ &\quad \times \left(y(n) - H_1^T X_1(n) - H_2^T X_2(n)\right)\} \end{aligned}$$

$$\begin{aligned} F_2(H_2) &= \frac{\partial J}{\partial H_2}(H_1) &(33)\\ &= -2\mathcal{E}\{\left(X_2(n) + \left(\frac{\partial H_1}{\partial H_2}\right)^T X_1(n)\right)\\ &\quad \times \left(y(n) - H_1^T X_1(n) - H_2^T X_2(n)\right)\} \end{aligned}$$

The Wiener-Hopf equation (27) easily provides the missing terms :

$$\frac{\partial H_1}{\partial H_2} = -R_{x_1x_1}^{-1} R_{x_1x_2} \qquad (34)$$

$$\frac{\partial H_2}{\partial H_1} = -R_{x_2x_2}^{-1} R_{x_2x_1} \qquad (35)$$

It is easily checked by replacing (34) (resp. (35)) in (32) (resp. (33)) that F_1 (resp. F_2) now depends only on H_1 (resp. H_2), which is not the case in the classical M-C LMS algorithm.

After the classical stochastic approximation applied on (32) and (33), we obtain the improved two-channel LMS :

$$H_1(n+1) = H_1(n) + \mu Z_1(n+1)e(n+1) \qquad (36)$$

$$H_2(n+1) = H_2(n) + \mu Z_2(n+1)e(n+1) \qquad (37)$$

where:

$$0 < \mu < 2/\left(L(\sigma_{z_1}^2 + \sigma_{z_2}^2)\right) \qquad (38)$$

and:

$$\begin{aligned} Z_1(n) &= X_1(n) - R_{x_1x_2}R_{x_2x_2}^{-1}X_2(n) &(39)\\ Z_2(n) &= H_2(n) - R_{x_2x_1}R_{x_1x_1}^{-1}X_1(n) &(40) \end{aligned}$$

Note that this algorithm has a form very similar to that of the two-channel RLS.

5 The M-C APA algorithm

5.1 A straightforward M-C APA

A simple trick for obtaining the mono-channel APA [6] is to search for the algorithm of a stochastic gradient type cancelling N a posteriori errors [7]. This requirement results in an underdetermined set of linear equations of which the minimum-norm solution is chosen. In the following, this technique is extended in order to fit to our problem.

By definition, the set of N a priori errors and N a posteriori errors are:

$$E(n+1) = Y(n+1) - \underline{X}^T(n+1)H(n) \qquad (41)$$
$$E_a(n+1) = Y(n+1) - \underline{X}^T(n+1)H(n+1) \qquad (42)$$

where:

$$\underline{X}(n+1) = \begin{matrix} \underline{X_1}(n+1) \\ \vdots \\ \underline{X_P}(n+1) \end{matrix} \qquad (43)$$

is a matrix of size $PL \times N$, the $L \times N$ matrix $\underline{X_i}(n+1)$ is made from the last input vectors $X_i(n+1)$ $(i = 1, 2, \cdots, P)$ and $Y(n+1)$ (resp. $E(n+1)$) is the vector of the N last samples of the reference signal $y(n+1)$ (resp. error signal $e(n+1)$).

Using (41) and (42) plus the requirement that $E_a(n+1) = 0$, we obtain :

$$\underline{X}^T(n+1)\Delta H(n+1) = E(n+1) \qquad (44)$$

where $\Delta H(n+1) = H(n+1) - H(n)$.

Eq. (44) (N equations in PL unknowns, $N \leq PL$) is an underdetermined set of linear equations. Hence, it has an infinite number of solutions, out of which the minimum-norm solution is chosen, so that the adaptive filter has smooth variations. This results in:

$$H(n+1) = H(n) + \underline{X}(n+1)(\underline{X}^T(n+1)\underline{X}(n+1))^{-1}E(n+1) \qquad (45)$$

However, nothing in this straightforward APA has solved the problem outlined in the above sections: the normalization matrix $\underline{X}^T(n+1)\underline{X}(n+1) = \sum_{i=1}^{P} \underline{X_i}^T(n+1)\underline{X_i}(n+1)$ does not involve the cross-correlation elements of the P input signals, and indeed, this algorithm converges slowly [7].

5.2 Improved version for $P = 2$

A simple way of improving the previous adaptive algorithm is to use the orthogonality and decorrelation properties of section 4, which appear in this context. Let us derive the new algorithm by requiring a condition similar to the one used in the improved M-C LMS case but weaker. Just use the constraint that ΔH_i be orthogonal to $\underline{X_j}$. As a result, we take into account separately the contributions of each input signal. This constraint reads:

$$C_1(n+1) = \underline{X_2}^T(n+1)\Delta H_1(n+1) = 0_{N\times 1} \qquad (46)$$
$$C_2(n+1) = \underline{X_1}^T(n+1)\Delta H_2(n+1) = 0_{N\times 1} \qquad (47)$$

From (44), (46), and (47) we obtain the improved M-C APA [7].

6 Conclusion

This paper has presented a general view on multi-channel adaptive algorithms in the context of acoustic echo cancellation. We have analyzed the drawback of M-C identification and have showed that in a real situation it does not occur. However, since the input signals are strongly cross-correlated, adaptive algorithms should take into account these statistics to have a good convergence rate. We have given an original form of the M-C RLS by factorizing the covariance matrix and from this new factorization we have derived the improved M-C LMS algorithm and the improved M-C APA. We have also seen the links that exist between all these algorithms.

References

[1] M. G. Bellanger, "Adaptive Digital Filters and Signal Analysis", New York : Marcel Dekker Inc., 1987.

[2] M. M. Sondhi, D. R. Morgan, "Acoustic Echo Cancellation for Stereophonic Teleconferencing", IEEE Workshop on Audio and Acoustics, Mohonk Moutain House, NY, May 1991.

[3] 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.

[4] Y. Mahieux, A. Gilloire, F. Khalil, "Annulation d'Echo en Teleconference Stereophonique", GRETSI 1993, FRANCE.

[5] J. Benesty, F. Amand, A. Gilloire, Y. Grenier, "Adaptive Filtering Algorithms for Stereophonic Acoustic Echo Cancellation", IEEE ICASSP, Detroit, USA, 1995.

[6] K. Ozeki, T. Umeda, "An Adaptive Filtering Algorithm Using an Orthogonal Projection to an Affine Subspace and its Properties", Elec. and Com. in Japan, Vol. J67-A, N5, pp. 126-132, Feb. 1984.

[7] J. Benesty, P. Duhamel, Y. Grenier, "A Multi-Channel Affine Projection Algorithm with Applications to Multi-Channel Acoustic Echo Cancellation", IEEE Trans. on SPL, Feb. 1996.