Convergence Analysis of a Multi-Channel Acoustic Echo Canceller
多チャンネルエコーキャンセラの収束解析

Akihiro Hirano
平野 晃宏

C&C Media Research Laboratories
NEC Corporation
日本電気株式会社
*C&C*メディア研究所

ABSTRACT

This paper analyzes convergence characteristics of a multi-channel acoustic echo canceller. A general procedure for analyzing averaged tap-weights and their convergence condition are shown. As examples of analyses, convergence for both uncrosscorrelated reference input signals and crosscorrelated signals are examined. The tap weights converge to their optimum values if the input signals are not crosscorrelated. When the reference input signals are delayed and attenuated versions of a white noise, the tap weights do not converge to the optimum values. Computer simulation results confirm the analyses.

あらまし

線形結合型多チャンネルエコーキャンセラを対象とした、収束特性の解析結果を報告する。タップ重みの期待値とその収束条件を解析する一般的手法を示す。具体的な解析例として、入力信号間が相互相関を持たない場合と、強い相互相関を持つ場合を解析する。相互相関がない場合には、タップ重みは最適値に収束する。各チャンネルの参照入力信号が、単一の白色信号に遅延と振幅差を与えたものである場合には、タップ重みは、最適値であるエコーパスのインパルス応答には収束しない。計算機シミュレーションにより、解析の妥当性を示す。

1. Introduction

Echo cancellers are used to reduce echoes in a wide range of applications, such as TV conference systems and hands-free telephones. To realistic TV conference systems, multi-channel audio, at least stereophonic, is essential. For multi-channel teleconference systems, multi-channel acoustic echo cancellers have been studied[1-6].

In multi-channel echo cancellers, the influence of strong crosscorrelation between input signals is one of the most serious problems[6-10]. In order to overcome this problem, improved echo cancellation algorithms have been proposed [4-5]. However, convergence of multi-channel echo cancellers for cross-correlated input signals has not been studied in details. Reported convergence analysis is only for a two or a three channel case [6-9].

This paper investigates convergence characteristics of a multi-channel acoustic echo canceller for arbitrary number of channels, i.e. M channels. Section 2 briefly reviews the multi-channel acoustic echo canceller based on linear combination[1]. A general procedure for a convergence analysis of this echo canceller is formulated in Section 3, followed by detailed analyses for an uncrosscorrelated signals and a cross-correlated signals. Computer simulation results will validate the analyses.

2. Multi-Channel Acoustic Echo Canceller

Let us concentrate on a multi-channel acoustic echo canceller based on linear combination[1]. Figure 1 depicts a block diagram of an M-channel echo canceller. The echo canceller consists of $M \times M$ adaptive filters corresponding to $M \times M$ echo paths from M loudspeakers to M microphones. An adaptive filter with a tap-weight vector $\boldsymbol{w}_{i,j}{}^{(n)}$ ($i = 1,\cdots,M$; $j = 1,\cdots,M$) estimates the impulse response of an echo path $\boldsymbol{h}_{i,j}$ from i-th loudspeaker to j-th microphone. A superscript (n) denotes the time index.

The input signal vector $\boldsymbol{x}^{(n)}$ is defined by

$$\boldsymbol{x}^{(n)} = \left[\boldsymbol{x}_1^{(n)T} \ \boldsymbol{x}_2^{(n)T} \cdots \boldsymbol{x}_M^{(n)T} \right]^T \qquad (1)$$

where $[\cdot]^T$ denotes the transpose of a matrix $[\cdot]$. $\boldsymbol{x}^{(n)}$ consists of M input signal vectors $\boldsymbol{x}_i^{(n)}$, each contains N input samples for an N-tap case. The subscripts i of $\boldsymbol{x}_i^{(n)}$ denotes the channel number. Tap-weight matrix $\mathbf{W}^{(n)}$ is given by

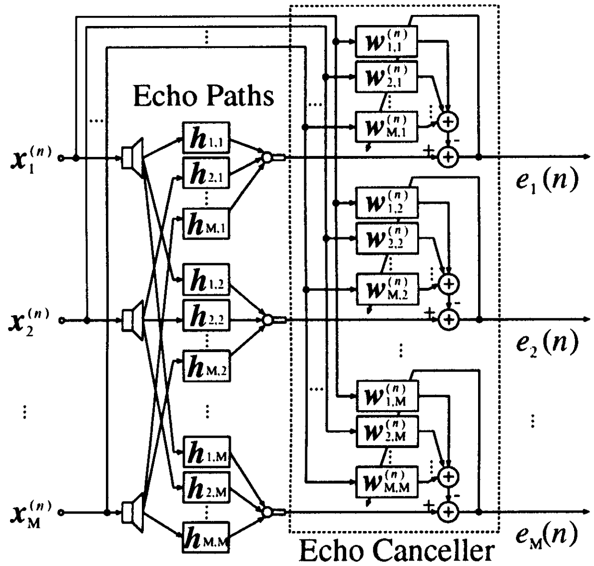


Fig. 1. Block diagram of M-channel acoustic echo canceller.

$$\mathbf{W}^{(n)} = \begin{bmatrix} w_{1,1}^{(n)} & w_{1,2}^{(n)} & \cdots & w_{1,M}^{(n)} \\ w_{2,1}^{(n)} & w_{2,2}^{(n)} & \cdots & w_{2,M}^{(n)} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ w_{M,1}^{(n)} & w_{M,2}^{(n)} & \cdots & w_{M,M}^{(n)} \end{bmatrix}. \quad (2)$$

Using the echo-path impulse-response matrix $\mathbf{H}$ defined by

$$\mathbf{H} = \begin{bmatrix} \boldsymbol{h}_{1,1} & \boldsymbol{h}_{1,2} & \cdots & \boldsymbol{h}_{1,M} \\ \boldsymbol{h}_{2,1} & \boldsymbol{h}_{2,2} & \cdots & \boldsymbol{h}_{2,M} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \boldsymbol{h}_{M,1} & \boldsymbol{h}_{M,2} & \cdots & \boldsymbol{h}_{M,M} \end{bmatrix}, \quad (3)$$

the tap-weight error matrix $\Theta^{(n)}$ is given by

$$\Theta^{(n)} = \mathbf{H} - \mathbf{W}^{(n)}. \quad (4)$$

The error vector $\boldsymbol{e}^{(n)}$ is calculated by

$$\boldsymbol{e}^{(n)} = \left[e_1(n) \;\; e_2(n) \cdots e_M(n) \right]^T = \Theta^{(n)^T} \boldsymbol{x}^{(n)}. \quad (5)$$

$\boldsymbol{e}^{(n)}$ consists of M residual echo signals.

Assuming the LMS (Least Mean Squares) algorithm[11], the tap-weight matrix $\mathbf{W}^{(n)}$ is updated by

$$\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} + \mu \boldsymbol{x}^{(n)} (\boldsymbol{e}^{(n)^T} + \boldsymbol{v}^{(n)^T}). \quad (6)$$

In (6), a positive constant μ is a step-size which controls the convergence. $\boldsymbol{v}^{(n)}$ is an additive noise vector defined by

$$\boldsymbol{v}^{(n)} = \left[v_1(n) \;\; v_2(n) \cdots v_m(n) \right]^T. \quad (7)$$

The additive noise $\boldsymbol{v}^{(n)}$ is assumed to be independent of the input signal $\boldsymbol{x}^{(n)}$.

3. Convergence Analysis of Averaged Tap-Weight Error

A general procedure to analyze the ensemble average of the tap-weight error matrix defined by

$$\mathbf{M}^{(n)} = E[\Theta^{(n)}] \quad (8)$$

will be shown. The updating equation for $\Theta^{(n)}$ is derived from (4), (5), and (6) as

$$\Theta^{(n+1)} = \Theta^{(n)} - \mu \boldsymbol{x}^{(n)} \boldsymbol{x}^{(n)^T} \Theta^{(n)} - \mu \boldsymbol{x}^{(n)} \boldsymbol{v}^{(n)^T}. \quad (9)$$

By taking an ensemble average of (9), difference equation for $\mathbf{M}^{(n)}$ becomes

$$\mathbf{M}^{(n+1)} = (\mathbf{I}_{MN} - \mu \mathbf{R}) \mathbf{M}^{(n)} \quad (10)$$

where $\mathbf{I}_{MN}$ is a $MN \times MN$ unit matrix, $\mathbf{R}$ is an "extended" covariance matrix defined by

$$\mathbf{R} = E[\boldsymbol{x}^{(n)} \boldsymbol{x}^{(n)^T}]. \quad (11)$$

$\mathbf{R}$ consists of both the auto-covariance matrices, $E[\boldsymbol{x}_i^{(n)} \boldsymbol{x}_i^{(n)^T}]$, and the cross-covariance $E[\boldsymbol{x}_i^{(n)} \boldsymbol{x}_j^{(n)^T}]$ $(i \neq j)$.

By introducing an orthonormal matrix $\mathbf{P}$ which diagonalizes $\mathbf{R}$, the difference equation for $\mathbf{M}^{(n)}$ becomes

$$\tilde{\mathbf{M}}^{(n+1)} = (\mathbf{I}_{MN} - \mu \Lambda) \tilde{\mathbf{M}}^{(n)}. \quad (12)$$

In (12), a new variable $\tilde{\mathbf{M}}^{(n)}$ and an eigenmatrix Λ are defined by

$$\tilde{\mathbf{M}}^{(n)} = \mathbf{P} \mathbf{M}^{(n)} \quad (13)$$

and

$$\Lambda = \mathbf{P} \mathbf{R} \mathbf{P}^{-1}, \quad (14)$$

respectively.

From (12), the solution of $\tilde{\mathbf{M}}^{(n)}$ is derived as

$$\tilde{\mathbf{M}}^{(n)} = (\mathbf{I}_{MN} - \mu \Lambda)^n \tilde{\mathbf{M}}^{(0)} \quad (15)$$

where $\tilde{\mathbf{M}}^{(0)}$ is the initial value of $\tilde{\mathbf{M}}^{(n)}$. (13) and (15) lead to the solution for the averaged tap-weight $\mathbf{M}^{(n)}$ given by

$$\mathbf{M}^{(n)} = \mathbf{P}^{-1} \tilde{\mathbf{M}}^{(n)} = \mathbf{P}^{-1} (\mathbf{I}_{MN} - \mu \Lambda)^n \mathbf{P} \mathbf{M}^{(0)}. \quad (16)$$

The convergence condition for $\mathbf{M}^{(n)}$ is

$$0 < \mu < \frac{2}{\lambda_{\max}} \quad (17)$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}$. Since $\lambda_{\max}$ is always less than $trace(\mathbf{R})$,

$$0 < \mu < \frac{2}{trace(\mathbf{R})} \quad (18)$$

is the sufficient condition for convergence.

4. Delatiled Analyses for Specific Input Signals

4.1. Uncrosscorrelated Signals

If the input signals $x_i^{(n)}$ $(i = 1, \cdots, M)$ are not crosscorrelated, the "extended" covariance matrix $\mathbf{R}$ becomes

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{1,1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{2,2} & \cdots & \mathbf{0} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_{M,M} \end{bmatrix}. \tag{19}$$

By substituting (19) into (10), M^2 independent equations

$$\boldsymbol{m}_{i,j}^{(n+1)} = (\mathbf{I}_N - \mu \mathbf{R}_{i,i}) \boldsymbol{m}_{i,j}^{(n)} \quad (i = 1, \cdots, M; j = 1, \cdots, M) \tag{20}$$

are derived. In (20), $\boldsymbol{m}_{i,j}^{(n)}$ is defined by

$$\boldsymbol{m}_{i,j}^{(n)} = E[\boldsymbol{h}_{i,j} - \boldsymbol{w}_{i,j}{}^{(n)}]. \tag{21}$$

(20) is completely same as the difference equation for the monaural LMS algorithm. Therefore, if the convergence condition similar to a monaural case is satisfied, the averaged tap-weights will converge to their optimum values.

4.2. Crosscorrelated Signals

In the following analysis, a single-talker case is assumed in which one talker in the distant room is speaking. All signals sent to the near-end room contain the same speech signal. Thus, crosscorrelated signals are generated. Note that such a situation is commonly encountered in many teleconferences.

The relation between M input signals and a source signal is given by

$$x_1(n) = a_1 x(n) \tag{22}$$

$$x_i(n) = a_i x(n - \sum_{j=1}^{i-1} n_j) \quad (i = 2, 3, \cdots, M) \tag{23}$$

where n_j is a time delay between $x_j(n)$ and $x_{j+1}(n)$, a_i is an attenuation factor. $x(n)$ is also assumed to be zero mean, white-Gaussian process with a unit variance and to be independent of $\Theta^{(n)}$. The independent assumption of $\boldsymbol{x}^{(n)}$ and $\Theta^{(n)}$ is common in many analyses and is valid if the step size μ is small[11].

For input signals shown in (22) and (23), the auto-covariance matrices and the cross-covariance matrices are given by

$$\mathbf{R}_{i,i} = a_i{}^2 \mathbf{I}_N \tag{24}$$

$$\mathbf{R}_{i,j} = \begin{bmatrix} \mathbf{0}_{n_i+\cdots+n_{j-1}, N-n_i-\cdots-n_{j-1}} & \mathbf{0}_{n_i+\cdots+n_{j-1}, n_i+\cdots+n_{j-1}} \\ a_i a_j \mathbf{I}_{N-n_i-\cdots-n_{j-1}} & \mathbf{0}_{N-n_i\cdots-n_{j-1}, n_i+\cdots+n_{j-1}} \end{bmatrix} \quad (i < j) \tag{25}$$

$$\mathbf{R}_{j,i} = \mathbf{R}_{i,j}{}^T \qquad (i < j) \tag{26}$$

where $\mathbf{0}_{i,j}$ is an $i \times j$ zero matrix. By substituting (24) – (26) into (10) and by changing the order of rows or columns, the difference equation is divided into $2M - 1$ equations:

$$\mathbf{M}_i^{(n+1)} = (\mathbf{I} - \mu \mathbf{R}_i) \mathbf{M}_i^{(n)} \quad (i = 1, \cdots, 2M - 1). \tag{27}$$

$\mathbf{M}_i^{(n)}$ is a sub-matrix of $\mathbf{M}^{(n)}$. For $i = 1, \cdots, M$, $\mathbf{M}_i^{(n)}$is defined by

$$\mathbf{M}_i^{(n)} = \begin{bmatrix} \boldsymbol{m}_{1,1,o_{1,i-1},o_{1,i}-1}^{(n)} & \cdots & \boldsymbol{m}_{1,M,o_{1,i-1},o_{1,i}-1}^{(n)} \\ \boldsymbol{m}_{2,1,o_{2,i-1},o_{2,i}-1}^{(n)} & \cdots & \boldsymbol{m}_{2,M,o_{2,i-1},o_{2,i}-1}^{(n)} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \boldsymbol{m}_{i,1,o_{i,i-1},o_{i,i1}-1}^{(n)} & \cdots & \boldsymbol{m}_{i,M,o_{i,i-1},o_{i,i1}-1}^{(n)} \end{bmatrix}. \tag{28}$$

In (28), $\boldsymbol{m}_{i,j,k,l}^{(n)}$ is a sub-vector of $\boldsymbol{m}_{i,j}^{(n)}$ defined by

$$\boldsymbol{m}_{i,j,k,l}^{(n)} = [m_{i,j,k}(n)\ m_{i,j,k+1}(n) \cdots m_{i,j,l}(n)]^T \tag{29}$$

where $m_{i,j,k}(n)$ is $(k+1)$-th element of $\boldsymbol{m}_{i,j}^{(n)}$. $o_{i,j}$ is

$$o_{i,j} = \begin{cases} \sum_{k=i}^{j} n_k & (i \le j) \\ N & (i > j) \end{cases}. \tag{30}$$

For $i = M+1, \cdots, 2M-1$, $\mathbf{M}_i^{(n)}$ is defined as in (31) where $p_{i,j}$ is

$$p_{i,j} = \begin{cases} N - \sum_{k=i}^{j} n_k & (i \le j) \\ 0 & (i > j) \end{cases}. \tag{32}$$

$\mathbf{R}_i$ is defined by

$$\mathbf{M}_{M+i}^{(n)} = \begin{bmatrix} \boldsymbol{m}_{i+1,1,p_{i,i},p_{i,i+1}-1}^{(n)} & \cdots & \boldsymbol{m}_{i+1,M,p_{i,i},p_{i+1,i}-1}^{(n)} \\ \boldsymbol{m}_{i+2,1,p_{i,i+1},p_{i+1,i+1}-ni+1-1}^{(n)} & \cdots & \boldsymbol{m}_{i+2,M,p_{i,i+1},p_{i+1,i+1}-1}^{(n)} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \boldsymbol{m}_{M,1,p_{i,M-1},p_{i+1,M-1}-1}^{(n)} & \cdots & \boldsymbol{m}_{M,M,p_{i,M-1},p_{i+1,M-1}-1}^{(n)} \end{bmatrix} \tag{31}$$

$$\mathbf{R}_i = \begin{bmatrix} {a_1}^2\mathbf{I}_{n_i} & a_1a_2\mathbf{I}_{n_i} & \cdots & a_1a_i\mathbf{I}_{n_i} \\ a_1a_2\mathbf{I}_{n_i} & {a_2}^2\mathbf{I}_{n_i} & \cdots & a_2a_i\mathbf{I}_{n_i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_1a_i\mathbf{I}_{n_i} & a_2a_i\mathbf{I}_{n_i} & \cdots & {a_i}^2\mathbf{I}_{n_i} \end{bmatrix}$$

$$(i = 1, 2, \cdots, M) \tag{33}$$

and

$$\mathbf{R}_{M+i} = \begin{bmatrix} {a_{i+1}}^2\mathbf{I}_{n_i} & a_{i+1}a_{i+2}\mathbf{I}_{n_i} & \cdots & a_{i+1}a_M\mathbf{I}_{n_i} \\ a_{i+1}a_{i+2}\mathbf{I}_{n_i} & {a_{i+2}}^2\mathbf{I}_{n_i} & \cdots & a_{i+2}a_M\mathbf{I}_{n_i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{i+1}a_M\mathbf{I}_{n_i} & a_{i+2}a_M\mathbf{I}_{n_i} & \cdots & {a_M}^2\mathbf{I}_{n_i} \end{bmatrix}$$

$$(i = 1, 2, \cdots, M-1). \tag{34}$$

For $i = M$, n_M defined by

$$n_M = N - \sum_{i=1}^{M-1} n_i \tag{35}$$

is used. $\mathbf{R}_i$ is a sub-matrix of the "extended" covariance matrix $\mathbf{R}$.

The difference equations for $\mathbf{M}_1^{(n)}$ and $\mathbf{M}_{2M-1}^{(n)}$ are completely same as those for the monaural LMS. Therefore, $\mathbf{M}_1^{(n)}$ and $\mathbf{M}_{2M-1}^{(n)}$ have unique solutions and converge to the optimum values. For $\mathbf{M}_2^{(n)}$ and $\mathbf{M}_{2M-2}^{(n)}$, the difference equations are similar to those for a two-channel case. $\mathbf{M}_2^{(n)}$ and $\mathbf{M}_{2M-2}^{(n)}$ do not converge to the optimum.

Convergence analysis of $\mathbf{M}_i^{(n)}$ for $i = 3, \cdots, M$ requires diagonalization of $\mathbf{R}_i$, i.e., diagonalization of an $Mn_i \times Mn_i$ matrix. Such a diagonalization can be carried out by introducing an orthonormal matrix in (36) where

$$S_{i,j} = \sum_{k=i}^{j} a_k^2 \tag{37}$$

$$q_{i,j} = \frac{1}{\sqrt{S_{i,j}}}. \tag{38}$$

By using $\mathbf{P}_i$, $\mathbf{R}_i$ is diagonalized as

$$\Lambda_i = \mathbf{P}_i\mathbf{R}_i{\mathbf{P}_i}^{-1} = \begin{bmatrix} S_{1,i}\mathbf{I}_{n_i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{39}$$

Obviously, $\mathbf{R}_i$ is singular, and therefore, $\mathbf{M}_i^{(n)}$ do not have a unique solution.

The convergence condition for $\mathbf{M}_i^{(n)}$ is determined as

$$0 < \mu < \frac{2}{\sum_{j=1}^{i} a_j^2} \tag{40}$$

because maximum eigenvalue of $\mathbf{R}_i$ is $S_{1,i} = \sum_{j=1}^{i} a_j^2$. If this condition is satisfied, $\mathbf{M}_i^{(n)}$ converges to (41). Analyses for $i = M+1, \cdots, 2M-3$ can be carried out in the same manner as those for $i = 3, \cdots, M$. Since the averaged tap-weight error do not converge to zero, the tap-weight matrix never converge to the optimum.

5. Computer Simulations

Computer simulations have been carried out and their results have been compared with the analytical results. Both uncrosscorrelated and crosscorrelated cases were examined. The number of channels M and the number of taps N were 4 and 20, respectively. The echo paths were given by

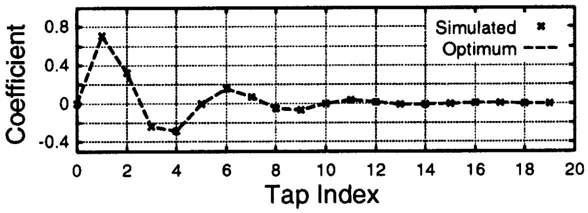
$$h_{1,1,k} = e^{-0.3k}\sin(0.4\pi k) \tag{42}$$

$$h_{2,1,k} = e^{-0.3k}\sin(0.3\pi k) \tag{43}$$

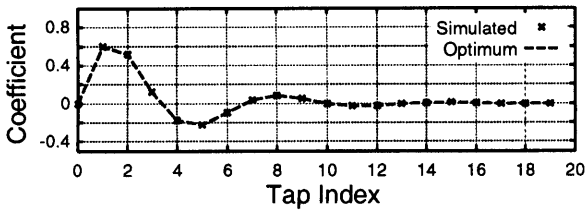
$$h_{3,1,k} = e^{-0.4k}\sin(0.3\pi k) \tag{44}$$

$$\mathbf{P}_i = \begin{bmatrix} q_{1,i}a_1\mathbf{I}_{n_i} & q_{1,i}a_2\mathbf{I}_{n_i} & q_{1,i}a_3\mathbf{I}_{n_i} & \cdots & q_{1,i}a_{i-1}\mathbf{I}_{n_i} & q_{1,i}a_i\mathbf{I}_{n_i} \\ q_{1,i}q_{2,i}S_{2,i}\mathbf{I}_{n_i} & -q_{1,i}q_{2,i}a_1a_2\mathbf{I}_{n_i} & -q_{1,i}q_{2,i}a_1a_3\mathbf{I}_{n_i} & \cdots & -q_{1,i}q_{2,i}a_1a_{i-1}\mathbf{I}_{n_i} & -q_{1,i}q_{2,i}a_1a_i\mathbf{I}_{n_i} \\ \mathbf{0} & q_{2,i}q_{3,i}S_{3,i}\mathbf{I}_{n_i} & -q_{2,i}q_{3,i}a_2a_3\mathbf{I}_{n_i} & \cdots & -q_{2,i}q_{3,i}a_2a_{i-1}\mathbf{I}_{n_i} & -q_{2,i}q_{3,i}a_2a_i\mathbf{I}_{n_i} \\ \mathbf{0} & \mathbf{0} & q_{3,i}q_{4,i}S_{4,i}\mathbf{I}_{n_i} & \cdots & -q_{3,i}q_{4,i}a_3a_{i-1}\mathbf{I}_{n_i} & -q_{3,i}q_{4,i}a_3a_i\mathbf{I}_{n_i} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -q_{j-1,i}q_{j,i}a_{j-1}a_{i-1}\mathbf{I}_{n_i} & -q_{j-1,i}q_{j,i}a_{j-1}a_i\mathbf{I}_{n_i} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & q_{i-1}q_iS_{i,i}\mathbf{I}_{n_i} & -q_{i-1,i}q_{i,i}a_{i-1}a_i\mathbf{I}_{n_i} \end{bmatrix} \tag{36}$$

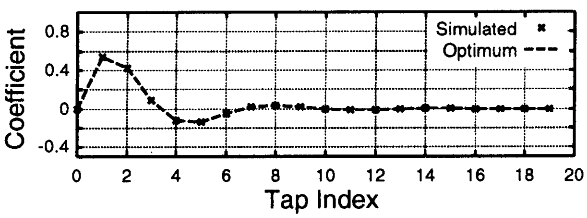
$$\mathbf{M}_i^{(\infty)} = {\mathbf{P}_i}^{-1}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(i-1)n_i} \end{bmatrix}\mathbf{P}_i\mathbf{M}_i^{(0)} = \frac{1}{S_{1,i}}\begin{bmatrix} (S_{1,i} - {a_1}^2)\mathbf{I}_{n_i} & -a_1a_2\mathbf{I}_{n_i} & \cdots & -a_1a_M\mathbf{I}_{n_i} \\ -a_1a_2\mathbf{I}_{n_i} & (S_{1,i} - {a_2}^2)\mathbf{I}_{n_i} & \cdots & -a_2a_M\mathbf{I}_{n_i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -a_1a_i\mathbf{I}_{n_i} & -a_2a_i\mathbf{I}_{n_i} & \cdots & (S_{1,i} - {a_i}^2)\mathbf{I}_{n_i} \end{bmatrix}\mathbf{M}_i^{(0)} \tag{41}$$



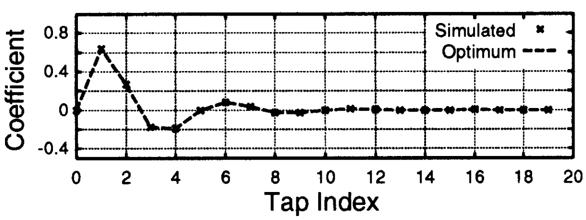
(a) $E[w_{1,1}^{(10000)}]$



(a) $E[w_{2,1}^{(10000)}]$

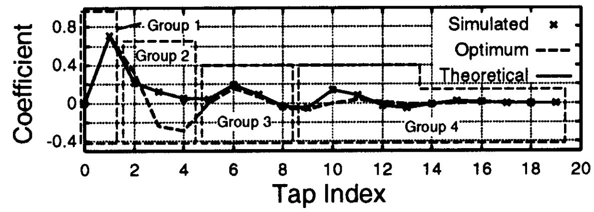


(a) $E[w_{3,1}{}^{(10000)}]$

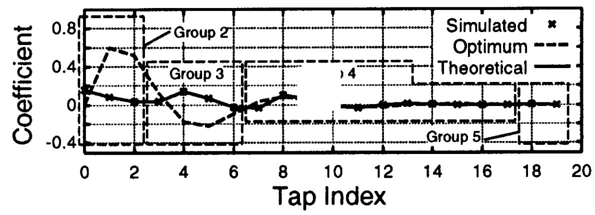


(b) $E[w_{4,1}(10000)]$

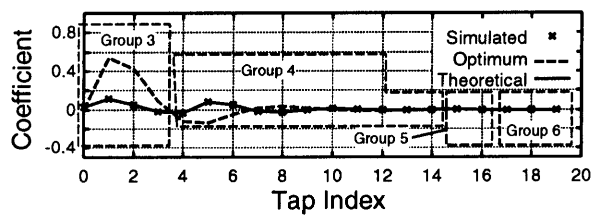
Fig. 2. Tap weights after converge for uncrosscorrelated signals.



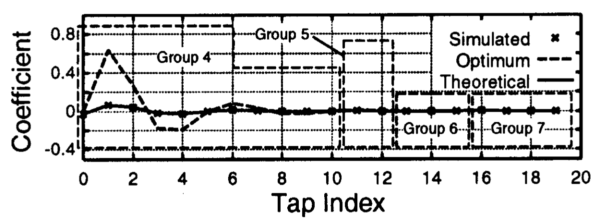
(a) $E[w_{1,1}^{(10000)}]$



(a) $E[w_{2,1}^{(10000)}]$



(a) $E[w_{3,1}{}^{(10000)}]$



(b) $E[w_{4,1}(10000)]$

Fig. 3. Tap weights after converge for crosscorrelated signals.

$$h_{4,1,k} = e^{-0.4k}\sin(0.4\pi k) \tag{45}$$

where $h_{i,j,k}$ is the k-th element of the echo path vector $\boldsymbol{h}_{i,j}$.

For the uncrosscorrelated case, each reference input signal was independent white-Gaussian process with a unit variance. The crosscorrelated signals were generated by delaying and by attenuating a white-Gaussian signal. The attenuation and the delay parameters were $a_1 = 1.0$, $a_2 = 0.7$, $a_3 = 0.6$, $a4 = 0.5$, $n1 = 2$, $n2 = 3$, $n3 = 4$. Independent white-Gaussian noises have been added to the echoes as the additive noise. The variance of the additive noise was 0.01. The step size μ was settled as 0.01. An average of 1000 independent runs has been calculated.

Averaged tap-weight vectors after convergence for uncrosscorrelated case are shown in Fig. 2. Tap weights converge to the optimum values, i.e., the impulse responses of the echo paths. Figure 3 depicts the results for the strongly crosscorrelated signals. In Fig. 3, "Group i" means that the tap weights in this group are elements of $\mathbf{M}_i^{(n)}$. As shown by the analysis, the tap weights in Group 1 and Group 7 converge to the optimum values. However, all the other weights do not converge to the optimum. The convergence values agree with analytical results shown in (41).

6. Conclusion

A convergence analysis for a multi-channel acoustic echo canceller has been presented. Analysis procedure on the averaged tap-weights and their convergence condition have been formulated. The tap weights converge to their optimu values if the reference input signals are not crosscorrelated. For a single talker case, in which all reference input signals to the adaptive filters are assumed to be delayed and attenuated versions of a white-Gaussian signal, a part of the tap weights does not converge to the optimum. Computer simulation results confirm the analysis.

Acknowledgments

The author is indebted to Dr. Shin'ichi Koike, chief engineer of NEC Corporation, for helpful discussions and valuable comments. He also wishes to thank Dr. Takao Nishitani, the deputy general manager of Silicon System Research Laboratories, NEC Corporation, and Akihiko Sugiyama, a principal researcher of C&C Media Research Laboratories, NEC Corporation, for his guidance and continuous encouragement.

References

[1] T. Fujii and S. Shimada, "A Note on Multi-Channel Echo Cancellers," Technical Reports of IEICE on CS, pp. 7-14, Jan. 1984 (in Japanese).

[2] M. M. Sondhi and D. R. Morgan, "Acoustic Echo Cancellation for Stereophonic Teleconferencing," Proc. of IEEE ASSP Workshop Applied Signal Processing Audio Acoustics, 1991.

[3] A. Hirano and A. Sugiyama, "A Compact Multi-Channel Echo Canceller with a Single Adaptive Filter per Channel," Proc. of ISCAS '92, pp. 1922-1925, 1992.

[4] Y. Joncour and A. Sugiyama, "A Stereo Echo Canceler with Correct Echo-Path Identification," Technical Report of IEICE, Vol. DSP97-1, pp. 1-8, 1997

[5] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the problem of stereophonic acoustic echo cancellation," Proc. of ICASSP '97, pp. 303-306, 1997

[6] A. Hirano and A. Sigiyama, "Convergence Characteristics of a Multi-Channel Echo Canceller with Strongly Cross-correlated Input Signals – Analytical Results –," Proc. of 6th DSP Symposium, pp. 144-149, November 1991.

[7] A. Hirano and S. Koike, "Convergence Analysis of a Stereophonic Acoustic Echo Canceller Part I: Convergence Characteristics of Tap Weights," Proc. of 11th DSP Symposium, pp. 569-574, November 1996.

[8] S. Koike and A. Hirano, "Convergence Analysis of a Stereophonic Acoustic Echo Canceller Part II: – Mean Squared Error, Convergence Condition and Error Surface –," Proc. of 11-th DSP Symposium, pp. 575-580, November 1996 (in Japanese).

[9] A. Hirano, "Convergence Analysis of a Three-Channel Echo Canceller," Proc. of the 1997 IEICE General Conference, Vol. A, pp. 155, April, 1997.

[10] M. M. Sondhi and D. R. Morgan, "Stereophonic Acoustic Echo Cancellation – An Overview of the Fundamental Problem," IEEE SP Letters, vol. 2, no. 8, pp. 148-151, August 1995.

[11] B. Widrow and S. D. Stearns, "Adaptive Signal Processing," Englewood Cliffs, NJ: Prentice-Hall, 1985.

Appendix A. Derivation of $\mathbf{P}_i$

Diagonalization of $\mathbf{R}_i$ can be simplified to diagonalization of

$$\mathbf{R}_i' = \begin{bmatrix} a_1{}^2 & a_1a_2 & \cdots & a_1a_i \\ a_1a_2 & a_2{}^2 & \cdots & a_2a_i \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_1a_i & a_2a_i & \cdots & a_i{}^2 \end{bmatrix}. \quad \text{(A1)}$$

By introduction of a vector

$$\boldsymbol{a} = [a_1\ a_2 \cdots a_i]^T, \quad \text{(A2)}$$

$\mathbf{R}_i'$ is given by

$$\mathbf{R}_i' = \boldsymbol{a}\boldsymbol{a}^T = [a_1\boldsymbol{a}\ a_2\boldsymbol{a} \cdots a_i\boldsymbol{a}]. \quad \text{(A3)}$$

From (A3), it is obvious that the rank of the matrix $\mathbf{R}_i'$ is 1 and therefore, there is only one non-zero eigenvalue. Since

$$\mathbf{R}_i'\boldsymbol{a} = (\boldsymbol{a}\boldsymbol{a}^T)\boldsymbol{a} = \left(\sum_{j=1}^{i} a_j^2\right)\boldsymbol{a}, \quad \text{(A4)}$$

the non-zero eigenvalue λ is given by

$$\lambda = \sum_{j=1}^{i} a_j^2 \quad \text{(A5)}$$

and the eigenvector for λ is $\boldsymbol{a}$.

Other $i-1$ eigenvectors corresponding to $\lambda = 0$ can be selected from arbitrary vectors $\boldsymbol{b}$ which satisfies

$$\boldsymbol{a}^T\boldsymbol{b} = 0. \quad \text{(A6)}$$

Such vectors can be derived by Gram-Schmidt diagonalization of vectors

$$\boldsymbol{u}_j = [\mathbf{0}_{1,j}\ 1\ \mathbf{0}_{1,i-j-1}]^T \quad (j = 2, \cdots, i). \quad \text{(A7)}$$

An eigenvector $\boldsymbol{v}_2$ is calculated from $\boldsymbol{u}_2$ by

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2^t\boldsymbol{a}}{|\boldsymbol{u}_2||\boldsymbol{a}|}\frac{\boldsymbol{a}}{|\boldsymbol{a}|}. \quad \text{(A8)}$$

Similarly, the next eigenvector $\boldsymbol{v}_3$ is determined by

$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3^t\boldsymbol{a}}{|\boldsymbol{u}_3||\boldsymbol{a}|}\frac{\boldsymbol{a}}{|\boldsymbol{a}|} - \frac{\boldsymbol{u}_3^t\boldsymbol{v}_2}{|\boldsymbol{u}_3||\boldsymbol{v}_2|}\frac{\boldsymbol{v}_2}{|\boldsymbol{v}_2|}. \quad \text{(A9)}$$

Repeating this procedure will generate $i-1$ eigenvectors for $\lambda = 0$. Vectors $\boldsymbol{a}$ and $\boldsymbol{v}_j$ construct an orthonormal matrix

$$\mathbf{P}_i' = [\frac{\boldsymbol{a}}{|\boldsymbol{a}|}\ \frac{\boldsymbol{v}_2}{|\boldsymbol{v}_2|} \cdots \frac{\boldsymbol{v}_i}{|\boldsymbol{v}_i|}]^T. \quad \text{(A10)}$$

Expansion of $\mathbf{P}_i'$ to $\mathbf{P}_i$ is strait-forward.