STEREO ECHO CANCELLATION ALGORITHM USING IMAGINARY INPUT-OUTPUT RELATIONSHIPS

Suehiro Shimauchi and Shoji Makino

NTT Human Interface Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan
sima@splab.hil.ntt.jp

ABSTRACT

A new stereo echo cancellation algorithm using imaginary input-output relationships is proposed. This algorithm is based on the idea of approximating the input-output relationships for the reversed stereo signals. Unlike conventional methods, which use the input-output relationships in only one situation, our algorithm uses the simultaneous equations for the input-output relationships under two different conditions. Consequently, convergence to the unique steady-state solution is faster than with the conventional algorithm. Computer simulations using data recorded in a conference room demonstrate the effectiveness of this algorithm.

1. INTRODUCTION

A stereo teleconferencing system provides greater presence in teleconferencing compared with a monaural system. It helps listeners distinguish who is talking at the other end by means of spatial information. The most important problem in stereo echo cancellation is that the adaptive filter often misconverges or, if not, its convergence speed is very slow because of the cross-correlation between stereo signals [1][2][3]. Several methods for overcoming this problem have been proposed, which utilize uncorrelated elements of stereo signals [1] and variations in cross-correlation between stereo signals [3]. In this paper, we propose a new stereo echo cancellation algorithm from another viewpoint. It is based on the following idea: if input-output relationships for the stereo signals under two different conditions are known, then the simultaneous equations for two different stereo signals can be obtained, which would provide a unique steady-state solution.

2. STEREO ECHO CANCELLATION AND ITS PROBLEM

A stereo (two-channel) telecommunication system is shown in Fig. 1. There are four echo paths between

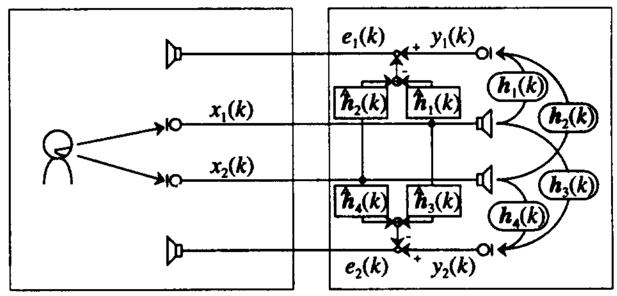

Figure 1: A stereo telecommunication system.

the loudspeakers and microphones: $\boldsymbol{h}_1(k)$, $\boldsymbol{h}_2(k)$, $\boldsymbol{h}_3(k)$, and $\boldsymbol{h}_4(k)$. The echoes of each microphone input are modeled as

$$y_1(k) = \boldsymbol{x}_1^T(k)\boldsymbol{h}_1(k) + \boldsymbol{x}_2^T(k)\boldsymbol{h}_2(k) \tag{1}$$

$$y_2(k) = \boldsymbol{x}_1^T(k)\boldsymbol{h}_3(k) + \boldsymbol{x}_2^T(k)\boldsymbol{h}_4(k), \tag{2}$$

where $\boldsymbol{x}_1(k) = [x_1(k), x_1(k-1), \ldots, x_1(k-L+1)]$, $\boldsymbol{x}_2(k) = [x_2(k), x_2(k-1), \ldots, x_2(k-L+1)]$, L is the filter tap length, k is a discrete time index, and ^{T} indicates a transpose. A stereo echo canceller makes echo replicas $\hat{y}_1(k)$ and $\hat{y}_2(k)$ by convolving stereo signals $\boldsymbol{x}_1(k)$ and $\boldsymbol{x}_2(k)$ with echo path estimating filters $\hat{\boldsymbol{h}}_1(k)$, $\hat{\boldsymbol{h}}_2(k)$, $\hat{\boldsymbol{h}}_3(k)$, and $\hat{\boldsymbol{h}}_4(k)$, as follows:

$$\hat{y}_1(k) = \boldsymbol{x}_1^T(k)\hat{\boldsymbol{h}}_1(k) + \boldsymbol{x}_2^T(k)\hat{\boldsymbol{h}}_2(k) \tag{3}$$

$$\hat{y}_2(k) = \boldsymbol{x}_1^T(k)\hat{\boldsymbol{h}}_3(k) + \boldsymbol{x}_2^T(k)\hat{\boldsymbol{h}}_4(k). \tag{4}$$

The estimation errors are derived as

$$\begin{aligned} e_1(k) &= y_1(k) - \hat{y}_1(k) \\ &= \boldsymbol{x}_1^T(k)\Delta\hat{\boldsymbol{h}}_1(k) + \boldsymbol{x}_2^T(k)\Delta\hat{\boldsymbol{h}}_2(k) \end{aligned} \tag{5}$$

$$\begin{aligned} e_2(k) &= y_1(k) - \hat{y}_1(k) \\ &= \boldsymbol{x}_1^T(k)\Delta\hat{\boldsymbol{h}}_3(k) + \boldsymbol{x}_2^T(k)\Delta\hat{\boldsymbol{h}}_4(k), \end{aligned} \tag{6}$$

where $\Delta\hat{\boldsymbol{h}}_1(k)$, $\Delta\hat{\boldsymbol{h}}_2(k)$, $\Delta\hat{\boldsymbol{h}}_3(k)$, and $\Delta\hat{\boldsymbol{h}}_4(k)$ are the error vectors of $\hat{\boldsymbol{h}}_1(k)$, $\hat{\boldsymbol{h}}_2(k)$, $\hat{\boldsymbol{h}}_3(k)$, and $\hat{\boldsymbol{h}}_4(k)$ from the true echo path impulse responses. The conventional normalized least mean squares (NLMS) algorithm estimates the error vectors $\Delta\hat{\boldsymbol{h}}_1(k)$, $\Delta\hat{\boldsymbol{h}}_2(k)$, $\Delta\hat{\boldsymbol{h}}_3(k)$, and

0-7803-3192-3/96 $5.00©1996 IEEE

$\Delta\hat{\boldsymbol{h}}_4(k)$ as the minimum-norm solutions of Eqs. (5) and (6), and uses them as the adjustment vectors of the echo path estimating filters, that is,

$$\begin{bmatrix}\hat{\boldsymbol{h}}_1(k+1)\\ \hat{\boldsymbol{h}}_2(k+1)\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{h}}_1(k)\\ \hat{\boldsymbol{h}}_2(k)\end{bmatrix} + \mu \begin{bmatrix}\Delta\hat{\boldsymbol{h}}_1(k)\\ \Delta\hat{\boldsymbol{h}}_2(k)\end{bmatrix} \tag{7}$$

$$\begin{bmatrix}\hat{\boldsymbol{h}}_3(k+1)\\ \hat{\boldsymbol{h}}_4(k+1)\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{h}}_3(k)\\ \hat{\boldsymbol{h}}_4(k)\end{bmatrix} + \mu \begin{bmatrix}\Delta\hat{\boldsymbol{h}}_3(k)\\ \Delta\hat{\boldsymbol{h}}_4(k)\end{bmatrix}, \tag{8}$$

where μ is the step size.

The echo path estimating filters should be estimated as $\hat{\boldsymbol{h}}_1(k) = \boldsymbol{h}_1(k)$, $\hat{\boldsymbol{h}}_2(k) = \boldsymbol{h}_2(k)$, $\hat{\boldsymbol{h}}_3(k) = \boldsymbol{h}_3(k)$, and $\hat{\boldsymbol{h}}_4(k) = \boldsymbol{h}_4(k)$ to achieve stable echo cancellation. However, the cross-correlation between $\boldsymbol{x}_1(k)$ and $\boldsymbol{x}_2(k)$ affects the echo path estimation. Therefore the echo path estimating filters are often estimated incorrectly. This is regarded as the most important problem in stereo echo cancellation [1][2][3].

3. NEW ALGORITHM

3.1. Basic concept

We consider the case where the input-output relationships for inputs $\boldsymbol{x}'_1(k)$ and $\boldsymbol{x}'_2(k)$, having a different cross-correlation from $\boldsymbol{x}_1(k)$ and $\boldsymbol{x}_2(k)$, can be obtained as follows:

$$y'_1(k) = \boldsymbol{x}'^T_1(k)\boldsymbol{h}_1(k) + \boldsymbol{x}'^T_2(k)\boldsymbol{h}_2(k) \tag{9}$$

$$y'_2(k) = \boldsymbol{x}'^T_1(k)\boldsymbol{h}_3(k) + \boldsymbol{x}'^T_2(k)\boldsymbol{h}_4(k). \tag{10}$$

Letting $\hat{y}'_1(k)$ and $\hat{y}'_2(k)$ be replicas of $y'_1(k)$ and $y'_2(k)$ respectively, the residual echoes $e'_1(k) = y'_1(k) - \hat{y}'_1(k)$ and $e'_2(k) = y'_2(k) - \hat{y}'_2(k)$ are derived as

$$e'_1(k) = \boldsymbol{x}'^T_1(k)\Delta\hat{\boldsymbol{h}}_1(k) + \boldsymbol{x}'^T_2(k)\Delta\hat{\boldsymbol{h}}_2(k) \tag{11}$$

$$e'_2(k) = \boldsymbol{x}'^T_1(k)\Delta\hat{\boldsymbol{h}}_3(k) + \boldsymbol{x}'^T_2(k)\Delta\hat{\boldsymbol{h}}_4(k). \tag{12}$$

Then the simultaneous equations for the stereo signals under two different conditions can be obtained from Eqs. (5), (6), (11), and (12) as

$$\begin{bmatrix}\boldsymbol{x}^T_1(k) & \boldsymbol{x}^T_2(k)\\ \boldsymbol{x}'^T_1(k) & \boldsymbol{x}'^T_2(k)\end{bmatrix}\begin{bmatrix}\Delta\hat{\boldsymbol{h}}_1(k)\\ \Delta\hat{\boldsymbol{h}}_2(k)\end{bmatrix} = \begin{bmatrix}e_1(k)\\ e'_1(k)\end{bmatrix} \tag{13}$$

$$\begin{bmatrix}\boldsymbol{x}^T_1(k) & \boldsymbol{x}^T_2(k)\\ \boldsymbol{x}'^T_1(k) & \boldsymbol{x}'^T_2(k)\end{bmatrix}\begin{bmatrix}\Delta\hat{\boldsymbol{h}}_3(k)\\ \Delta\hat{\boldsymbol{h}}_4(k)\end{bmatrix} = \begin{bmatrix}e_2(k)\\ e'_2(k)\end{bmatrix}. \tag{14}$$

In this case, the echo path estimating filters converge to the true echo path impulse responses faster than when only the input-output relationships under one condition are used. This was also pointed out in Ref. [2]. However, it is difficult to simultaneously obtain the input-output relationships under two different conditions.

3.2. Imaginary input-output relationships for the reversed stereo signals

Our new idea is to use the input-output relationships for the reversed stereo signals $\boldsymbol{x}'_1(k) = \boldsymbol{x}_2(k)$ and $\boldsymbol{x}'_2(k) = \boldsymbol{x}_1(k)$. These relationships are usually unknown, so below we consider how to approximate them. Some hint is given by the following assumptions: $\boldsymbol{h}_1(k) \simeq \boldsymbol{h}_4(k)$ and $\boldsymbol{h}_2(k) \simeq \boldsymbol{h}_3(k)$, which indicate a particular situation where the loudspeakers and the microphones are arranged symmetrically. From these assumptions and from Eqs. (1) and (3), the input-output relationships for the reversed stereo signals are approximated as

$$y'_1(k) = \boldsymbol{x}^T_2(k)\boldsymbol{h}_1(k) + \boldsymbol{x}^T_1(k)\boldsymbol{h}_2(k) \simeq y_2(k) \tag{15}$$

$$y'_2(k) = \boldsymbol{x}^T_2(k)\boldsymbol{h}_3(k) + \boldsymbol{x}^T_1(k)\boldsymbol{h}_4(k) \simeq y_1(k). \tag{16}$$

Moreover, we introduce assumption error vectors $\boldsymbol{f}_1(k)$ and $\boldsymbol{f}_2(k)$ to compensate for the assumption errors, because the assumptions are not valid in actual situations. Thus the equations $\boldsymbol{h}_1(k) - \boldsymbol{h}_4(k) = \boldsymbol{f}_1(k)$ and $\boldsymbol{h}_2(k) - \boldsymbol{h}_3(k) = \boldsymbol{f}_2(k)$ are used. Therefore, instead of Eqs. (15) and (16), the imaginary echoes $y'_1(k)$ and $y'_2(k)$ for the reversed stereo signals of $\boldsymbol{x}_1(k)$ and $\boldsymbol{x}_2(k)$ can be described as

$$\begin{aligned} y'_1(k) &= \boldsymbol{x}^T_2(k)\boldsymbol{h}_1(k) + \boldsymbol{x}^T_1(k)\boldsymbol{h}_2(k)\\ &= y_2(k) + \boldsymbol{x}^T_2(k)\boldsymbol{f}_1(k) + \boldsymbol{x}^T_1(k)\boldsymbol{f}_2(k) \end{aligned} \tag{17}$$

$$\begin{aligned} y'_2(k) &= \boldsymbol{x}^T_2(k)\boldsymbol{h}_3(k) + \boldsymbol{x}^T_1(k)\boldsymbol{h}_4(k)\\ &= y_1(k) - \boldsymbol{x}^T_1(k)\boldsymbol{f}_1(k) - \boldsymbol{x}^T_2(k)\boldsymbol{f}_2(k). \end{aligned} \tag{18}$$

Denoting the replicas of the imaginary echoes $y'_1(k)$ and $y'_2(k)$ respectively as $\hat{y}'_1(k)$ and $\hat{y}'_2(k)$, the imaginary residual echoes $e'_1(k) = y'_1(k) - \hat{y}'_1(k)$ and $e'_2(k) = y'_2(k) - \hat{y}'_2(k)$ are

$$\begin{aligned} e'_1(k) &= \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{h}}_1(k) + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{h}}_2(k)\\ &= e_2(k) + \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{f}}_1(k) + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{f}}_2(k) \end{aligned} \tag{19}$$

$$\begin{aligned} e'_2(k) &= \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{h}}_3(k) + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{h}}_4(k)\\ &= e_1(k) - \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{f}}_1(k) - \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{f}}^T_2(k), \end{aligned} \tag{20}$$

where $\Delta\hat{\boldsymbol{h}}_1(k)$, $\Delta\hat{\boldsymbol{h}}_2(k)$, $\Delta\hat{\boldsymbol{h}}_3(k)$, and $\Delta\hat{\boldsymbol{h}}_4(k)$ are the error vectors of $\hat{\boldsymbol{h}}_1(k)$, $\hat{\boldsymbol{h}}_2(k)$, $\hat{\boldsymbol{h}}_3(k)$, and $\hat{\boldsymbol{h}}_4(k)$ from the true echo path impulse responses, and $\Delta\hat{\boldsymbol{f}}_1(k) = \boldsymbol{f}_1(k) - \{\hat{\boldsymbol{h}}_1(k) - \hat{\boldsymbol{h}}_4(k)\}$ and $\Delta\hat{\boldsymbol{f}}_2(k) = \boldsymbol{f}_2(k) - \{\hat{\boldsymbol{h}}_2(k) - \hat{\boldsymbol{h}}_3(k)\}$. Since $\Delta\hat{\boldsymbol{f}}_1(k)$ and $\Delta\hat{\boldsymbol{f}}_2(k)$ include unknowns $\boldsymbol{f}_1(k)$ and $\boldsymbol{f}_2(k)$, we cannot determine $e'_1(k)$ and $e'_2(k)$ exactly. Therefore we solve Eqs. (19) and (20) for $e_2(k)$ and $e_1(k)$ as follows.

$$\begin{aligned} e_2(k) = {} & \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{h}}_1(k) + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{h}}_2(k)\\ & - \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{f}}_1(k) - \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{f}}_2(k) \end{aligned} \tag{21}$$

$$\begin{aligned} e_1(k) = {} & \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{h}}_3(k) + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{h}}_4(k)\\ & + \boldsymbol{x}^T_1(k)\Delta\hat{\boldsymbol{f}}_1(k) + \boldsymbol{x}^T_2(k)\Delta\hat{\boldsymbol{f}}_2(k). \end{aligned} \tag{22}$$

Finally, from Eqs. (5), (6), (21), and (22), the following simultaneous equations are obtained for the new algorithm:

$$\begin{bmatrix} x_1^T(k) & x_2^T(k) & 0 & 0 & 0 & 0 \\ 0 & 0 & x_1^T(k) & x_2^T(k) & 0 & 0 \\ x_2^T(k) & x_1^T(k) & 0 & 0 & -x_2^T(k) & -x_1^T(k) \\ 0 & 0 & x_2^T(k) & x_1^T(k) & x_1^T(k) & x_2^T(k) \end{bmatrix} \begin{bmatrix} \Delta\hat{h}_1(k) \\ \Delta\hat{h}_2(k) \\ \Delta\hat{h}_3(k) \\ \Delta\hat{h}_4(k) \\ \Delta\hat{f}_1(k) \\ \Delta\hat{f}_2(k) \end{bmatrix}$$

$$= \left[e_1(k), e_2(k), e_2(k), e_1(k)\right]^T . \tag{23}$$

3.3. Extension with the projection algorithm

Our algorithm's effectiveness is independent of the stereo projection algorithm [3]. Therefore, we can extend Eq. (23) by using the concept of the projection algorithm [4]. The input stereo signal matrixes $X_1(k)$ and $X_2(k)$ are given by

$$X_1(k) = [x_1(k), x_1(k-1), \ldots, x_1(k-p+1)] \tag{24}$$

$$X_2(k) = [x_2(k), x_2(k-1), \ldots, x_2(k-p+1)], \tag{25}$$

and the error vectors $e_1(k)$ and $e_2(k)$ are denoted as

$$e_1(k) = [e_1(k),(1-\mu)e_1(k-1),\ldots,(1-\mu)^{p-1}e_1(k-p+1)]^T \tag{26}$$

$$e_2(k) = [e_2(k),(1-\mu)e_2(k-1),\ldots,(1-\mu)^{p-1}e_2(k-p+1)]^T \tag{27}$$

where p is the projection order and μ is the step size. Applying Eqs. (24) - (27) to Eq. (23), we obtain the following simultaneous equations.

$$\begin{bmatrix} X_1^T(k) & X_2^T(k) & 0 & 0 & 0 & 0 \\ 0 & 0 & X_1^T(k) & X_2^T(k) & 0 & 0 \\ X_2^T(k) & X_1^T(k) & 0 & 0 & -X_2^T(k) & -X_1^T(k) \\ 0 & 0 & X_2^T(k) & X_1^T(k) & X_1^T(k) & X_2^T(k) \end{bmatrix} \begin{bmatrix} \Delta\hat{h}_1(k) \\ \Delta\hat{h}_2(k) \\ \Delta\hat{h}_3(k) \\ \Delta\hat{h}_4(k) \\ \Delta\hat{f}_1(k) \\ \Delta\hat{f}_2(k) \end{bmatrix}$$

$$= \left[e_1^T(k), e_2^T(k), e_2^T(k), e_1^T(k)\right]^T . \tag{28}$$

3.4. Weighting parameters

We introduce several weighting parameters to Eq. (28) in order to control the estimating weight between the actual input-output relationships and the imaginary input-output relationships, and also to link it with the stereo projection algorithm and one kind of decorrelation algorithm [2]. As a result, Eq. (28) becomes

$$M^T(k) \begin{bmatrix} \Delta\hat{h}_1(k) \\ \Delta\hat{h}_2(k) \\ \Delta\hat{h}_3(k) \\ \Delta\hat{h}_4(k) \\ \Delta\hat{f}_1(k) \\ \Delta\hat{f}_2(k) \end{bmatrix} = e(k), \tag{29}$$

where

$$M(k) = \begin{bmatrix} X_1(k) & 0 & X_2(k) & 0 \\ X_2(k) & 0 & X_1(k) & 0 \\ 0 & X_1(k) & 0 & X_2(k) \\ 0 & X_2(k) & 0 & X_1(k) \\ 0 & 0 & -\alpha X_2(k) & \alpha X_1(k) \\ 0 & 0 & -\alpha X_1(k) & \alpha X_2(k) \end{bmatrix} \tag{30}$$

$$e(k) = \begin{bmatrix} e_1(k) \\ e_2(k) \\ \alpha\{e_2(k) - R_{12}R^{-1}e_1(k)\} + \beta R_{12}R^{-1}e_1(k) \\ \alpha\{e_1(k) - R_{12}R^{-1}e_2(k)\} + \beta R_{12}R^{-1}e_2(k) \end{bmatrix} \tag{31}$$

$$R = X_1^T(k)X_1(k) + X_2^T(k)X_2(k) \tag{32}$$

$$R_{12} = X_1^T(k)X_2(k) + X_2^T(k)X_1(k). \tag{33}$$

Then, the adjustment vectors $\Delta\hat{h}_1(k)$, $\Delta\hat{h}_2(k)$, $\Delta\hat{h}_3(k)$, and $\Delta\hat{h}_4(k)$ are obtained from the minimum-norm solution of Eq. (29), that is,

$$\begin{bmatrix} \Delta\hat{h}_1(k) \\ \Delta\hat{h}_2(k) \\ \Delta\hat{h}_3(k) \\ \Delta\hat{h}_4(k) \end{bmatrix} = \begin{bmatrix} X_1(k) & 0 \\ X_2(k) & 0 \\ 0 & X_1(k) \\ 0 & X_2(k) \end{bmatrix} \begin{bmatrix} q_1(k) \\ q_2(k) \end{bmatrix} + \begin{bmatrix} X_2(k) & 0 \\ X_1(k) & 0 \\ 0 & X_2(k) \\ 0 & X_1(k) \end{bmatrix} \begin{bmatrix} q_3(k) \\ q_4(k) \end{bmatrix}, \tag{34}$$

where

$$\begin{bmatrix} q_1(k) \\ q_2(k) \end{bmatrix} = \begin{bmatrix} R^{-1}e_1(k) \\ R^{-1}e_2(k) \end{bmatrix} - \alpha AB^{-1} \begin{bmatrix} e_2(k) - R_{12}R^{-1}e_1(k) \\ e_1(k) - R_{12}R^{-1}e_2(k) \end{bmatrix} + (1-\beta)AB^{-1} \begin{bmatrix} R_{12}R^{-1}e_1(k) \\ R_{12}R^{-1}e_2(k) \end{bmatrix} \tag{35}$$

$$\begin{bmatrix} q_3(k) \\ q_4(k) \end{bmatrix} = \alpha B^{-1} \begin{bmatrix} e_2(k) - R_{12}R^{-1}e_1(k) \\ e_1(k) - R_{12}R^{-1}e_2(k) \end{bmatrix} - (1-\beta)B^{-1} \begin{bmatrix} R_{12}R^{-1}e_1(k) \\ R_{12}R^{-1}e_2(k) \end{bmatrix} \tag{36}$$

$$A = \begin{bmatrix} R^{-1}R_{12} & 0 \\ 0 & R^{-1}R_{12} \end{bmatrix} \tag{37}$$

$$B = \begin{bmatrix} (1+\alpha^2)R - R_{12}R^{-1}R_{12} & -\alpha^2 R_{12} \\ -\alpha^2 R_{12} & (1+\alpha^2)R - R_{12}R^{-1}R_{12} \end{bmatrix}. \tag{38}$$

$\Delta\hat{f}_1(k)$ and $\Delta\hat{f}_2(k)$ do not have to be calculated explicitly. When $\alpha = 0$ and $\beta = 1$, the adjustment vectors are the same as those of the stereo projection algorithm, including the stereo NLMS algorithm. When $\alpha = 0$ and $\beta = 0$, the adjustment vectors are the same as those of one kind of decorrelation algorithm, which decorrelates $[X_1^T(k), X_2^T(k)]^T$ from $[X_2^T(k), X_1^T(k)]^T$. When $\alpha = 1$ and $\beta = 0$, the decorrelation of redundancy between the actual input-output relationships and the imaginary ones concerning $[X_1^T(k), X_2^T(k)]^T$ is introduced into Eq. (28). Thus, we can control the adaptation by choosing suitable values of α and β.

4. COMPUTER SIMULATIONS

In our computer simulations, we used 500 taps in each echo path estimating filter. The sampling frequency was 8 kHz. The input stereo signals were a single talker's speech recorded by two microphones. The performance was evaluated by the coefficient error between $[\boldsymbol{h}_1^T(k), \boldsymbol{h}_2^T(k)]^T$ and $[\hat{\boldsymbol{h}}_1^T(k), \hat{\boldsymbol{h}}_2^T(k)]^T$ which is given by

$$10\log_{10}\frac{||\boldsymbol{h}_1(k)-\hat{\boldsymbol{h}}_1(k)||^2+||\boldsymbol{h}_2(k)-\hat{\boldsymbol{h}}_2(k)||^2}{||\boldsymbol{h}_1(k)||^2+||\boldsymbol{h}_2(k)||^2}[\mathrm{dB}].$$

The true echo path impulse responses used here, which are shown in Fig. 2, were measured in a conference room with a reverberation time of 150 ms and with the loudspeakers and the microphones arranged roughly symmetrically. As shown in Figs. 2(e) and 2(f), there were some errors from the assumption: $\boldsymbol{h}_1(k) \simeq \boldsymbol{h}_4(k)$ and $\boldsymbol{h}_2(k) \simeq \boldsymbol{h}_3(k)$.

We show the results which were obtained with different values of α and β. Figures 3(a) and 3(c) correspond to the results with the stereo NLMS algorithm and the stereo projection algorithm respectively. Figure 3(d) corresponds to the results with the decorrelation algorithm, which decorrelates $[\boldsymbol{X}_1^T(k), \boldsymbol{X}_2^T(k)]^T$ from $[\boldsymbol{X}_2^T(k), \boldsymbol{X}_1^T(k)]^T$. It shows that the decorrelation method may improve the convergence to the true solution, although it has some unstability. Figure 3(b) corresponds to using Eq. (28). It shows that the imaginary input-output relationships are not used effectively in Eq. (28). Figure 3(e) corresponds to using Eq. (28) and the decorrelation of redundancy between the actual input-output relationships and the imaginary ones concerning $[\boldsymbol{X}_1^T(k), \boldsymbol{X}_2^T(k)]^T$. Figure 3(e) shows the best results of these shown in Fig. 3.

5. CONCLUSIONS

We proposed a new stereo echo cancellation algorithm, which is based on the idea of using the input-output relationships for the reversed stereo signals. This algorithm is linked to the stereo projection algorithm and the decorrelation algorithm. Computer simulations with data recorded in a conference room show that the imaginary input-output relationships can be used effectively to estimate the true echo path impulse responses by decorrelating redundancy between the actual and imaginary relationships.

ACKNOWLEDGMENTS

We would like to thank Dr. N. Kitawaki, Mr. J. Kojima, Dr. Y. Kaneda, Mr. Y. Haneda, and Mr. M. Tanaka for discussions and many suggestions. We also wish to express our gratitude to Dr. J. Benesty and Dr. A. Gilloire for fruitful discussions.

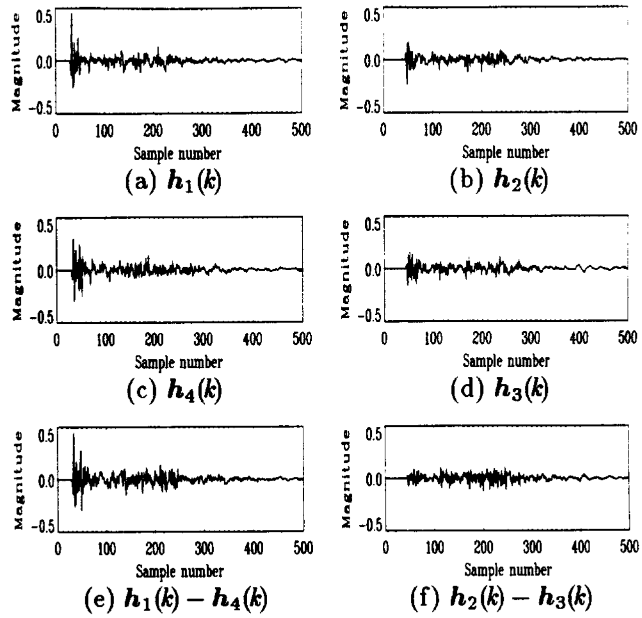

Figure 2: Acoustic echo path impulse responses used in simulations.

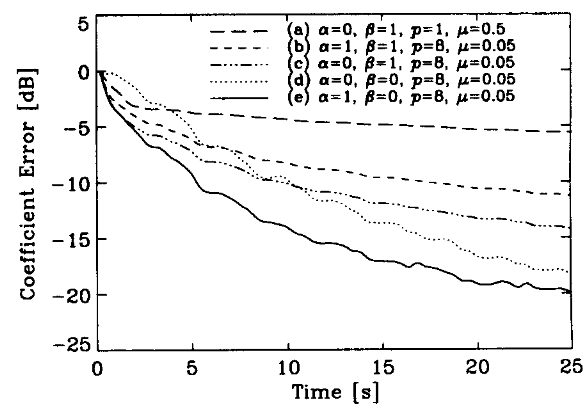

Figure 3: Coefficient error convergence.

REFERENCES

[1] J. Benesty, F. Amand, A. Gilloire, and Y. Grenier, "Adaptive Filtering Algorithms for Stereophonic Acoustic Echo Cancellation," *Proc. ICASSP95*, vol. 5, pp. 3099-3102 (1995).

[2] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic Acoustic Echo Cancellation -An Overview of the Fundamental Problem," *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 148-151 (1995).

[3] S. Shimauchi and S. Makino, "Stereo Projection Echo Canceller with True Echo Path Estimation," *Proc. ICASSP95*, vol. 5, pp. 3059-3062 (1995).

[4] S. Makino and Y. Kaneda, "Exponentially Weighted Step-size Projection Algorithm for Acoustic Echo Cancellers," *IEICE Trans.*, Vol. E75-A, No. 11, pp. 1500-1508, (1992).