A SOLUTION TO ECHO PATH IMBALANCE PROBLEM IN STEREO ECHO CANCELLATION

Satoru Emura, Yoichi Haneda, and Akitoshi Kataoka

NTT Cyber Space Laboratories, NTT Corporation
3-9-11, Midoricho, Musashino-shi, Tokyo 180-8585, Japan
{emura.satoru, haneda.yoichi, kataoka.akitoshi}@lab.ntt.co.jp

ABSTRACT

We propose a novel step-size control method that treats the echo path imbalance problem in stereo echo cancellation. The imbalance between two echo paths in magnitude slows down the convergence rate of the adaptive filter for stereo echo cancellation. To speed up the convergence rate, the proposed method extracts the information about filter-coefficient errors of each unknown echo path, and controls the step-size of each channel accordingly.

1. INTRODUCTION

For advanced full-duplex stereophonic teleconferencing, stereo echo cancellation has been studied [1–4]. Though the well-known nonuniqueness problem of multi-echo-path identification was overcome by preprocessing of the received stereo signal [1–3], an adaptive algorithm with a high convergence rate is still required [4].

In search of such an adaptive algorithm, the echo paths became recognized as a factor of slow convergence. When the two echo paths are very different in magnitude, the estimate of the echo path with low magnitude is too sensitive [4], and the convergence rate of the adaptive filter becomes slow. This problem occurs when a stereo echo canceller is combined with the directional microphone, which is equipped with most monaural hands-free conferencing systems.

As a solution of this problem, the update of each echo-path estimate with nonuniform step-size [5,6] can be considered. That is, adjusting the echo-path estimate with large errors by a large step-size, and the echo-path estimate with small errors by a small step-size. However, the difficulty lies in knowing the error of estimates of the unknown echo paths.

In this paper, we propose a novel step-size control method that treats the echo path imbalance problem. This algorithm has two features. One is that the noncausal portion is introduced in the adaptive filter and coefficient error of this portion is extrapolated to the information about the filter-coefficient-vector error of each channel. The other is that the step-size of each channel is controlled according to this information.

2. STEREO ECHO CANCELLER UNDER ECHO PATH IMBALANCE

A typical system for stereo-echo cancellation is shown in Fig. 1. Stereo signals $u_1(k)$ and $u_2(k)$, picked up by two microphones in the transmission room, are nonlinearly preprocessed as

$$x_1(k) = u_1(k) + g_1[u_1(k)],\ x_2(k) = u_2(k) + g_2[u_2(k)], \quad (1)$$

where k is the time index, and $g_1[\bullet]$ and $g_2[\bullet]$ are nonlinear functions for preprocessing [3]. The distortion due to this nonlinearity needs to be hardly perceptible.

The input signal vectors $\mathbf{x}_j(k)\,(j=1,2)$ are defined as

$$\begin{aligned}\mathbf{x}_j(k) &= \mathbf{u}_j(k) + \mathbf{u}_{gj}(k)\\ \mathbf{u}_j(k) &= \left[u_j(k) \quad \cdots \quad u_j(k-L+1)\right]^T \\ \mathbf{u}_{gj}(k) &= \left[g_j[u_j(k)] \quad \cdots \quad g_j[u_j(k-L+1)]\right]^T, \\ &(j=1,2)\end{aligned} \quad (2)$$

where L is the length of the adaptive-filter-coefficient vectors $\hat{\mathbf{h}}_1(k)$ and $\hat{\mathbf{h}}_2(k)$. The error $e(k)$ is expressed as

$$e(k) = y(k) - \hat{y}(k) \quad (3)$$

$$\hat{y}(k) = \hat{\mathbf{h}}_1^T(k)\mathbf{x}_1(k) + \hat{\mathbf{h}}_2^T(k)\mathbf{x}_2(k), \quad (4)$$

where $\hat{y}(k)$ is an echo replica.

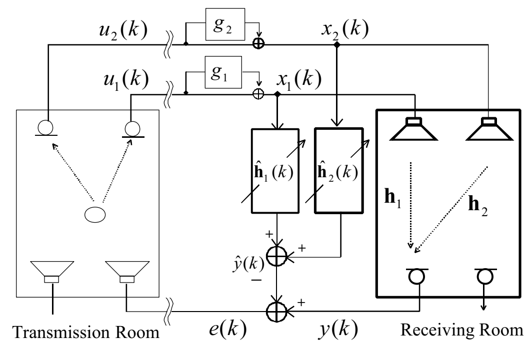

Fig.1 A system for stereo-echo cancellation.

2.1. Echo path imbalance problem

Here, we show that the imbalance of echo paths slows down the convergence rate of the adaptive filter.

The update of coefficient vectors is generally given, by using the direction of adjustment $\mathbf{z}_1(k)$ and $\mathbf{z}_2(k)$, as

$$\begin{aligned}\hat{\mathbf{h}}_1(k+1)&=\hat{\mathbf{h}}_1(k)+\ \mu(k)e(k)\mathbf{z}_1(k)\\ \hat{\mathbf{h}}_2(k+1)&=\hat{\mathbf{h}}_2(k)+\ \mu(k)e(k)\mathbf{z}_2(k).\end{aligned} \tag{5}$$

In the well-known normalized least-mean-squares (NLMS) algorithm, $\mathbf{z}_1(k)$ and $\mathbf{z}_2(k)$ are equal to $\mathbf{x}_1(k)$ and $\mathbf{x}_2(k)$. The step-size is given as

$$\mu(k)=\frac{\mu_0}{\mathbf{z}_1^T(k)\mathbf{x}_1(k)+\mathbf{z}_2^T(k)\mathbf{x}_2(k)+\delta}, \tag{6}$$

where μ_0 $(0<\mu_0<2)$ is a step-size parameter and δ is a small constant for regularization.

Consider the case where the ratio of the echo path magnitude is given as $\|\mathbf{h}_1\|\approx 9\|\mathbf{h}_2\|$ and the ratio of input signal powers is given as $\|\mathbf{x}_1(k)\|^2\approx\|\mathbf{x}_2(k)\|^2$. We see

$$\|\hat{\mathbf{h}}_1(k+1)-\hat{\mathbf{h}}_1(k)\|\approx\|\hat{\mathbf{h}}_2(k+1)-\hat{\mathbf{h}}_2(k)\| \tag{7}$$

for the increment vectors of the NLMS algorithm.

At the initial state ($\hat{\mathbf{h}}_1(k)=\hat{\mathbf{h}}_2(k)\approx\mathbf{0}$), the error of the estimate $\|\mathbf{h}_1-\hat{\mathbf{h}}_1(k)\|$ is nearly 9 times as large as $\|\mathbf{h}_2-\hat{\mathbf{h}}_2(k)\|$. However, this difference is not reflected in the increments of $\hat{\mathbf{h}}_1(k)$ and $\hat{\mathbf{h}}_2(k)$. Hence, $\hat{\mathbf{h}}_2(k)$ is updated by a relatively large increment vector and its convergence becomes slow. On the other hand, when the step-size parameter μ_0 is set small, the convergence of $\hat{\mathbf{h}}_1(k)$ becomes too slow.

The same phenomena occur in the conventional two-channel affine projection algorithm with low order [2], where the direction of adjustment $\mathbf{z}_1(k)$ is given as a linear combination of multiple delayed input signal vectors of $\mathbf{x}_1(k)$, and $\mathbf{z}_2(k)$ as those of $\mathbf{x}_2(k)$.

To overcome this problem, the update of each echo-path estimate with nonuniform step-size [5,6] can be considered. That is, adjusting the echo-path estimate with large errors by a large step-size, and the echo-path estimate with small errors by a small step-size.

However, the difficulty lies in knowing how large these errors are. It seems impossible to know the error of estimates of the unknown echo paths.

3. PROPOSED ALGORITHM

We propose a novel step-size control method, in which we extend the *generalized weighting function* technique [7] (the *delay coefficients method* in [8]) to the multichannel case for extracting information about the error of estimates of unknown echo paths and apply this information to the control of the step-size.

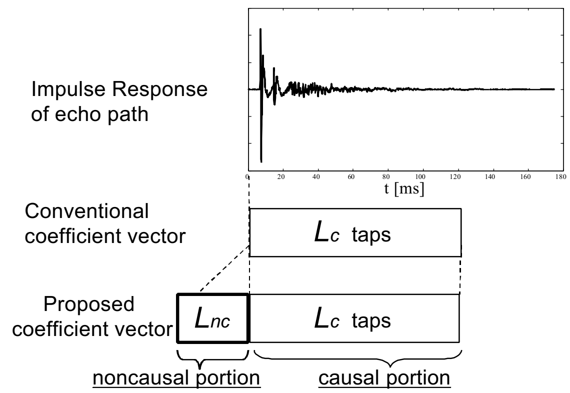

Fig.2 Impulse response of an echo path, the corresponding filter-coefficient vectors in the conventional adaptive algorithms and the proposed method.

3.1. Misalignment estimation via noncausal portion

In the conventional two-channel adaptive algorithms, the adaptive filter estimates the causal portion of the impulse responses of the echo paths. We firstly propose to add a noncausal portion to the multichannel adaptive filter. Since the true coefficients corresponding to the noncausal portion are known to be zero, this known portion of the coefficient error vector can be extrapolated to the coefficient error of the causal portion [7].

For that purpose, we introduce an artificial delay of L_{nc} taps into the microphone signal, and estimate echo paths from this delayed microphone signal and the input signal vectors. The filter-coefficient vector of the j-th $(j=1,2)$ channel is given as

$$\hat{\mathbf{h}}_j=\ [\underbrace{\hat{h}_j(-L_{nc})\cdots\hat{h}_j(-1)}_{\textit{Noncausal portion}}\ \ \underbrace{\hat{h}_j(0)\cdots\hat{h}_j(L_c-1)}_{\textit{Causal portion}}]^T, \tag{8}$$

where the noncausal and causal portion have L_{nc} and L_c taps as shown in Fig. 2. The echo replica $\hat{y}(k)$ is generated as

$$\hat{y}(k)=\hat{\mathbf{h}}_1^T(k)\begin{bmatrix}x_1(k)\\ \vdots\\ x_1(k-L_{nc})\\ \vdots\\ x_1(k-L_{nc}-L_c+1)\end{bmatrix}+\hat{\mathbf{h}}_2^T(k)\begin{bmatrix}x_2(k)\\ \vdots\\ x_2(k-L_{nc})\\ \vdots\\ x_2(k-L_{nc}-L_c+1)\end{bmatrix}. \tag{9}$$

The error signal is given as the difference between $\hat{y}(k)$ and the delayed microphone signal $y(k-L_{nc})$.

$$e(k)=y(k-L_{nc})-\hat{y}(k) \tag{10}$$

We use the following function $\sigma[\bullet]$ to extract the information from coefficient error of the noncausal portion.

$$\sigma[\hat{\mathbf{h}}_j]=\sum_{n=-L_{nc}}^{-1}\left|\hat{h}_j(n)\right|^2 \tag{11}$$

3.2. Step-size control

We secondly propose to control the step-size for each echo path estimate according to the information extracted from the noncausal portion of the adaptive filter.

The filter-coefficient vector corresponding to each echo path is updated as

$$\hat{\mathbf{h}}_j(k+1) = \hat{\mathbf{h}}_j(k) + \mu_j(k)e(k)\mathbf{z}_j(k) \;\; (j=1,2) \qquad (12)$$

where the step-size for each channel $\mu_j(k)$ is given as

$$\mu_j(k) = \frac{\mu_0 \, \sigma[\hat{\mathbf{h}}_j(k)]}{\|\mathbf{z}_j(k)\| \sum_{i=1}^{2} \sigma[\hat{\mathbf{h}}_i(k)]\mathbf{z}_i^T(k)\mathbf{x}_i(k) / \|\mathbf{z}_i(k)\|}. \qquad (13)$$

This control of the step-size is obtained from the following two constraints

$$\Delta\hat{\mathbf{h}}_1^T(k)\mathbf{x}_1(k) + \Delta\hat{\mathbf{h}}_2^T(k)\mathbf{x}_2(k) = \mu_0 \, e(k) \qquad (14)$$

and

$$\frac{\|\Delta\hat{\mathbf{h}}_1(k)\|}{\|\Delta\hat{\mathbf{h}}_2(k)\|} = \frac{\mu_1(k)e(k)\|\mathbf{z}_1(k)\|}{\mu_2(k)e(k)\|\mathbf{z}_2(k)\|} = \frac{\sigma[\hat{\mathbf{h}}_1(k)]}{\sigma[\hat{\mathbf{h}}_2(k)]}, \qquad (15)$$

where $\Delta\hat{\mathbf{h}}_j(k) = \hat{\mathbf{h}}_j(k+1) - \hat{\mathbf{h}}_j(k)$.

Note that the power imbalance of stereo input signal [9] is already taken into account because the power imbalance is reflected to $\|\mathbf{z}_1(k)\|$ and $\|\mathbf{z}_2(k)\|$.

The proposed method can be easily generalized to the multichannel case. Note that we can eliminate the artificial delay introduced for estimating the noncausal portion by using an echo canceller with two echo path models [10]. This echo canceller uses the first echo path model for adaptively estimating the echo path and the second model for synthesizing an echo replica to cancel out the echo. This echo canceller can eliminate the artificial delay by removing the noncausal portion from the second model and copying only the causal portion from the first to the second model.

4. SIMULATION

We confirmed the validity of the proposed method through computer simulation. The signal source $\boldsymbol{s}$ in the transmission room was a speech signal. The two microphone signals in the transmission room were obtained by convolving $\boldsymbol{s}$ with two impulse responses of 700 taps, which were measured in an actual room and scaled to have the same magnitude. The microphone output signal $y(k)$ in the receiving room was obtained by summing the two convolutions $(h_1 * x_1)$ and $(h_2 * x_2)$, where h_1 and h_2 were also measured in an actual room and were truncated to 700 taps. A white-noise signal was added to $y(k)$ as ambient noise to simulate 45-dB SNR. The sampling frequency was 8 kHz. We applied a half-wave rectifier nonlinearity [3]

$$g_1[u] = \alpha(u + |u|)/2, \;\; g_2[u] = \alpha(u - |u|)/2 \qquad (16)$$

with $\alpha = 0.3$ to the received stereo signal.

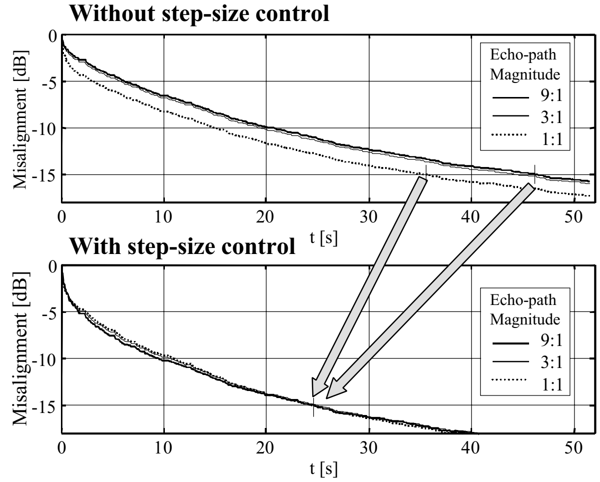

Fig. 3 Behavior of the misalignment of the frequency-domain two-channel LMS adaptive algorithm with and without the proposed method of step-size control, under three conditions in which echo-path magnitude ratios were 9:1, 3:1, and 1:1.

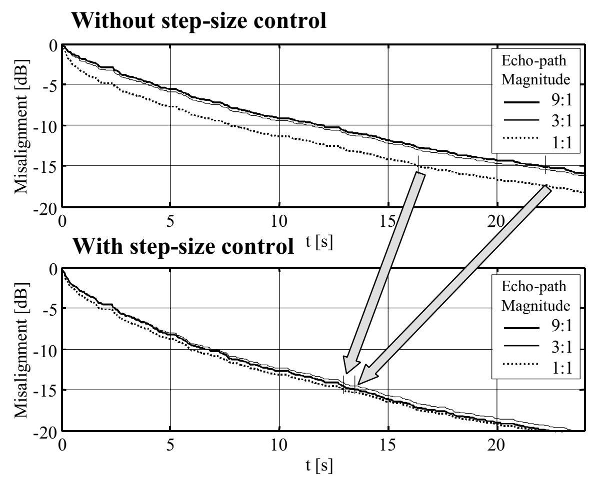

Fig. 4 Behavior of the misalignment of the frequency-domain two-channel enhanced adaptive algorithm (attenuation factor $\beta = 0.1$) with and without the proposed method of step-size control, under three conditions in which echo-path magnitude ratios were 9:1, 3:1, and 1:1.

We assigned L_c=512 taps and L_{nc}=0 taps for the causal and noncausal portions of the adaptive filter in the conventional method. We assigned L_c=512 taps and L_{nc}=128 taps for the causal and noncausal portions in the proposed method. In both methods, we used the step-size parameter $\mu_0 = 0.5$ and updated the adaptive filter every 512 samples.

We checked the effect of the proposed method when it is combined with the frequency-domain adaptive algorithm derived by straightforwardly extending the monaural constrained FLMS [11] to two channel, where

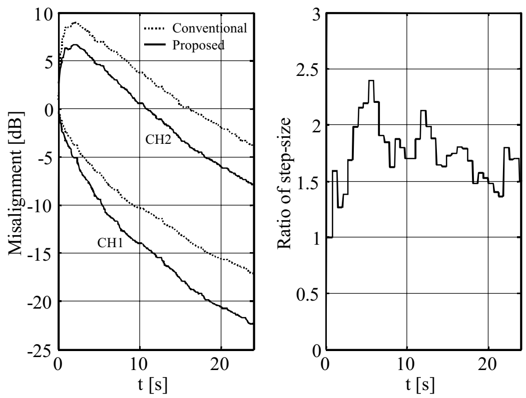

Fig.5 Behavior of the misalignment of each channel and the ratio of step-size where the echo-path magnitude ratio was 9:1 and the frequency-domain two-channel enhanced adaptive algorithm (attenuation factor $\beta = 0.1$) was used.

filter coefficients in the frequency domain are updated as in the two-channel NLMS algorithm. Step-size μ_j in each frequency bin is controlled every frame by using $\sigma[\hat{\mathbf{h}}_j]$ which is updated every 10 frame from the estimated impulse responses in the time domain. We measured its convergence by using the misalignment defined as

$$\frac{\| \mathbf{h}_1 - \hat{\mathbf{h}}_1(k) \|^2 + \| \mathbf{h}_2 - \hat{\mathbf{h}}_2(k) \|^2}{\| \mathbf{h}_1 \|^2 + \| \mathbf{h}_2 \|^2} \tag{17}$$

under three conditions in which the echo-path magnitude ratios $\| \mathbf{h}_1 \| : \| \mathbf{h}_2 \|$ were set to 9:1, 3:1, and 1:1. We used the time spent for the misalignment to reach –15dB (T_{-15}) as the performance index.

The behavior of the misalignment is shown in Fig. 3. It can be seen that the proposed method improved the convergence rate in all three situations. In addition, the proposed method reduced the difference of the convergence rate caused by echo path imbalance. When the echo-path magnitude ratio was 9:1, T_{-15} was reduced to 56%. When this magnitude was 1:1, T_{-15} was reduced to 68%.

Next, we checked the proposed method when it is combined with the constrained frequency-domain two-channel enhanced adaptive algorithm [12]. The behavior of the misalignment is shown in Fig. 4, where the attenuation factor of the enhanced adaptive algorithm was set to $\beta = 0.1$. It can be seen that the proposed method improved the convergence rate in all three situations and reduced the difference of the convergence rate caused by echo path imbalance. When the echo-path magnitude ratio was 9:1, T_{-15} was reduced to 61%. When this magnitude was 1:1, T_{-15} was reduced to 76%.

Figure 5 shows the relative filter-coefficient error of each channel and the ratio of step-size $\mu_1(k)/\mu_2(k)$ where the frequency-domain enhanced adaptive algorithm was used and echo-path magnitude was 9:1. In both channels, the relative filter-coefficient errors, i.e., $\| \mathbf{h}_1 - \hat{\mathbf{h}}_1(k) \|^2 / \| \mathbf{h}_1 \|^2$ and $\| \mathbf{h}_2 - \hat{\mathbf{h}}_2(k) \|^2 / \| \mathbf{h}_2 \|^2$ were improved.

6. SUMMARY

We have proposed a novel step-size control method that treats the echo path imbalance problem in stereo echo cancellation. Computer simulations demonstrated that this method improves the convergence rate of misalignment when it is combined with the frequency-domain two-channel LMS algorithm and the enhanced adaptive algorithm.

REFERENCES

[1] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic Acoustic Echo Cancellation – An Overview of the Fundamental Problem," *IEEE Signal Processing Letters*, 2, pp. 148–151, 1995.

[2] S. Shimauchi and S. Makino, "Stereo Projection Echo Canceller with True Echo Path Estimation," *Proc. ICASSP95*, pp. 3059–3062, 1995.

[3] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A Better Understanding and an Improved Solution to the Specific Problems of Stereophonic Acoustic Echo Cancellation," *IEEE Trans. Speech Audio Processing*, 6, pp. 156–165, 1998.

[4] T. Gaensler and J. Benesty, "Multichannel Acoustic Echo Cancellation: What's new?," *Proc. IWAENC2001*, pp. 14–19, 2001.

[5] S. Makino, Y. Kaneda, and N. Koizumi, "Exponentially Weighted Stepsize NLMS Adaptive Filter Based on the Statistics of a Room Impulse Response," *IEEE Trans. Speech Audio Processing*, 1, pp. 101–108, 1993.

[6] D. L. Duttweiler, "Proportionate Normalized Least-Mean-Squares Adaptation in Echo Cancelers," *IEEE Trans. Speech Audio Processing*, 8, pp. 508–518, 2000.

[7] A. Noda, "A Generalized Weighting Function with Pseudo-Predictive Portion," *IEEE Trans. Automatic Control*, 19, pp. 618–619, 1974.

[8] C. Breining, et. al., "Acoustic Echo Control," *IEEE SP Mag.*, 16, 4, pp. 42–69, 1999.

[9] A. Nakagawa and Y. Haneda, "A Study of an Adaptive Algorithm for Stereo Signals with a Power Difference," *Proc. ICASSP2002*, pp. 1914–1916, 2002

[10] K. Ochiai, T. Araseki, and T. Ogihara, "Echo Canceler with Two Echo Path Models," *IEEE Trans. on Communications*, 25, 6, pp. 589–595, 1977.

[11] E. R. Ferrara, Jr., "Frequency-Domain Adaptive Filtering," *Adaptive Filters*, C. F. N. Cowen and P. M. Grant, Eds, Englewood Cliffs, NJ: Prentice Hall, pp. 145–179, 1985.

[12] S. Emura, Y. Haneda, and S. Makino, "Enhanced Frequency-Domain Adaptive Algorithm for Stereo Echo Cancellation," *Proc. ICASSP2002*, pp. 1901–1904, 2002.