ACOUSTIC ECHO CANCELLATION USING SUB-ADAPTIVE FILTER

Satoshi Ohta, Yoshinobu Kajikawa, Yasuo Nomura

Faculty of Engineering, Kansai University

ABSTRACT

In this paper, we propose an acoustic echo cancellation (AEC) using a sub-adaptive filter. In the AEC, the step-size parameter of the adaptive filter must be varied according to the situations where a double talk and an echo path change occur. The proposed AEC can appropriately control the step-size parameter even if the double talk and the echo path change simultaneously occur because the optimal step-size parameter can be obtained according to the output of the sub-adaptive filter and the echo path change detector is controlled through the double talk detector. Hence, the proposed AEC can realize superior convergence property to the conventional one. Simulation results demonstrate that the proposed AEC can achieve higher ERLE and faster convergence than the conventional one.

***Index Terms*—** Acoustic echo cancellation, sub-adaptive filter, double-talk detection, echo path change detector

1. INTRODUCTION

In hands-free telecommunication systems talking through a loudspeaker and a microphone, the acoustic echo deteriorates the speech quality, and then the acoustic echo canceller(AEC) is used as one of technologies that remove the echo [1, 2]. The principle of the AEC is to estimate the impulse response of an acoustic echo path using an adaptive filter, generate a pseudo-echo, and subtract it from the echo. However, the acoustic echo path changes frequently and the microphone picks up near-end talker's signal and ambient noise in addition to the echo in the application to real environments. In this case, it is necessary to control the step-size parameter according to the situations and determine whether the double-talk occurs or the echo path changes. For this purpose, the step-size control using Fuzzy control has been already proposed [3]. Moreover, we have already proposed a technique using the sub-adaptive filter(Sub-ADF) which assists the acoustic echo canceller to control the step-size parameter [4]. However, it is impossible to control the step-size parameter appropriately when the double-talk and the echo path change simultaneously occur because the Sub-ADF and the main adaptive filter(ADF) differ in the convergence speed. In this paper, we therefore propose an improvement of the AEC using Sub-ADF where the initial coefficients of Sub-ADF are modified and a double-talk detector is introduced into the echo path change detector.

2. STEP-SIZE CONTROL USING SUB-ADF

2.1. Optimal step-size parameter

An optimal step-size parameter when the input signal of the adaptive filter is white Gaussian noise is expressed as follows [2, 5]:

$$\alpha_{opt}(n) = \frac{E\left\{\varepsilon^2(n)\right\}}{E\left\{e^2(n)\right\}} = \frac{E\left\{\varepsilon^2(n)\right\}}{E\left\{\varepsilon^2(n)\right\} + E\left\{w^2(n)\right\}} \quad (1)$$

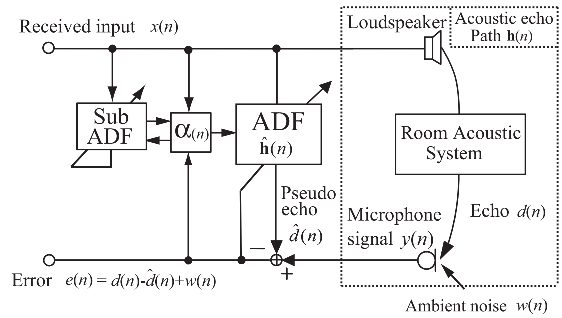

Fig. 1. Structure of the AEC.

where $e(n)$ is the error signal and $\alpha_{opt}(n)$ is decreased with time so that an estimate error $\varepsilon(n)$ becomes small. In addition, if the acoustic echo path changes, the AEC can track the change because the step-size parameter is increased according to the increase of the estimate error. On the other hand, the step-size parameter is suddenly decreased when the ambient noise $w(n)$ is increased due to double talking. However, it is impossible to apply (1) to the real system because the estimate error $\varepsilon(n)$ is unknown in the real system and the expectation operation is used in (1). Hence, the estimate error must be approximately calculated in some way.

2.2. Acoustic echo canceller using Sub-ADF

In order to obtain an approximation of the estimate error, it is necessary to use an adaptive filter which is independent of ambient noise. Therefore, we introduce the Sub-ADF separately from the main adaptive filter(ADF). The composition of the AEC with the Sub-ADF is shown in Figure 1. In Fig. 1, the desired signal of the Sub-ADF is set to zero and the coefficients of the Sub-ADF are initialized to nonzero values. In a word, the coefficients of the Sub-ADF converge from nonzero to zero as the update advances. As a result, the optimal step-size parameter can be approximately calculated as follows:

$$\alpha(n) = \frac{Py_{sub}(n)}{Pe(n)} \quad (2)$$

$$Py_{sub}(n) = \gamma Py_{sub}(n-1) + (1-\gamma)\left\{\hat{\mathbf{h}}_{\mathrm{sub}}(n)^T \mathbf{x}(n)\right\}^2 \quad (3)$$

$$Pe(n) = \gamma Pe(n-1) + (1-\gamma)e(n)^2 \quad (4)$$

$\hat{\mathbf{h}}_{\mathrm{sub}}(n)$: Coefficients of Sub-ADF
N_{sub} : Tap length of Sub-ADF
$\gamma = 0.95$

The step-size parameter of the ADF is always optimized according to (2) where the estimate error of the Sub-ADF is substituted for that of ADF.

2.3. Initial coefficients of Sub-ADF

In this method, the accuracy of step-size control can be improved as the initial value of the Sub-ADF is close to the impulse response of the actual acoustic echo path. Although the acoustic echo path is unknown unless it is measured, it is generally attenuated exponentially. Thus, the initial coefficients of the Sub-ADF are also attenuated exponentially. In particular, the rate of the attenuation is computed from an approximation of the reverberation time, and then the initial coefficients are computed by attenuating random values based on this decay rate.

3. ECHO PATH CHANGE DETECTOR

3.1. Principle of echo path change detector

However, the Sub-ADF cannot track the echo path change because the Sub-ADF is independent of the acoustic echo path. Hence, the step-size parameter becomes small even if the echo path changes. It is consequently necessary to detect the echo path change and then initialize the coefficients of the Sub-ADF. In ordinary rooms, the acoustic impulse responses exhibit exponential decay properties [6]. In a similar fashion, the amount of echo path change also exhibits exponential decay properties. Hence, it is thought that the ADF coefficients also attenuate exponentially during their updates. Therefore, we define the amount of the change of the ADF coefficients, called change parameter $\mathrm{CP}(k)$, as follows:

$$\mathrm{CP}(k) = \sum_{n=kJ+1}^{(k+1)J} \frac{\sum_{i=1}^{N/4} \Delta\hat{h}_i(n)^2}{J \sum_{i=1}^{N} \Delta\hat{h}_i(n)^2} \tag{5}$$

$\Delta\hat{h}_i(n)$: Update amount of ADF
J : Detection interval

The change parameter $\mathrm{CP}(k)$ grows large when the echo path changes because the echo path change has great influence on the first quarter of the ADF coefficients. Hence, if the change parameter $\mathrm{CP}(k)$ becomes large, that is, the echo path changes, then the coefficients of the Sub-ADF are initialized, so that the ADF can track the echo path change quickly. Concretely, the echo path change is detected according to whether $\mathrm{CP}(k)$ is more than a threshold ε_1.

3.2. Handling of double-talk

The echo path change detector shown in 3.1 can accurately detect the change under single-talk conditions, but not under double-talk conditions. This is because the change parameter $\mathrm{CP}(k)$ also becomes large due to the near-end talker's signal. It is necessary to stop detecting the echo path change at the time of double talking to prevent false detection. In this paper, the double-talk is detected by the correlation between the pseudo echo signal $\hat{d}(n)$ and the microphone signal $y(n)$. First of all, the sum of products of the pseudo echo signal and the microphone signal is calculated as follows:

$$\begin{aligned} P_{dy}(k) &= \sum_{n=kJ+1}^{(k+1)J} \hat{d}(n)y(n) \\ &= \sum_{n=kJ+1}^{(k+1)J} \hat{d}(n)\left\{d(n)+w(n)\right\} \end{aligned} \tag{6}$$

The sum of products makes the effect of ambient noise $w(n)$ small, and then $P_{dy}(k)$ is close to the power of the pseudo echo signal because $w(n)$ is uncorrelated to the echo signal under single-talk and the steady state of the ADF. On the other hand, the change of $P_{dy}(k)$ becomes large under double-talk because the near end talker's signal is nonstationary. Now, $P_{dy}(k)$ is normalized by the power of the pseudo echo signal as follows:

$$P_{DT}(k) = \frac{P_{dy}(k)}{\sum_{n=kJ+1}^{(k+1)J} \hat{d}(n)^2} \tag{7}$$

The detection parameter $P_{DT}(k)$ remains at about '1' during the single-talk, while indicates some different values from '1' during the double-talk. Hence, the double-talk is detected according to $P_{DT}(k)$, that is, whether $|P_{DT}(k) - 1.0|$ is more than a threshold ε_2.

4. TAP LENGTH OF SUB-ADF

The proposed method uses two adaptive filters and updates the filter coefficients by NLMS algorithm. Hence, the computational complexity for updating the two ADFs is $6N$. In addition, the complexity for calculation of equation (5) is N, and that of equations (6) and (7) is small. Therefore, the computational complexity of the proposed method is $7N$. Thus, the complexity for updating the Sub-ADF increases as the number of taps N becomes large. Hence, the number of taps of the Sub-ADF must be reduced so that the approximation accuracy of the estimate error doesn't deteriorate.

However, the convergence speed of the ADF becomes slow as the tap length becomes large. If the tap length of the Sub-ADF is made smaller than that of the ADF, the Sub-ADF will converge more quickly than the ADF. In other words, we cannot control the step-size parameter optimally because of the difference of convergence rate between Sub-ADF and ADF. According to the ratio of the tap length between Sub-ADF and ADF, the step-size parameter of the Sub-ADF is therefore calculated as follows:

$$\mathrm{SSP}_{sub}(k) = \mathrm{SSP}_{\mathrm{ADF}}(k) \times \frac{N_{sub}}{N} \tag{8}$$

where $\mathrm{SSP}_{\mathrm{ADF}}(k)$ is the step-size parameter of the ADF, and N and N_{sub} are the tap lengths of ADF and Sub-ADF, respectively. Hence, the step-size parameter can be controlled appropriately, even if the tap length of the Sub-ADF is small.

5. SIMULATION RESULTS

In this chapter, some simulation results demonstrate the effectiveness of the proposed AEC. The block diagram of the proposed AEC is shown in Figure 2 and the simulation conditions are shown in Table 1. The acoustic echo path used for the simulation was measured in a conference room whose reverberation time is 300ms according to the measurement procedure recommended by ITU-T-P34 as shown in Figure 3. The echo path change situation is the movement from Measurement 1 to Measurement 2 in Fig. 3. Simulation results are shown as

$$\mathrm{ERLE} = 10 \cdot \log_{10} \frac{\sum \left\{d(n)\right\}^2}{\sum \left\{d(n) - \hat{d}(n)\right\}^2} \tag{9}$$

Table 1. Simulation conditions.

Input signal	Voice (female)
Noise signal (Single talk)	White noise
Noise signal (Double talk)	Voice (male)
SNR (Single talk)	18 [dB]
SNR (Double talk)	-5 [dB]
Tap length of ADF	2048
Tap length of Sub-ADF	2048
Threshold ε_1	0.6
Threshold ε_2	0.5
Update algorithm	NLMS

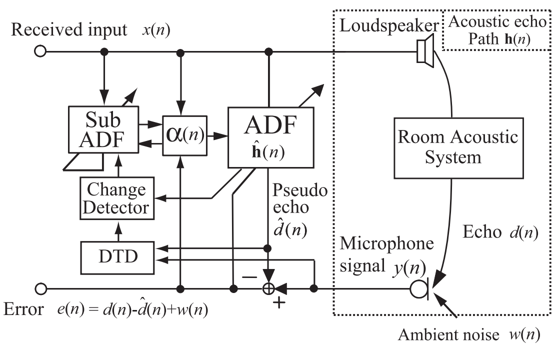

Fig. 2. Structure of the proposed AEC.

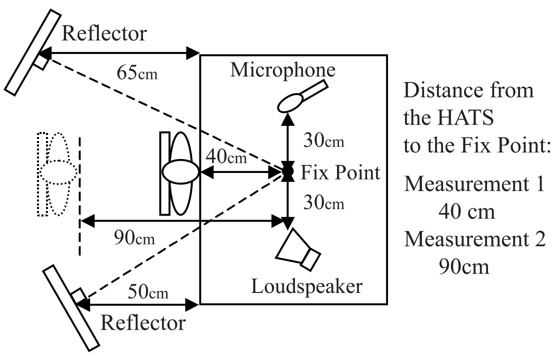

Fig. 3. Measurement setup for the acoustic echo path.

For comparison of a result, the method using Fuzzy control is shown as a conventional method. In addition, the parameters of the conventional method are set so that the best convergence property is obtained.

5.1. Single-talk situation

The results under the single-talk are shown in Figure 4. From Fig. 4, the proposed AEC can improve the convergence speed than the conventional one. This is because the step-size parameter is controlled appropriately in accordance with the convergence property of the ADF. Moreover, the conventional method sets the step-size parameter comparatively large, even after the update progresses in order to prepare for the echo path change. Hence, the accuracy of the conventional method does not advance. In contrast, since the proposed method uses the echo path change detector, it is not necessary to perform such processing. From these reasons, it is thought that the proposed method has superior performance to the conventional one.

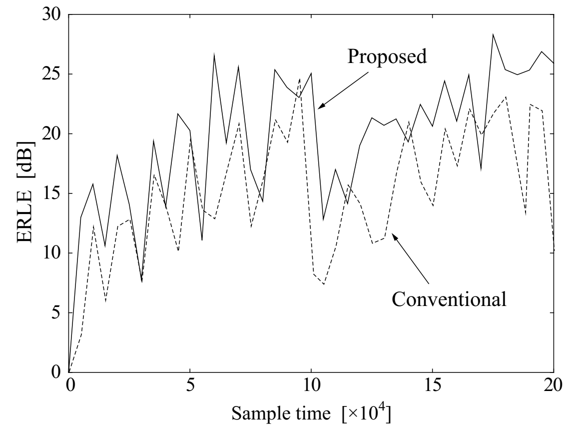

Fig. 4. Convergence properties of ERLE(single-talk).

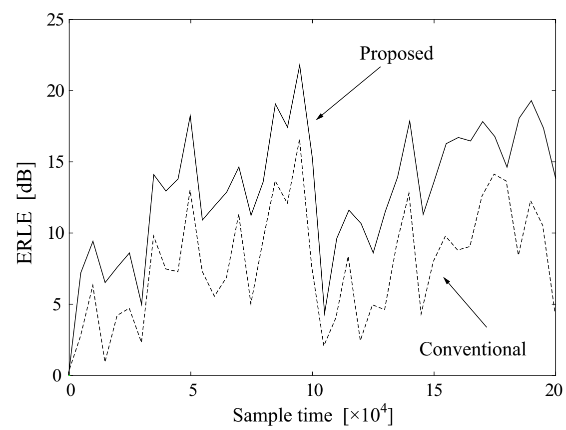

Fig. 5. Convergence properties of ERLE(double-talk).

5.2. Double-talk situation

Figure 5 shows simulation results under the double-talk. It can be seen from Fig. 5 that the proposed AEC can achieve higher ERLE and converge faster than the conventional one. This is because the proposed AEC can prevent misjudgments of the echo path change using the double talk detector. On the other hand, since the conventional method sets the step-size parameter comparatively large, it is greatly influenced by the near-end signal. From the above results, the effectiveness of the proposed AEC is demonstrated.

5.3. Effectiveness of the double-talk detector

Figure 6 shows the transition of the double talk detection parameter. In Fig. 6, the near-end talker's signal is shown together in order to demonstrate the effectiveness of the double talk detector. It can be seen from Fig. 6 that the detection parameter remains at about 1.0 under single-talk situations (i.e. no near-end talker's signal). On the other hand, the detection parameter changes greatly under double-talk situations. Hence, it is understood that the proposed double talk detector can detect the double-talk situation exactly.

Figure 7 shows the results of the echo path change and double talk detection. The threshold value of the echo path change detector

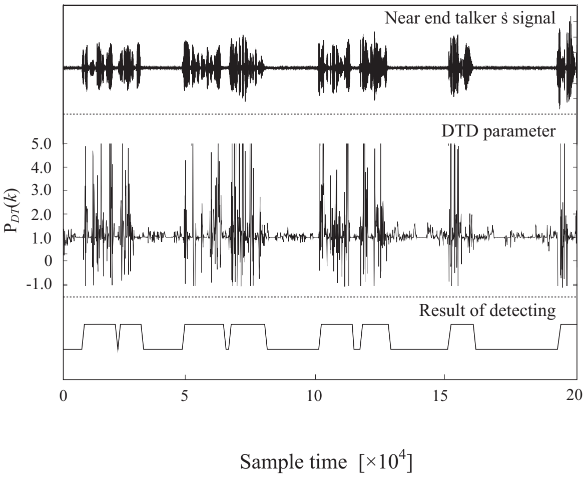

Fig. 6. Voice waveform and characteristic of double talk detector.

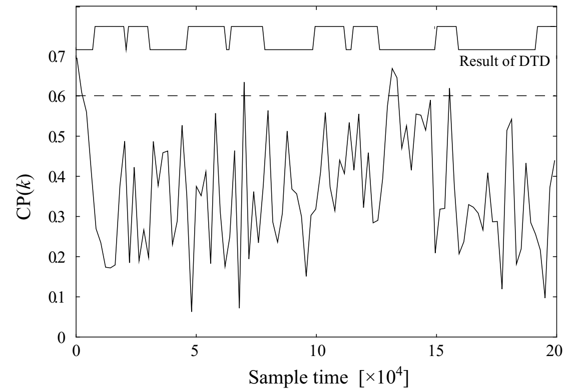

Fig. 7. Detection results of echo path change and double talk.

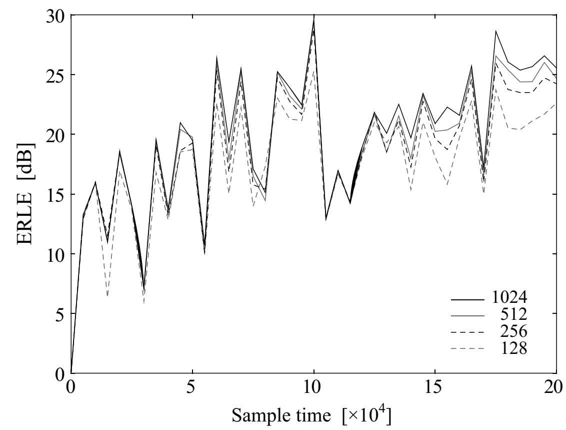

Fig. 8. Comparison of the Sub-ADF taps.

is 0.6. In Fig. 7, the proposed method is not influenced even if the echo path change parameter exceeds the threshold during the double talk, because the double talk detector prohibits the initialization of the Sub-ADF.

The echo path change parameter is increased at the time of about 130,000 samples which the double talk ends and then the echo path change is detected. From these results, the proposed method can track the echo path change as soon as possible, while suppressing the influence of the double talk.

5.4. Tap Length of Sub-ADF

In this section, we examine the influence of the number of taps of the Sub-ADF on the convergence property. Figure 8 shows simulation results in the case where the number of taps of the Sub-ADF is set to 512, 256, 128, and 32 taps, respectively. The step-size parameter of the Sub-ADF is computed by the equation (8).

It is found from Fig. 8 that almost the same convergence property is obtained in cases of 512, 256 and 128 taps. On the other hand, the convergence property in case of 64 taps deteriorates overall compared with the other taps. This is because the error between the initial value of the Sub-ADF and the actual echo path becomes large, as the number of taps of the Sub-ADF becomes small.

Next, It is found from Figure 8 that the degradation of the convergence property is small in case of 1024, 512 and 256. However, the convergence property in case of 128 taps deteriorates overall. From these results, it turned out the number of taps can be reduced to 1/4 of that of the ADF without the degradation of the performance.

6. CONCLUSIONS

In this paper, we have proposed an AEC using the Sub-ADF. At first, the coefficients of the Sub-ADF have been initialized to have an exponential decay property. In addition, the double talk detector has been introduced to prevent misjudgment of the echo path change under double-talk situations. The simulations results have demonstrated that the proposed AEC can achieve higher ERLE and converge faster than the conventional one under both single-talk and double-talk situations. Moreover, we have examined the appropriate tap length of the Sub-ADF for obtaining good convergence speed. The results have shown that the number of taps of the Sub-ADF can be reduced by a quarter of that of the ADF. Hence, the computational complexity of the proposed method can be reduced.

7. REFERENCES

[1] Y. A. Huang and J. Benesty, *Audio Signal Processing*, Kluwer Academic Publishers, London, 2004.

[2] C. Breining and *et al.*, "Acoustic echo control," *IEEE Signal Processing Magazine*, vol. 16, no. 4, pp. 42–69, Jul. 1999.

[3] C. Breining, "A robust fuzzy logic-based step-gain control for adaptive filters in acoustic echo cancellation," *IEEE Transaction on Speech and Audio Processing*, vol. 9, no. 2, pp. 162–167, Feb. 2001.

[4] M. Otani, Y. Kajikawa, and Y. Nomura, "An acoustic echo cancellation using sub-adaptive filter," *Electronics and Communications in Japan, Part 3*, vol. 90, no. 2, pp. 9–21, Feb. 2007.

[5] A. Mader, H. Puder, and G. U. Shmidt, "Step-size control for acoustic echo cancellation filter – an overview," *Signal Processing*, vol. 80, no. 9, pp. 1697–1719, Sep. 2000.

[6] S. Makino, Y. Kaneda, and N. Koizumi, "Exponentially weighted step-size nlms adaptive filter based on the statistics of a room impulse response," *IEEE Transaction on Speech and Audio Processing*, vol. 1, no. 1, pp. 101–108, Jan. 1993.