# PERFORMANCE IMPROVEMENT OF DOUBLE-TALK DETECTION ALGORITHM IN THE ACOUSTIC ECHO CANCELLER

*Si Ho Kim*, Hong Seok Kwon*, Keun Sung Bae*, Kyung Jin Byun**, Kyung Soo Kim***

*School of Electronic & Electrical Engineering, Kyungpook National University, Taegu, Korea
**Integrated Circuits Research Dept., Electronics and Telecommunications Research Institute, Taejon, Korea
si5@mmirl1.knu.ac.kr

## ABSTRACT

This paper deals with a delay problem in the endpoint detection of double-talk detection algorithm in the acoustic echo canceller. In case that past power is much bigger than current power like at the end of double-talking period, the power estimated using forgetting factor decreases slowly to cause the delay problem in the endpoint detection. In this paper two methods are proposed to solve this problem. One is replacing the current power periodically by a new average power, and the other is removing the past power term in a recursive equation or replacing it by other values. The simulation results show that proposed methods outperform conventional method in the endpoint detection of double-talking periods without increasing the computational burden much more.

## 1. INTRODUCTION

An adaptive echo canceller (AEC)[1] updates the tap coefficients of an adaptive filter to model echo path using an error signal *e(k)* as shown in Fig. 1. If the tap coefficients are updated during the double-talking situation, which means that microphone input signal includes both near-end talker signal and echo signal, they can fluctuate greatly or diverge to misestimate the impulse response of echo path. Hence AEC should stop the filter adaptation during the double-talking period. So far many double-talk detection algorithms have been proposed, and they can be divided into three categories. The first is the method based on the energy comparison of input signals[2] and the second is the method of comparing the linear prediction coefficient of microphone input signal and far-end signal[3]. The third is the method of using the cross-correlation coefficient of far-end signal and error signal[4]. The final method of them is the first approach of using correlation for double-talk detection but it is computational extensive. Hence these days the method of using cross-correlation of microphone input signal and error signal is frequently used because its computational burden is low and detection performance is high[5]. When we use the correlation for double-talk detection, we need to calculate the power which is used to calculate the correlation coefficient.

This work was funded by Integrated Circuits Research Dept., Electronics and Telecommunications Research Institute in Korea.

Usually the power is estimated recursively using a forgetting factor to reduce required memory space and computational load. In that case, past power affects the estimation of current power. Especially, when past power is much bigger than current power like at the end of double-talking period, the power estimated by recursive equation decreases slowly under the influence of past big power and it causes the delay problem in the endpoint detection. In this paper two methods are proposed to solve the delay problem in the endpoint detection. One is replacing the current power periodically by a new average power, and the other is removing the past power term in a recursive equation or replacing it by other values.

This paper is organized as follows. The double-talk detection algorithms and the delay problem in the endpoint detection are explained in the following section, and the methods to solve this problem are proposed in next section. Then we show the experimental results and the performance of proposed methods comparing with that of the conventional method. Finally conclusion is given in the last section.

## 2. DOUBLE-TALK DETECTION ALGORITHM AND DELAY PROBLEM IN THE ENDPOINT DETECTION

Fig. 1 shows the basic structure of an adaptive acoustic echo canceller. The far-end signal *x(k)* enters microphone through echo path and it makes the acoustic echo. The adaptive filter estimates an echo replica $\hat{y}(k)$ and subtracts it from the echo signal *y(k)* so that the echo is cancelled. An NLMS(Normalized Least Mean Square) algorithm is usually used for adaptation because of its simplicity and good performance. The tap

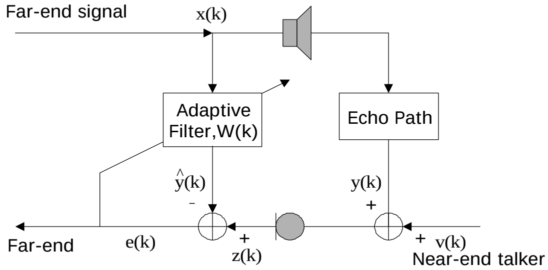


Fig. 1. The structure of adaptive echo canceller

coefficients of an adaptive filter are computed by Eq. (1)[6]. If the microphone input signal *z(k)* consists of an echo signal only, the echo replica $\hat{y}(k)$ converges to *y(k)* as the tap coefficients are updated recursively using Eq. (1). But if microphone input signal includes near-end talker signal as well as an echo signal, the error signal grows so big that the variation of the coefficients increases or diverges. It makes the adaptive filter to misestimate the echo signal. To prevent this, the AEC needs a double-talk detector that detects whether a near-end talker signal exists or not.

As we mentioned in the previous section, the method of using cross-correlation of microphone input signal and error signal is frequently used for double-talk detection these days. It uses the orthogonal principle, that is, in the state of converging, the correlation value between input signal and error signal becomes nearly zero if the error signal consists of only residual echo signal, but close to one if it contains near-end talker signal. Then double-talk detector can determine whether near-end talker signal exists or not by comparing the correlation with an appropriate threshold.

$$W(k+1) = W(k) + \mu(k)e(k)X(k) \tag{1}$$

$$\mu(k) = \frac{\mu_0}{X^T(k)X(k)}$$

The cross-correlation coefficient, used as a detection parameter of near-end talker signal, is calculated by Eq. (2) where $\sigma^2_{de}$ means the correlation between input signal and error signal. $\sigma^2_d$ and $\sigma^2_e$ are the power of microphone input signal and error signal, respectively. The powers are calculated recursively by Eq. (3), (4), and (5) with a forgetting factor $\alpha$.

$$Cor(k) = \frac{\sigma^2_{de}(k)}{\sqrt{\sigma^2_d(k)\ \sigma^2_e(k)}} \tag{2}$$

$$\sigma^2_d(k) = (1-\alpha)\cdot\sigma^2_d(k-1) + \alpha\cdot d(k)^2 \tag{3}$$

$$\sigma^2_e(k) = (1-\alpha)\cdot\sigma^2_e(k-1) + \alpha\cdot e(k)^2 \tag{4}$$

$$\sigma^2_{de}(k) = (1-\alpha)\cdot\sigma^2_{de}(k-1) + \alpha\cdot d(k)e(n) \tag{5}$$

It is useful to estimate the current power using Eq. (3), (4), (5) as long as the power of a signal does not change abruptly. But if we estimate the current power recursively like at the end of double-talking period, where past power is much bigger than current power, the estimated current power decreases slowly and it causes the power misestimation. This is due to the fact that the past power becomes the dominant component in the current power estimation if past power is very large compared with current power. Therefore the correlation coefficient calculated by these powers decreases slowly and it causes the delay problem in the endpoint detection of double-talking period. The delay problem is not important if echo path does not change during double-talking period or after. But it could be important to quickly detect the endpoint in case that echo path changes frequently like in the mobile communication.

## 3. PROPOSED DOUBLE-TALK DETECTION ALGORITHMS

As we explained in the previous section, the power misestimation problem in the endpoint of double-talking period is due to the fact that the estimated current power reflects continuously the old past power. Therefore, to solve this problem, the old past power must be removed and the latest power must be reflected greatly to estimate the current power. For this, two methods are proposed in this paper. One is that the current power is periodically replaced by a new average power and the other is that the past power in a recursive equation is periodically removed or replaced by other values.

### 3.1. Method of using a new average power

We can remove periodically the past big power by replacing the current power with a new average power at regular intervals during double-talking period. The flowchart of this algorithm is shown in Fig. 2.

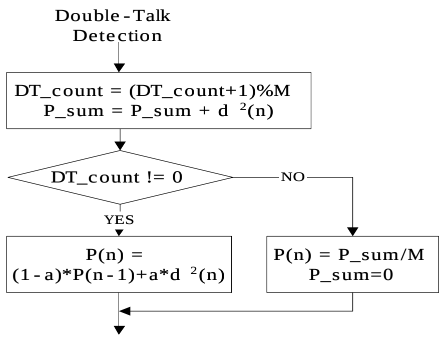


Fig. 2. Flowchart of the proposed method using a new average power

During double-talking period, the power of an input signal *d(n)* is calculated and accumulated to P_sum at every samples. At regular interval M, P_sum is divided by M and the current power is replaced by the new average power P_sum. In the rest of double-talk period, it is calculated by a recursive equation.

### 3.2. Methods of removing or replacing the past power

Recursive equations given in Eq. (3), (4), and (5) can be rewritten in the form of Eq. (6), where $\sigma^2_{k_0}$ denotes the estimated power at past sample $k_0$ and $\Delta\sigma^2_{k_0k}$ is the average power between $k_0+1$ and $k_0$.

$$\begin{aligned}\sigma^2_k &= (1-\alpha)\ \sigma^2_{k-1} + \alpha x^2_k \\ &= (1-\alpha)^{k-k_0}\ \sigma^2_{k_0} + \Delta\sigma^2_{k_0k}\end{aligned} \tag{6}$$

As the $k-k_0$ in Eq. (6) increases, the weight value $(1-\alpha)^{k-k_0}$ approaches zero so that $\sigma^2_{k_0}$ affects little to estimate $\sigma^2_k$. But if $\sigma^2_{k_0}$ is so big that $(1-\alpha)^{k-k_0}\ \sigma^2_{k_0}$ is bigger than $\Delta\,\sigma^2_{k_0k}$, the estimated current power $\sigma^2_k$ depends on $\sigma^2_{k_0}$ mostly. Therefore the estimated current power cannot decrease rapidly and it causes the delay problem in endpoint detection. We solved this problem by removing the big value $\sigma^2_{k_0}$ or replacing it by other values. In this paper, three values for replacing $\sigma^2_{k_0}$ are proposed and investigated. If we define $\sigma'^2_{k_0}$ as a new value for replacing $\sigma^2_{k_0}$, and $\sigma'^2_k$ as the newly calculated value for $\sigma^2_k$, then Eq. (6) can be represented as Eq. (7).

$$\sigma'^2_k = (1-\alpha)^{k-k_0}\ \sigma'^2_{k_0} + \Delta\,\sigma^2_{k_0k} \tag{7}$$

(1) Case Ⅰ: $\sigma'^2_{k_0} = 0$

In this case the current power is estimated only by the samples between $k_0$ and $k$ neglecting the power for the samples before $k_0$. It is possible to estimate the current power using only the samples between $k_0$ and $k$ if $k-k_0$ is long enough. From Eq. (6), $\Delta\,\sigma^2_{k_0k}$ can be represented as

$$\Delta\,\sigma^2_{k_0k} = \sigma^2_k - (1-\alpha)^{k-k_0}\ \sigma^2_{k_0} \tag{8}$$

Substituting $\sigma'^2_{k_0} = 0$ and Eq. (8) into Eq. (7), we obtain $\sigma'^2_k$ as given in Eq. (9).

$$\sigma'^2_k = \sigma^2_k - (1-\alpha)^{k-k_0}\ \sigma^2_{k_0} \tag{9}$$

(2) Case Ⅱ: $\sigma'^2_{k_0} = \Delta\,\sigma^2_{k_0k}$

In the previous case we could reduce the influence of the past big power by using only $\Delta\,\sigma^2_{k_0k}$. But in that case the estimated current power depends on only $\Delta\,\sigma^2_{k_0k}$ so the estimated power may change abruptly. Thus it needs to estimate smoothly and naturally the current power as adding an appropriate past power. First we uses $\Delta\,\sigma^2_{k_0k}$ as an appropriate value for $\sigma'^2_{k_0}$. Substituting $\sigma'^2_{k_0} = \Delta\,\sigma^2_{k_0k}$ and Eq. (8) into Eq. (7), we obtain $\sigma'^2_k$ as follows.

$$\sigma'^2_k = (\sigma^2_k - (1-\alpha)^{k-k_0}\ \sigma^2_{k_0})(1+(1-\alpha)^{k-k_0}) \tag{10}$$

(3) Case Ⅲ: $\sigma'^2_{k_0} = \sigma'^2_k$

Similarly we use $\sigma'^2_k$ as an appropriate value for $\sigma'^2_{k_0}$ to smooth the estimated power with reducing the influence of the past big power. Substituting $\sigma'^2_{k_0} = \sigma'^2_k$ and Eq. (8) into Eq. (7), we obtain $\sigma'^2_k$ as given in Eq. (11).

$$\sigma'^2_k = \frac{\sigma^2_k - (1-\alpha)^{k-k_0}\ \sigma^2_{k_0}}{1 - (1-\alpha)^{k-k_0}} \tag{11}$$

The previous three cases show that the Eq. (9), (10), and (11) can remove the power for the samples before $k_0$. Hence we can reduce periodically the influence of the past big power by using one of Eq. (9), (10), and (11) at regular intervals $k-k_0$, where $k-k_0$ is long enough for $\Delta\,\sigma^2_{k_0k}$ to reflect the current power well. The flowchart of this algorithm is shown in Fig. 3 and M denotes the value of $k-k_0$.

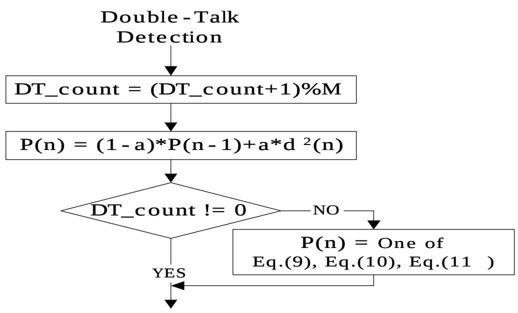


Fig. 3. Flowchart of the proposed method with removing or replacing the past power

## 4. EXPERIMENTAL RESULTS

In this section, we evaluated the performance of the proposed endpoint detection algorithms compared with conventional methods. We used the NLMS as an adaptation algorithm and the number of the tap was 256. Adaptation constant $\mu_0$ is 0.3 and forgetting factor $\alpha$ is 0.0039 which is the reciprocal of 256.

We set the threshold of correlation coefficient for double-talk detection to be 0.7. Fig. 4 shows the microphone input signal which contains echo signal with near-end talker signal. The near-end talker signal exists between sample 7000 and 11500.

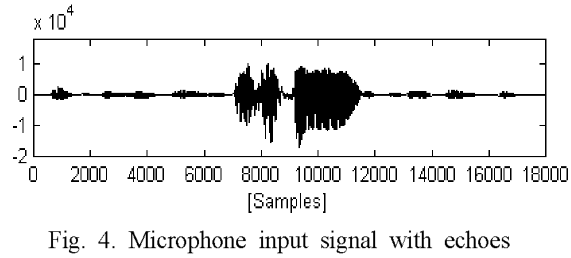


Fig. 4. Microphone input signal with echoes

We define the methods used in experiment as follows:

A : Conventional method using recursive equation only.

B : The proposed method of using a new average power.

C : The case of $\sigma'^2_{k_0} = 0$ in the secondly proposed method.

D : The case of $\sigma'^2_{k_0} = \Delta\sigma^2_{k_0 k}$ in the secondly proposed method.

E : The case of $\sigma'^2_{k_0} = \sigma'^2_k$ in the secondly proposed method.

In Fig. 5, the correlation value starts to decrease slowly at the sample 11500, which is the real endpoint of double-talking period, and meets the threshold at about 13000. Hence the double-talk detector decides 13000 as the endpoint but it is delayed about 1500. Fig. 6, however, shows that the correlation values decreases abruptly in the neighborhood of 11500. The

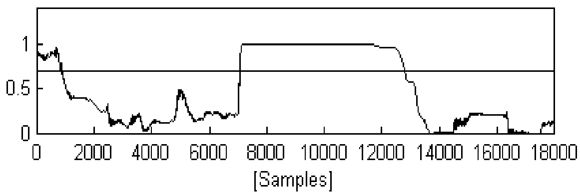


Fig. 5. Trajectory of cross-correlation coefficient for method A

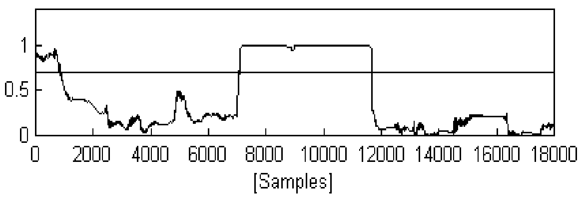


Fig. 6. Trajectories of cross-correlation coefficient for methods B, C, D, E

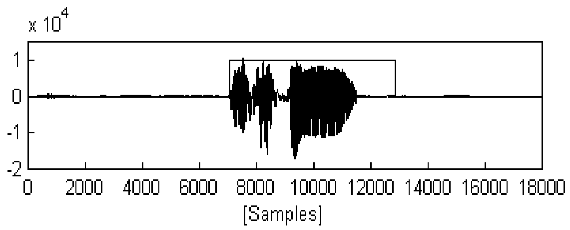


Fig. 7. Result of the double-talk detection for method A

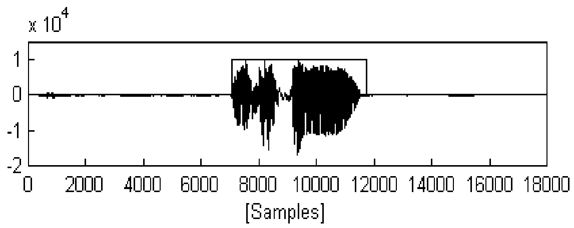


Fig. 8. Results of the double-talk detection for methods B, C, D, E

detection results displayed in Fig. 7 and 8 also show that the detection point by the proposed methods outperforms the conventional method. The proposed methods do not increase the computational burden much more because it is possible to implement it as adding the simple operation in conventional method.

## 5. CONCLUSION

The power estimated recursively decreases slowly at the end of double-talking period, where the past power is much bigger than current power, and it causes the delay problem in the endpoint detection. To solve this problem, we proposed two methods in this paper. One is that the current power is periodically replaced by a new average power and the other is that the past power in recursive equation is periodically removed or replaced by other values. The computer simulation verified that proposed methods outperform conventional method in the endpoint detection of double-talking periods without increasing the computational burden much more.

## REFERENCE


[1] Peter Heitkamper, "An Adaptation Control for Acoustic Echo Cancellers," IEEE Signal Processing Letters. vol.4, no.6, pp.170-172, Jun. 1997.

[2] D. G. Messerschmitt, "Echo Cancellation in Speech and Data Transmission," IEEE J. Select. Areas in Comm., vol. SAC-2, no.2, pp.283-297, Mar. 1984.

[3] Jae Ha Yoo and Sung Ho Cho, "A New Double Talk Detector Using The Lattice Predictors For An Acoustic Echo Canceller," Proceeding of IEEE TENCON'97 Speech and Image Technologies for Computing and Telecommunication vol.2, pp.483-486, 1997.

[4] Hua Ye, Bo-Xiu Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem," IEEE Trans. on Communications, vol.39, no.11, pp.1542-1545, Nov. 1991.

[5] Haeng Woo Lee, Myoung E Eun, Chong Kyo Kim, "A Double-Talk Processing Algorithm of Acoustic Echo Canceller," J. of Acoustical Society of Korea, vol.17, no.3, pp.10-15, 1998.

[6] S. Haykin, *Adaptive Filter Theory*, 3rd Edition, Prentice-Hall, 1996.