A New Type Echo Canceling by Using the Smart Acoustic Room (SAR) system & Correlation Function For the Double-Talk Condition

Rui Chen*, Prof. Mohammad Reza Alsharif*, Prof. KatsumiYamashita +

*Department of Information Engineering, University of the Ryukyus
1-Senbaru, Nishihara, Okinawa, Japan
asharif@ie.u-ryukyu.ac.jp

+Department of Electrical and Electronic Systems, Graduate School of Engineering,
Osaka Prefecture University, Osaka, Japan
yamashita@comm.ees.osakafu-u.ac.jp

Abstract: To solve the double-talk problem in the echo canceling, various algorithms such as CLMS, ECLMS FECLMS have been proposed by authors to challenge the double-talk in the echo canceling system. In this paper we propose a new type echo canceling by using the smart acoustic room (SAR) system & correlation function to challenge the double-talk problem. The proposed algorithm is combination of the ANC and the AEC theories. By using this algorithm, the signal is cancelled in the microphone position locally. The echo signal will not be generated in the telephone system. The computer simulation results support the theoretical findings and verify the robustness of the proposed algorithm in the double-talk situation.

Key words: ANC, AEC, Echo canceling, Smart acoustic room (SAR), Double-talk, Correlation function..

1-Intoduction

The conventional algorithm usually stops adaptation whenever double-talk sensor detects this condition. Stopping the tap adaptation is just a passive action to handle the double-talk condition and it causes lowering speed of adaptations and/or totally mislead when the echo path changed in the period of halting tap adaptation. Other works for challenging the problem of double-talk situation in the echo canceling can be found in [1], [2], and [3] that cause much more complexity adding to a simple LMS algorithm.

To solve the double-talk problem and reduce computational loads, FBAF algorithm, correlation LMS (CLMS) algorithm, Extended CLMS (ECLMS) algorithm and Frequency domain ECLMS (FECLMS) have been proposed [4], [5], [6] and [7]. We can continue the tap adaptation (non-freezing) even in the double-talk situation, without misleading the estimation process. However, the convergence of these algorithms was not perfectly satisfactory.

In this paper we propose a new type echo canceling by using the smart acoustic room (SAR) [8] & correlation function to challenge the double-talk problem. The algorithm uses two speakers and one microphone, by smartly control the acoustic impulse response the speaker signal will be cancelled at the microphone position locally. That is, the microphone cannot receive any echo signal. For the double-talk, the correlation function in the frequency domain also is used.

2-Double-talk condition

In the echo canceling system shown in Fig.1, the acoustic impulse response of the teleconference room is estimated by an adaptive algorithm such as LMS algorithm. The output of the FIR filter, $\tilde{y}(n)$, is presented by:

$$\tilde{y}(n) = \sum_{i=0}^{N-1} h_i x(n-i) \tag{1}$$

where N is the number of tap, h is the tap coefficient of the adaptive FIR filter and $x(n)$ is the far-end signal at sample n.

The echo signal is obtained from echo impulse response, r, as follows (N is the acoustic impulse response length):

$$y(n) = \sum_{i=0}^{N-1} r_i x(n-i) \tag{2}$$

The error signal, $e(n)$, is calculated as below:

$$e(n)=d(n)-\tilde{y}(n) \tag{3}$$

where $d(n)$ is microphone signal that usually contains the echo signal. The LMS algorithm is as follows:

$$h_i(n+1) = h_i(n) + 2\mu_0 . e(n)x(n-i) \tag{4}$$

where μ_0 is the step size for tap coefficients adaptation. If the near-end signal $s(n)$ is also presented during the echo canceling, then the microphone signal contains both the echo and the near-end signals.:

$$d(n)=y(n)+s(n) \quad (5)$$

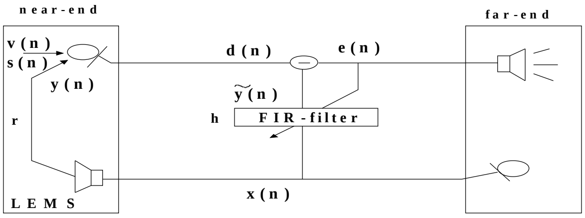

Fig.1. Echo canceller system

We call this condition as double-talk condition. As we know the LMS algorithm it has difficulty to work in the double-talk condition. To challenge the double-talk condition, the correlation function is used. Because the near-end signal *s(n)* is uncorrelated with the input signal, the correlation of the *s(n)* and *x(n)* nearly equal zero. So by using the correlation function algorithm, the gradient for tap adaptation does not carry the undesired near-end signal to misadjust the adaptive digital filter for echo path identification [5], [6], [7]. However, the steady state convergences of these algorithms are not satisfactory. Here we combine ANC with AEC to improve echo canceling performance. In ANC, as we know the acoustic noise is supposed to be cancelled by generating an opposite phase signal that is generated by adaptive filtering of reference (main) noise signal. Now, if we use this ANC structure at near end room in AEC system, then echo signal will be diminished at the microphone position. That is, a very week feed back exists between loudspeaker and microphone. In a sense, we cancel echo signal before it enters to microphone, acoustically by using ANC system. In the next section, we explain about not only a simple ANC, rather for well estimation of secondary path we have defined a SAR (Smart Acoustic Room) algorithm to challenge this problem, before using ANC to AEC that will be explained later in the following sections.

3- SAR Algorithm

The secondary path estimation is a big problem in ANC. Usually we estimate the secondary path by off-line, which is not an effective method. To challenge this problem the virtual microphone algorithm is proposed [9]. By using virtual microphone algorithm, we can estimate the main path and the secondary path simultaneous. In Fig.2, a SAR model by using the virtual microphone is shown. The source signal *x(n)* is for instance a record player output or any audio electric signal. This signal usually converted to acoustic signal through an amplifier and a loudspeaker in order to propagate in a room for listening. The acoustic paths from Speaker S_1 to the Microphone M is $w_1(n)$ and the one from Speaker S_2 is $w_2(n)$. We want to make a null point at the place of Microphone M. For this purpose, we put one adaptive filter estimator *h(n)* in order to predict the acoustic paths and to zero-enforce the signal of M. The signal of Microphone is called the error signal, *e(n)*, and it is obtained as follows:

$$e(n)=x(n)*w_1(n)+x(n)*h(n)*w_2(n) \quad (6)$$

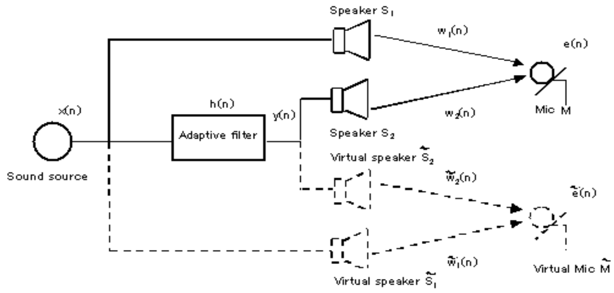

Fig.2. SAR model by using the virtual microphone

Aside of speakers S1 and S2, we imagine that we have two virtual speakers S̃1 and S̃2 in parallel with S1 and S2, respectively. Also, we define two virtual acoustic paths for S̃1 and S̃2 as w̃1(n) and w̃2(n) from each virtual speaker to a virtual microphone M̃ (see Fig.2). The signal of the virtual microphone is $\tilde{e}(n)$. According to Fig.2, we can write the following relation for the virtual paths:

$$\tilde{e}(n)=x(n)*\tilde{w}_1(n)+x(n)*h(n)*\tilde{w}_2(n) \quad (7)$$

If *h(n)* is adapted perfectly, then the virtual error signal will be diminished to zero. Therefore, in Z transform we have:

$$X(z)*\tilde{W}_1(z)+X(z)*H(z)*\tilde{W}_2(z)=0 \quad (8)$$

That is:

$$H(z)=-\frac{\tilde{W}_1(z)}{\tilde{W}_2(z)} \quad (9)$$

From Eq. (6) and (8), we conclude that:

$$\frac{W_1(z)}{W_2(z)}=\frac{\tilde{W}_1(z)}{\tilde{W}_2(z)} \Rightarrow \frac{W_1(z)}{\tilde{W}_1(z)}=\frac{W_2(z)}{\tilde{W}_2(z)}=a(z) \quad (10)$$

Function *a(z)* describes the relation between the real and virtual part of the system. Then we can use two simple LMS adaptive filters to estimate the impulse responses w_1 and w_2. For estimation the w_1, the error signal can be written:

$$E_{w1}(z)=[W_1(z)-a(z)\tilde{W}_1(z)]X(z)$$
$$=W_1(z)X(z)-\frac{W_2(z)}{\tilde{W}_2(z)}\tilde{W}_1(z)X(z)=E(z) \quad (11)$$

As the same for estimation the w_2., the error signal can be written:

$$E_{w1}(z) = -E(z) \tag{12}$$

That is, the acoustic paths $w_1(n)$ and $w_2(n)$ can be estimated by using the real error $e(n)$. In order to reduce the computational complexity at this time all the computation will be done in the frequency domain [7]. First the FFT of the input signal and echo signal are calculated.

$$F_x(n, p) = \sum_{k=0}^{N-1} x(n-k)W^{kp} \tag{13}$$

$$F_y(n, p) = \sum_{k=0}^{N-1} y(n-k)W^{kp} \tag{14}$$

Then the FFT transform of the error signal is calculated.

$$F_e(n, p) = \sum_{k=0}^{N-1} e(n-k)W^{kp} \tag{15}$$

So as the same as the FECLMS algorithm [6], the acoustic impulse response can be estimated by

$$\tilde{W}_{1_i}(n+1, p) = \tilde{W}_{1_i}(n, p) + \frac{2\mu F_e(n,p)F_x^*(n,p)}{1+tr[F_x(n,p)F_x(n,p)]} \tag{16}$$

$$\tilde{W}_{2_i}(n+1, p) = \tilde{W}_{2_i}(n, p) - \frac{2\mu F_e(n,p)F_y^*(n,p)}{1+tr[F_y(n,p)F_y(n,p)]} \tag{17}$$

The superscript * shows the Hermitian transposition and $tr[\cdot]$ means the trace operator. Finally the $h(n)$ can be calculated by using the inverse FFT transform.

4-Proposed algorithm (Combination of ANC with AEC)

In the Fig.3, the proposed echo canceling system by using the smart acoustic room & correlation function is shown. Also the proposed algorithm will be implement in the frequency domain.

For the double-talk condition the signal from the microphone will be defined as follows:

$$\begin{aligned} d(n) &= e(n) + s(n) \\ &= x(n) * w_1(n) + x(n) * h(n) * w_2(n) + s(n) \end{aligned} \tag{18}$$

First the auto-correlation of the input signal is calculated:

$$R_{xx}(n, k) = \sum_{j=0}^{n} x(j)x(j-k) \tag{19}$$

$$R_{yy}(n, k) = \sum_{j=0}^{n} y(j)y(j-k) \tag{20}$$

And then the cross-correlation function is calculated:

$$R_{ex}(n, k) = \sum_{j=0}^{n} d(j)x(j-k) \tag{21}$$

$$R_{ey}(n, k) = \sum_{j=0}^{n} d(j)y(j-k) \tag{22}$$

The fast Fourier transform is shown as below:

$$F_{xx}(n, p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} x(j)x(j-k)\right]W^{kp} \tag{23}$$

$$F_{yy}(n, p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} y(j)y(j-k)\right]W^{kp} \tag{24}$$

$$F_{ex}(n, p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} d(j)x(j-k)\right]W^{kp} \tag{25}$$

$$F_{ey}(n, p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} d(j)y(j-k)\right]W^{kp} \tag{26}$$

So the acoustic paths can be updated by:

$$\tilde{W}_{1_i}(n+1, p) = \tilde{W}_{1_i}(n) + \frac{2\mu F_{ex}(n,p)F_{xx}^*(n,p)}{1+tr[F_{xx}(n,p)F_{xx}(n,p)]} \tag{27}$$

$$\tilde{W}_{2_i}(n+1, p) = \tilde{W}_{2_i}(n, p) - \frac{2\mu F_{ey}(n,p)F_{yy}^*(n,p)}{1+tr[F_{yy}(n,p)F_{yy}(n,p)]} \tag{28}$$

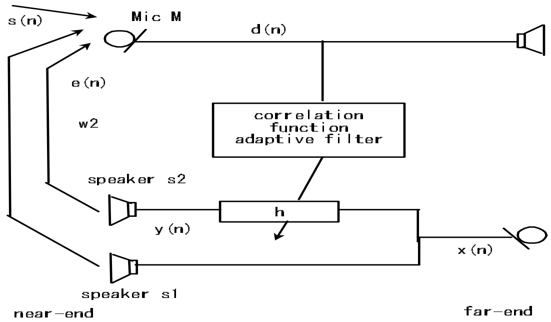

Fig.3 the proposed echo canceling system

5-Simulation results

To demonstrate the validity and the robustness of the proposed algorithm, we performed some computer simulations. The input signal $x(n)$ and the double-talk signal $s(n)$ are the speeches of woman in English as shown in Fig4, 5, respectively. The adaptive filter has 32 taps. The step size is 0.007. The acoustic echo impulse response, $w_{1,2}$, of the room are assumed to have exponential decaying shape that decreases to -60 dB after N samples as follows:

$$w_{1,2} = Randn[\exp(-8i/N)] \tag{29}$$

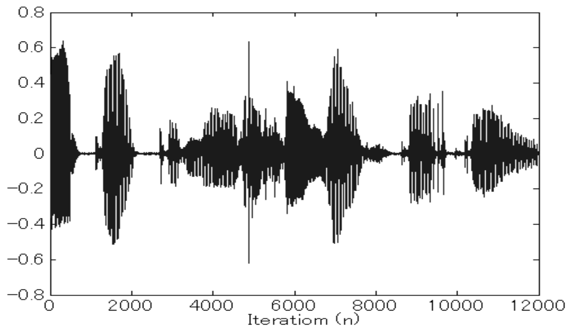

Fig 4. The input speech signal x(n)

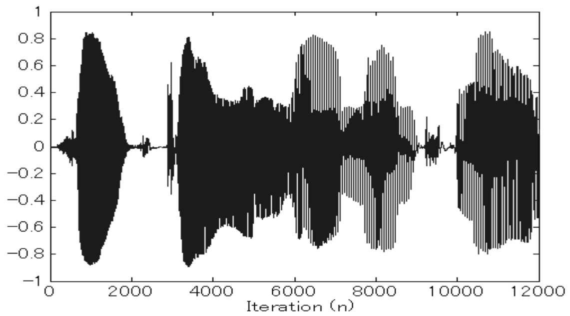

Fig 5. The double-talk speech signal s(n)

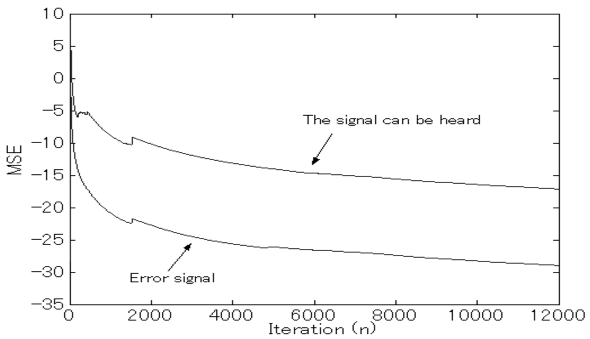

Fig.4 The MSE of the proposed algorithm in single-talk condition.

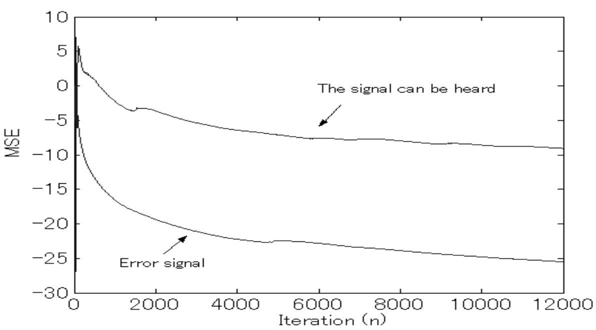

Fig.5 The MSE of the proposed algorithm in double-talk condition.

The simulation results show that how much dB echo is cancelled at the microphone location and at the same time how much db signal the person who is talking at the near-end room can be heard. In the Fig.4, the MSE of the proposed algorithm in the signal talk condition is shown. In the Fig.5, the MSE of the proposed algorithm in the double- talk condition is shown.

In the single-talk condition, the proposed algorithm converges to –29 dB of echo cancellation at the microphone and –17db signal can be heard. In the double-talk condition, the proposed algorithm converges to –26 dB and –8db signal can be heard. That is, the echo can be cancelled in microphone position by using the smart acoustic control & correlation function. And also the person who is talking in the near-end room can hear the signal from the speakers clearly.

6- conclusion

In this paper, a new type adaptive echo canceling by using the smart acoustic room system & correlation function is presented. The proposed algorithm is combined with the ANC and the AEC theories. By smartly control the acoustic impulse response the speaker signal will be cancelled at the microphone position locally. The simulation results show that the proposed algorithm has a good performance for the double-talk condition.

Reference

[1] P. Heitkamper, "An adaptation control for acoustic echo cancellers," *IEEE Signal Processing Letters*, Vol. 4, No. 6 pp.170-172, June 1997.

[2] T.Gansler, "A double-talk resistant sub-band echo canceller, "*Elsevier, Signal Processing* 65, pp.89-101, April 1997.

[3] T. Gansler, S. L Gay, M.M Sondhi, J. Benesty, "Double-talk robust fast converging algorithm for network echo cancellation," IEEE Tans, on Speech and Audio Processing, Vol.8, No.6 pp.656-663, November 2000.

[4] M.R. Asharif, F.Amano, "Acoustic echo-canceler using the FBAF algorithm," IEEE Trans.Communications, vol42, No.12, pp3090-3094, Dec.1994.

[5] M. R. Asharif, T. Hayashi, K. Yamashita, "Correlation LMS algorithm and its application to double-talk echo cancelling,'' *ELECTRONICS LETTERS*, vol.35, No.3, pp.194-195, 4th February 1999.

[6] M. R. Asharif, A. Shimabukuro, T. Hayashi, K. Yamashita, "Expanded CLMS algorithm for double-talk echo cancelling,'' *Proc IEEE, SMC'99*, Japan, Vol.1, pp. 998-1002, Oct. 1999.

[7] M. R. Asharif, M. Shiroma, K. Yamashita, "Design of echo canceler with frequency domain ECLMS algorithm," GlobeCom 2000, San Francisco, USA, November, 2000.

[8] M. R. Asharif, Ryota Higa, Rui Chen, "Smart acoustic room", The 2004 autumn meeting of the acoustical society of Japan, pp601-602, September 2004.

[9] Y. Ohana, T. Kohna, "Direct Fully Adaptive Active Noise Control Algorithms Without Identification of Secondary Path Dynamics", IEEE International Conference on Control Application, Scotland, 2002.