World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:5, No:3, 2011

An Improved Transfer Logic of the Two-Path Algorithm for Acoustic Echo Cancellation

Chang Liu and Zishu He

Open Science Index, Electronics and Communication Engineering Vol:5, No:3, 2011 publications.waset.org/11883/pdf

***Abstract*—**Adaptive echo cancellers with two-path algorithm are applied to avoid the false adaptation during the double-talk situation. In the two-path algorithm, several transfer logic solutions have been proposed to control the filter update. This paper presents an improved transfer logic solution. It improves the convergence speed of the two-path algorithm, and allows the reduction of the memory elements and computational complexity. Results of simulations show the improved performance of the proposed solution.

***Keywords*—**Acoustic echo cancellation, Echo return loss enhancement (ERLE), Two-path algorithm, Transfer logic

I. Introduction

ACOUSTIC echo cancellers (AECs) with adaptive filters are widely used to remove the acoustic echo resulting from the acoustic coupling between the loudspeaker and the microphone. However, a serious problem is the double-talk, i.e., the near-end and far-end talkers are active simultaneously. During the double-talk, the near-end speech that behaves as the uncorrelated noise may cause the adaptive filter to diverge. For this reason, several double-talk detectors (DTDs) have been proposed for halting the adaptation [1-4]. It is difficult to distinguish between the double-talk and echo path change for these DTDs. To overcome this problem, two-path adaptive echo canceller has been employed [5-11]. In this scheme, the echo canceller comprises a background (adaptive) filter and a foreground (non-adaptive) filter. The background filter adapts continuously, while the foreground filter is mostly non-adaptive and kept in fixed state. Both filters operate on the same input signals in order to cancel the same echo signal and the error signal of the foreground filter serves as the output. When the background filter is judged to perform better than the foreground filter, the coefficients of the background filter are copied into the foreground filter.

In the two-path algorithm, the copying of coefficients from the background filter to the foreground filter is controlled by the transfer logic. Only when all the transfer logic conditions are met, the coefficient copying is permitted. The conventional two-path (CTP) solution in [5] is based on the comparison of the output error of the foreground and background filter. This solution does not avoid the erroneous coefficient copying in the double-talk situation. An improvement of the conventional two-path solution (ITP) improves the convergence speed of the two-path algorithm and prevents from the erroneous coefficient copying [6]. However, it needs increased memory elements and computational complexity.

In this paper, we present an improved transfer logic solution for the two-path algorithm. This transfer logic solution relies on the comparison of the echo return loss enhancement (ERLE). It improves the convergence speed of the two-path algorithm, and allows the reduction of the memory elements and computational complexity.

This paper is organized as follows. Section II introduces two-path adaptive model and derives the previous transfer logic solutions. Section III presents the proposed transfer logic and computer simulations are illustrated in section IV. Conclusions are given in section V.

II. Two-Path Echo Canceller

A. Two-Path Structure

A block diagram of the two-path echo canceller is shown in Fig. 1. The signal $x(n)$ (often referred to as the far-end signal) excites the echo path and produces an echo signal $d(n)$. The microphone signal $z(n)$ includes not only the echo, but also the background noise $v(n)$ and possible near-end speech $s(n)$, i.e., $z(n) = d(n) + v(n) + s(n)$. The foreground filter which is modeled as a finite impulse response (FIR) filter, produces the acoustic estimated echo $\hat{d}_f(n)$. By subtracting the estimated echo from the microphone signal, the echo-cancelled error signal is obtained by

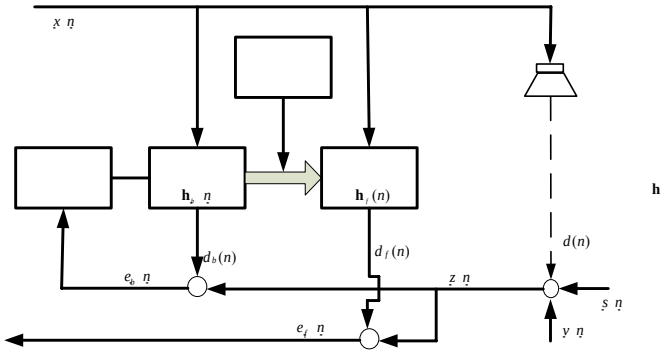


Fig.1 Two-path structure

$$e_f(n) = z(n) - \hat{d}_f(n) = z(n) - \hat{\mathbf{h}}_f(n)^T \mathbf{x}(n), \quad (1)$$

where $\hat{\mathbf{h}}_f(n) = [\hat{h}_{f,0}(n), \cdots, \hat{h}_{f,L-1}(n)]^T$ is the foreground filter and $\mathbf{x}(n) = [x(n), \cdots, x(n-L+1)]^T$ is the input vector, L is the filter length, $[\square]^T$ denotes the matrix transpose.

Chang Liu is with the University of Electronic Science and Technology, Chengdu, Sichuan, China (phone: +86-028-61852219; e-mail: chaneaaa@163.com).

Open Science Index, Electronics and Communication Engineering Vol:5, No:3, 2011 publications.waset.org/11883/pdf

For the background filter, the error signal is produced like (1)

$$e_b(n) = z(n) - \hat{d}_b(n) = z(n) - \hat{\mathbf{h}}_b(n)^T \mathbf{x}(n), \quad (2)$$

where $\hat{\mathbf{h}}_b(n) = [\hat{h}_{b,0}(n), \cdots, \hat{h}_{b,L-1}(n)]^T$ is the background filter. The background filter is continuously updated using the NLMS algorithm

$$\mathbf{w}_b(n+1) = \mathbf{w}_b(n) + \frac{\mu e_b^*(n)\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2 + \varepsilon}, \quad (3)$$

where $\|\cdot\|$ is the Euclidian norm of a vector, μ is a step-size and ε is a regularization parameter.

B. Conventional Two-path (CTP) Solution

In the CTP solution [5, 6], the copy of the filter coefficients is controlled by comparing the short term powers of the signals $e_f(n)$, $e_b(n)$ and $x(n)$. The update conditions are expressed by

$$\frac{P_{e_b}(n)}{P_{e_f}(n)} < \alpha_{e_b,e_f}, \quad (4)$$

$$\frac{P_{e_b}(n)}{P_x(n)} < \alpha_{e_b,x}, \quad (5)$$

where α_{e_b,e_f} and $\alpha_{e_b,x}$ are thresholds and the short term power can be estimated by

$$P_x(n) = \frac{1}{K}\sum_{k=0}^{K-1} x^2(n-k), \quad (6)$$

Thus, the foreground filter update performs in the following manner

$$\hat{\mathbf{h}}_f(n+1) = \begin{cases} \hat{\mathbf{h}}_b(n), & \text{if (4) and (5) are true} \\ \hat{\mathbf{h}}_f(n), & \text{otherwise} \end{cases}. \quad (7)$$

In (7), the copy of the filter coefficients is permitted from the background filter to the foreground filter if the conditions (4) and (5) are true. Condition (4) can be figured as the core condition and verifies whether the background filter performs better than the foreground filter, while condition (5) is used to judge the presence of double-talk which prevents deadlock in the echo path change situation.

A major drawback of the CTP is the tradeoff between the convergence speed and stability. During the double-talk, due to minor cancellation of the near-end speech, there is a risk of erroneous coefficient copying [6]. To avoid the risk of erroneous copying, the threshold α_{e_b,e_f} must be set to a suitable low level. However, setting α_{e_b,e_f} low reduces the convergence speed. Furthermore, too low $\alpha_{e_b,x}$ may halt the coefficient copying. The suitable threshold choice is setting the $\alpha_{e_b,x}$ as low as the application allows, and then setting α_{e_b,e_f} as high as possible for ensuring the convergence speed. Unfortunately, this choice of the thresholds which satisfies the robustness may sacrifice the convergence speed [6].

C. Improved Two-path (ITP) Solution

The ITP solution is proposed in [6] to increase the convergence speed of the two-path algorithm. It depends on a complementary update condition by use of estimating the normalized square deviation (NSD). The NSD of the background filter is calculated as

$$\xi_b(n) = \frac{\sum_{l=0}^{L-1}(h_l - \hat{h}_{b,l}(n))^2}{\|\mathbf{h}\|^2}, \quad (8)$$

where $\mathbf{h} = [h_0, \cdots, h_{b,L-1}]^T$ is the impulse response vector of the loudspeaker-enclosure-microphone (LEM).

The NSD of the foreground filter $\xi_f(n)$ can still be calculated in the manner of (8). Thus, the core condition which determines whether the foreground filter should be updated is given by

$$\xi_b(n) < \xi_f(n). \quad (9)$$

Due to the unknown LEM impulse response, $\mathbf{h}$, the NSD in (8) is not accessible. In order to obtain the estimated square deviation, an artificial delay of M samples is inserted into the signal path of $z(n)$ before the sub-traction yielding $e_b(n)$, as shown in Fig. 2. The output signal is artificially delayed by M samples. Hence, the background filter is increased by M taps, resulting in an extended background filter $\mathbf{h}_e(n)$, that is

$$\mathbf{h}_e(n) = \begin{bmatrix} \hat{\mathbf{h}}_I(n) \\ \hat{\mathbf{h}}_b(n) \end{bmatrix}, \quad (10)$$

where $\hat{\mathbf{h}}_I(n) = [\hat{h}_{I,0}(n), \cdots, \hat{h}_{I,M-1}(n)]^T$.

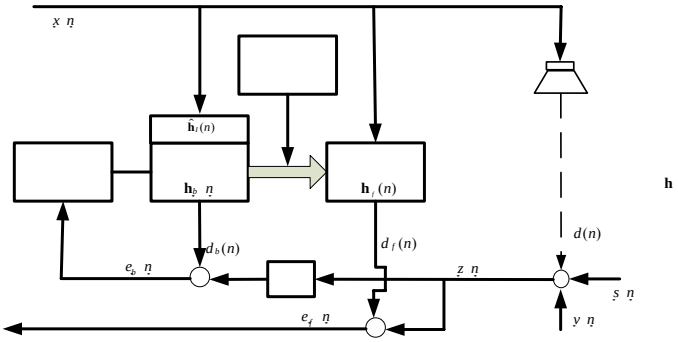


Fig.2 Two-path structure for ITP solution

Thus, the complementary update condition can be given by

$$\frac{\hat{\xi}_b(n)}{\hat{\xi}_f(n)} < \alpha_{b,f}, \quad (11)$$

where $\alpha_{b,f}$ is a threshold.

The foreground deviation estimate $\hat{\xi}_f(n)$ is updated by the previous deviation estimate of the background filter, this can be expressed by

$$\hat{\xi}_f(n+1) = \begin{cases} \hat{\xi}_b(n), & \text{if updating the foregroud filter is true} \\ \hat{\xi}_f(n), & \text{otherwise} \end{cases}. \quad (12)$$

The foreground filter update is then given by

$$\hat{\mathbf{h}}_f(n+1) = \begin{cases} \hat{\mathbf{h}}_b(n), & \text{if [(4) or (11)] and (5) are true} \\ \hat{\mathbf{h}}_f(n), & \text{otherwise} \end{cases}, \quad (13)$$

Note that condition (11) is only a complementary update condition. It helps increase the convergence speed, but can not

Open Science Index, Electronics and Communication Engineering Vol:5, No:3, 2011 publications.waset.org/11883/pdf

replace the previous condition (4). When the echo path change occurs, $\hat{\xi}_b(n)$ becomes large and the update is stalled. In this situation, the filter update must be triggered by the condition (4).

Due to the increased M taps for estimating the NSD, the ITP needs M more memory elements. The background filter requires about $2M$ extra additions and $2M$ extra multiplications at every adaptation as compared to the CTP solution.

III. Proposed Transfer Logic

In this section, we present the new transfer logic. It obtains a tradeoff between the convergence speed and computational complexity. The idea of the proposed one is based on the comparison between the ERLE of the background filter, $C_b(n)$, and reference ERLE, C_r. The reference ERLE, C_r, represents the best-attained canceling performance of the background and foreground filter. The ERLE of the background filter at time n can be computed by

$$C_b(n) = \frac{\sum_{k=0}^{K-1} |z(n-k)|^2}{\sum_{k=0}^{K-1} |e_b(n-k)|^2}, \tag{14}$$

where K denotes the filter update interval.

By comparing $C_b(n)$ with the reference value C_r, the decision of filter update is made. If $C_b(n)$ is larger than C_r, it indicates that the background filter has a better canceling performance than the present best one, and the filter update should be performed. Hence, the proposed update condition can be expressed by

$$\frac{C_b(n)}{C_r} > \alpha_0, \tag{15}$$

where α_0 is a threshold and should be set to $\alpha_0 \geq 1$.

The proposed condition (15) can replace the condition (11) to form the foreground filter update condition and it is given by

$$\hat{\mathbf{h}}_f(n+1) = \begin{cases} \hat{\mathbf{h}}_b(n), & \text{if [(4) or (15)] and (5) are true} \\ \hat{\mathbf{h}}_f(n), & \text{otherwise} \end{cases}. \tag{16}$$

If the foreground filter is updated, the current reference ERLE should be replaced by a new reference value which is the maximum ERLE of the background and foreground filter. The reference value update is given by

$$C_r = \begin{cases} \max\{C_b(n), C_f(n)\}, & \text{if updating the foregroud filter is true} \\ C_r, & \text{otherwise} \end{cases}, \tag{17}$$

where $C_f(n)$ is the ERLE of the foreground filter and can be computed by

$$C_f(n) = \frac{\sum_{k=0}^{K-1} |z(n-k)|^2}{\sum_{k=0}^{K-1} |e_f(n-k)|^2}. \tag{18}$$

Like the condition (11), (15) is only a complementary update condition. When the echo path change occurs, the condition (4) is used to trigger the filter update. During the double-talk, the near-end speech leads to the decrease of the ERLE of the background filter, even though there is a minor cancellation of the near-end speech. Hence, the foreground filter update can be halted in this situation.

Compared to the ITP solution, since there is unnecessary to extend the background filter by M coefficients, the proposed solution can save $2M$ additions and $2M$ multiplications at every adaptation.

IV. Simulation Results

Several simulations are presented to evaluate the performance of the proposed update condition. As shown in Fig. 3, speech signal is used as the input signal $x(n)$ and its sampling frequency is 8 kHz. For the doubletalk, the near end speech occurs at the *Ith* sample, and the signal $z(n)$ is obtained by

$$z(n) = \begin{cases} d(n) + v(n), & \text{if } n < I \\ d(n) + v(n) + as(n), & \text{otherwise} \end{cases}, \tag{19}$$

where a is the parameter that controls the near-end speech level.

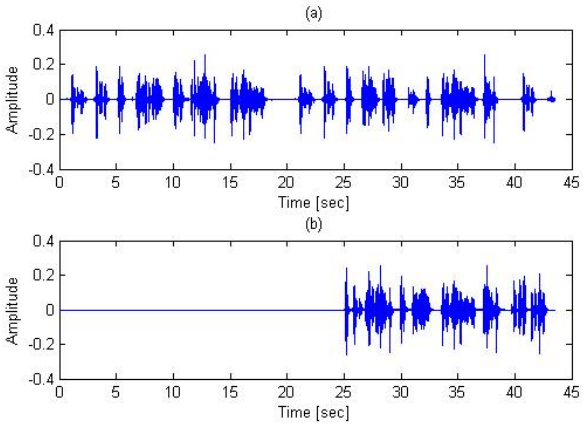


Fig. 3. Speech signals. (a) Far-end signal. (b) Near-end signal. Double talk occurs at about 26s.

In the echo path change situation, the echo path change occurs at the *Jth* sample, the signal $z(n)$ is given by

$$z(n) = \begin{cases} \mathbf{h}_1^T \mathbf{x}(n) + v(n), & \text{if } n < J \\ \mathbf{h}_2^T \mathbf{x}(n) + v(n), & \text{otherwise} \end{cases}, \tag{20}$$

where $\mathbf{h}_1$ and $\mathbf{h}_2$ are two different echo paths used in the simulations.

The unknown echo paths $\mathbf{h}_1$ and $\mathbf{h}_2$ are illustrated in Fig 4(a) and Fig 4(b) respectively. The length of them is 1024. The length of the background and foreground filters is identical with that of $\mathbf{h}_1$, $L = 1024$. The threshold $\alpha_{e_b,x}$ is set to $-18dB$ and α_{e_b,e_f} is set to $-12dB$ which is the highest possible setting for guaranteeing robustness during the double-talk. The thresholds $\alpha_{b,f}$ and α_0 are set to 0dB. In addition, the other parameters I, J, μ, κ, ε and δ can be configured depending on the application in practice. All parameter settings used in the simulations are shown in Table I.

Open Science Index, Electronics and Communication Engineering Vol:5, No:3, 2011 publications.waset.org/11883/pdf

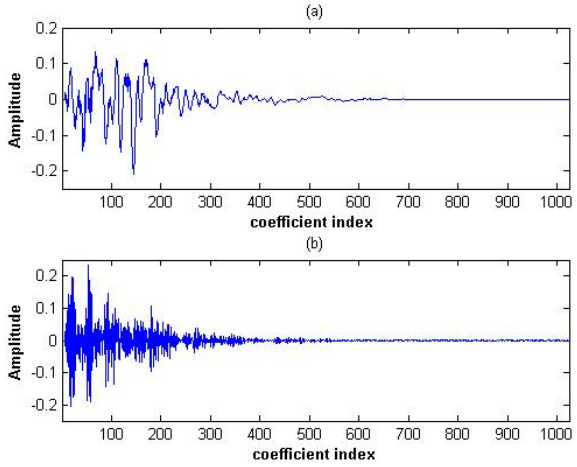


Fig. 4 Echo paths (a) $\mathbf{h}_1$. (b) $\mathbf{h}_2$

TABLE I
PARAMETER SETTINGS

Parameters	Value
L	1024
K	2000
M	20
μ	0.4
ε	0.05
κ	0
δ	0.01
α_{e_b,e_f}	-12dB
$\alpha_{e_b,x}$	-18dB
$\alpha_{b,f}$	0dB
α_0	0dB
I	208000
J	160000
a	0.25

Fig. 5 shows the tradeoff between the convergence speed of the foreground filter and its robustness to the doubletalk for the CTP solution. Two different α_{e_b,e_f} threshold settings are used. We can see in Fig. 5(a) that too high threshold- $\alpha_{e_b,e_f} = -6\text{dB}$ has no ability to detect the doubletalk but achieves the faster convergence speed, while in Fig. 5(b) with $\alpha_{e_b,e_f} = -12\text{dB}$, the slower foreground filter convergence speed and a larger steady-state NSD can be observed. However, this avoids the failure of detecting doubletalk. Furthermore, in the upper plot the foreground filter NSD can reach -25dB, and only -20dB NSD is obtained in the lower plot.

Fig. 6 shows the doubletalk situation for the ITP solution (a) and proposed solution (b). The threshold $\alpha_{b,f}$ and α_0 are set to 0dB. As can be seen in Fig. 6, both the ITP and proposed solution are robust during doubletalk which starts after 26 s. The foreground filter NSD of two solutions can reach -25dB.

Fig. 7 illustrates the ERLR performance of the proposed solution in the doubletalk situation. It is clear to see that after 26s, the ERLR of the background and foreground filter decreases because of the presence of the near-end speech. Therefore, as shown in Fig. 6(b), the proposed solution is able to detect this ERLE decreasing and halt the filter update in the doubletalk.

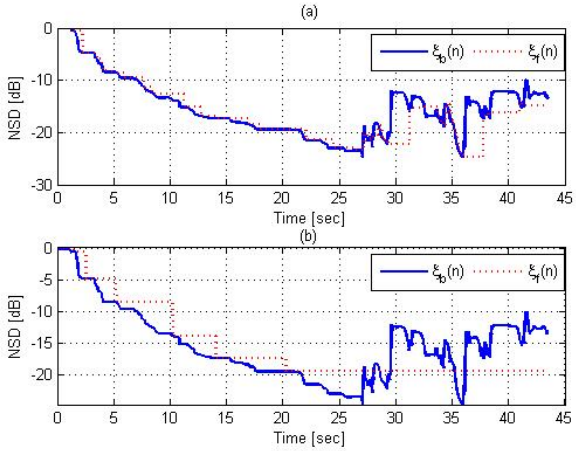


Fig. 5 NSD of background filter (solid line) and foreground filter (dotted line) in the CTP solution for two different α_{e_b,e_f} settings in a doubletalk situation and doubletalk occurs after 26 s. (a) $\alpha_{e_b,e_f} = -6\text{dB}$, non-robust. (b) $\alpha_{e_b,e_f} = -12\text{dB}$, robust

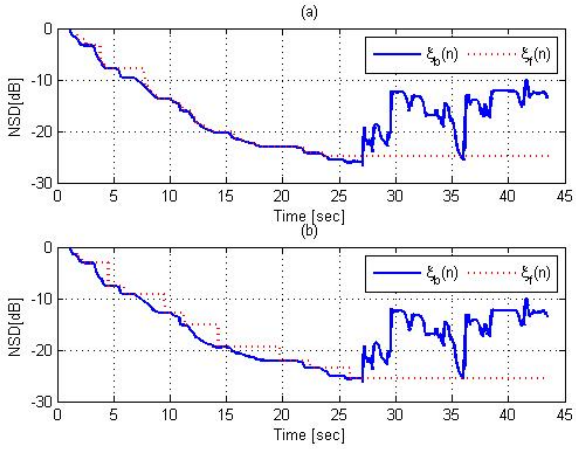


Fig. 6 NSD of background filter (solid line) and foreground filter (dotted line) for the ITP and proposed solution in a doubletalk situation. Doubletalk occurs after 26 s. (a) NSD for the ITP, $\alpha_{b,f} = 0\text{dB}$. (b) NSD for the proposed solution, $\alpha_0 = 0\text{dB}$

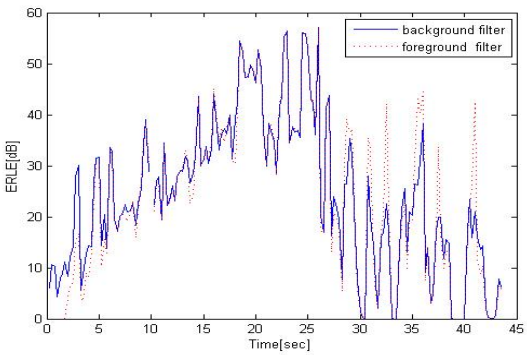


Fig. 7 ERLR of background filter (solid line) and foreground filter (dotted line) for the proposed solution in a doubletalk situation. Doubletalk occurs after 26 s

Open Science Index, Electronics and Communication Engineering Vol:5, No:3, 2011 publications.waset.org/11883/pdf

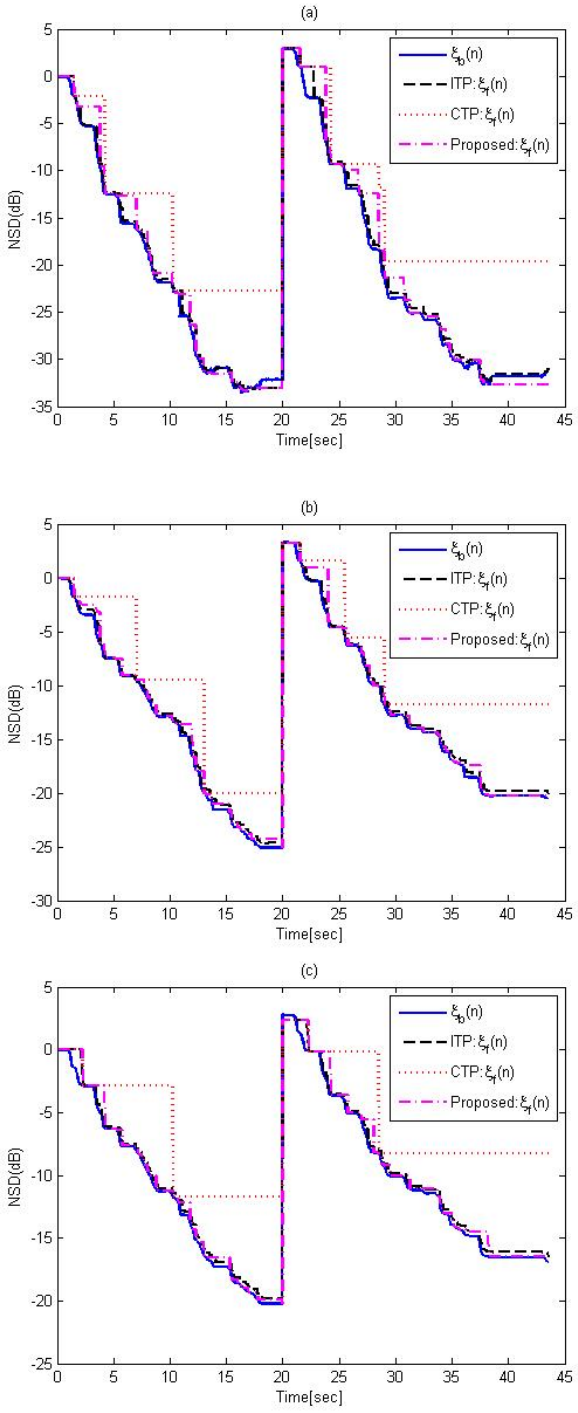


Fig. 8 NSD comparison of the CTP, ITP and proposed solutions for different step-size in the echo path change situation. Echo path change occurs from $\mathbf{h}_1$ to $\mathbf{h}_2$ at 20 s

(a) μ=0.4 . (b) μ=0.2 . (c) μ=0.1 .

Fig. 8 demonstrates the NSD comparison of three transfer logic solutions for different values of step size. The echo path change occurs at 20 s. It can be seen that both the ITP and proposed solution increase the convergence speed of the foreground filter, as compared to the CTP foreground filter.

V. Conclusion

An improved transfer logic solution of two-path algorithm for acoustic echo cancellation is proposed. This solution depends on the comparison between the ERLE of the background filter and reference ERLE. It not only increases the convergence speed of two-path algorithm, but also retains low computational complexity compared to the previous solutions. Simulation results show the efficiency of the proposed solution.

References

[1] J. Benesty, D. R. Morgan, and J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," *IEEE Transactions on. Speech and Audio Process*, vol. 8, no. 2, pp. 168-172, Mar. 2000.

[2] T. Gansler, H. Hansson, C. J. Ivarsson and G. Salomonsson, "A double-talk detector based on coherence," *IEEE Transactions on Communication* , vol. 44, no. 11, pp. 1421-1427, Nov. 1996.

[3] A. Mader, H. Puder, and G. U. Schmit, "Step-size control for acoustic cancellation filters–An overview," *Signal process*, vol. 80, pp. 1697-1719, 2000.

[4] D. L. Duttweiler, "A twelve-channel digital echo canceler," IEEE Transactions on Communication, vol. COM-26, no. 5, pp. 647-653, May. 1978.

[5] K. Ochiai, T. Araseki, and T. Ogihara, "Echo canceler with two echo path models," *IEEE Transactions on Communication*, vol. COM-25, no. 6, pp. 8-11, May. 1977.

[6] F. Lindstrom, C. Schuldt and I. Claesson, "An improvement of the two-path algorithm transfer logic for acoustic echo cancellation," *IEEE Transactions on Audio Speech and Language Process*., vol. 15, no. 4, pp. 1320-1326, May. 2007.

[7] F. Lindstrom, C. Schuldt and I. Claesson, "The two-path algorithm for line echo cancellation," *in Proceedings of the IEEE Tencon*, vol. 1, pp. 637-640, Nov .2004.

[8] J. Benesty and S. L. Gay, "An improved PNLMS algorithm, "An improved PNLMS algorithm," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP)* , vol. 2, pp. 857-860, May 1999.

[9] J. Liu, "A novel adaptation scheme in the nlms algorithm for the echo cancellation," *IEEE Signal Processing Letters*., vol. 8, no. 1, pp. 20-22, Jan. 2001.

[10] Y. Haneda, S. Makino, J. Kojima, and S. Shimauchi, "Implementation and evaluation of an acoustic echo canceler using the duo-filter control system," *in Proceedings of IWAENC* , pp.79-82 ,June. 1995,.

[11] F. Lindstrom, C. Schuldt, M. Langstrom and I. Claesson, "A method for reduced finite precision effects in parallel filtering echo cancellation," *IEEE Transactions on Circuits and System-I: Regular Papers.*,vol. 54, no. 9, pp. 2011-2018, Sep. 2007.