PAPER

A New Structure for Noise and Echo Cancelers Based on A Combined Fast Adaptive Filter Algorithm

Youhua WANG†, Kenji NAKAYAMA†† *and* Zhiqiang MA††, *Members*

SUMMARY This paper presents a new structure for noise and echo cancelers based on a combined fast adaptive algorithm. The main purpose of the new structure is to detect both the double-talk and the unknown path change. This goal is accomplished by using two adaptive filters. A main adaptive filter F_n, adjusted only in the non-double-talk period by the normalized LMS algorithm, is used for providing the canceler output. An auxiliary adaptive filter F_f, adjusted by the fast RLS algorithm, is used for detecting the double-talk and obtaining a near optimum tap-weight vector for F_n in the initialization period and whenever the unknown path has a sudden or fast change. The proposed structure is examined through computer simulation on a noise cancellation problem. Good cancellation performance and stable operation are obtained when signal is a speech corrupted by a white noise, a colored noise and another speech signal. Simulation results also show that the proposed structure is capable of distinguishing the near-end signal from the noise path change and quickly tracking this change.

key words: *adaptive filters, fast RLS algorithms, noise canceler, echo canceler, double-talk detection*

1. Introduction

Noise and echo cancellation problems have been very active research fields in the recent years, due to their variety of applications in communications, such as teleconference system and mobile phone system. The performance of noise and echo cancelers is usually evaluated by two important factors — convergence speed and residual error. In communication applications, the impulse response of the unknown path, in which noise or echo is transferred, is usually long and the number of tap weights for the adaptive filter is therefore large. Furthermore, the training signal used in these applications is often narrow-banded like speech signal. These make the convergence performance of the normalized LMS (NLMS) algorithm hard to be satisfactory. Seeking for improvement in convergence has been the main topic of some papers [1],[2].

The residual error is, however, closely related to the performance of a double-talk detector. It is known that the adjustment of the tap weights during the double-talk period will seriously degrade the performance. The conventional method, which compares the levels between the near-end and the far-end signals, cannot detect the double-talk precisely since the path loss, which is supposed to be known, is usually unknown. The use of two echo path models is an attractive method without using a double-talk detector [3],[4]. The basic idea of these methods is to form a foreground model and a background model. The result of the background model is transferred to the foreground model when the residual echo produced by the background model is smaller. One problem of these methods is that if there are some correlations between the far-end and the near-end signals during the double-talk period, a smaller residual echo model does not mean that the error of the tap-weight vector is also smaller [5]. So the transfer of the tap weights during this period may cause a wrong result. Another problem is that when the unknown path changes, the use of the NLMS algorithm may not be possible to distinguish this change from the double-talk due to its slow convergence rate.

We have proposed a method which combines the fast transversal filter (FTF) and the NLMS algorithms based on a single adaptive filter structure in order to achieve a fast convergence and stable performance with less computation [6]. With periodic reinitialization, the FTF algorithm can provide fast convergence and fast tracking when the unknown system changes fast. The degraded performance caused by the reinitialization process in the stationary state is compensated by using the NLMS algorithm. The improved performance in system identification has been shown in Ref. [6]. However, the structure of using only one adaptive filter is not suitable for noise and echo cancellation problems because it is difficult to distinguish the double-talk from the unknown path change.

In this paper, a new structure based on the combined algorithm is proposed, which attempts to solve the above mentioned problems. The main purpose of the new approach is to detect both the double-talk and the unknown path change. This goal is accomplished by using two adaptive filters. A main adaptive filter F_n, adjusted only in the non-double-talk period by the NLMS algorithm, is used for providing the canceler output. An auxiliary adaptive filter F_f, adjusted by the FTF algorithm, is used for detecting the double-talk and obtaining a near optimum tap-weight vector for F_n

Manuscript received June 6, 1994.
Manuscript revised November 22, 1994.
†The author is with the Graduate School of Natural Sci. & Tech., Kanazawa University, Kanazawa-shi, 920 Japan.
††The authors are with the Faculty of Technology, Kanazawa University, Kanazawa-shi, 920 Japan.

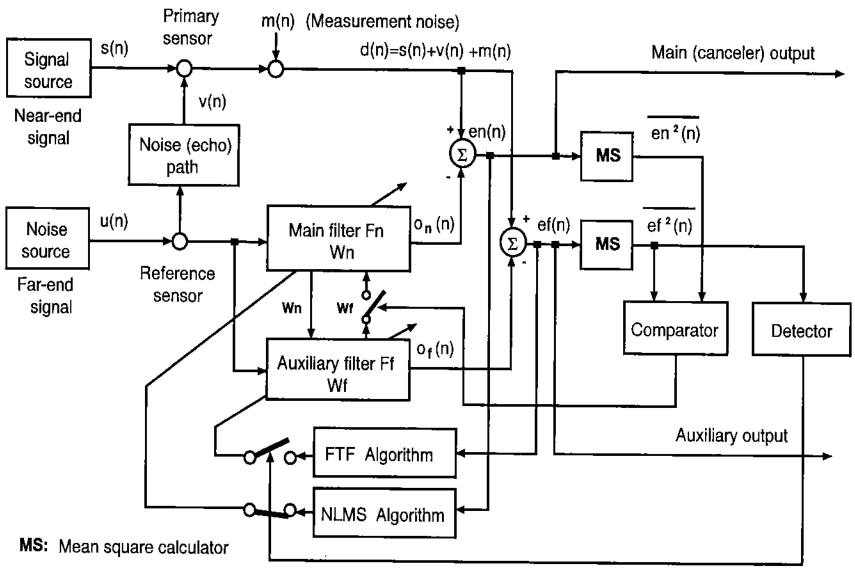


Fig. 1 New structure for noise and echo cancelers.

in the initialization period and whenever the unknown path has a sudden or fast change.

The remainder of this paper is organized as follows: The following section describes the proposed structure and its adaptive process. In Sect. 3, a theoretical analysis of the performance is presented. Section 4 contains several simulations on a noise cancellation problem. The results reveal the good performances and also validate the performance analysis. Comparison of the proposed method with the Sugaya's method proposed in Ref. [4] is also discussed.

2. A New Structure for Noise and Echo Cancellation

2.1 Structure Using Double Adaptive Filters

Figure 1 shows the proposed echo and noise canceler, which consists of two kinds of adaptive filters. One of them is a main adaptive filter F_n, whose tap-weight vector $\mathbf{w}_n$ is adjusted by the NLMS algorithm, and the other is an auxiliary adaptive filter F_f, whose tap-weight vector $\mathbf{w}_f$ is adjusted by the FTF algorithm. The output $o_n(n)$ from F_n is used as the canceler output. F_f is used in order to obtain fast convergence in the cases of the initial training, unknown path change and so on. The double-talk and the unknown path change are detected by using the mean square error (MSE) of F_f denoted $\overline{e_f^2(n)}$. Either F_n or F_f is adjusted by their own algorithm. This means they are not simultaneously adjusted. The tap-weight vectors $\mathbf{w}_n$ and $\mathbf{w}_f$ can be transferred to each other. The MSE of F_n denoted $\overline{e_n^2(n)}$ is compared with $\overline{e_f^2(n)}$ in order to select $\mathbf{w}_n$ or $\mathbf{w}_f$ that should be used.

2.2 Switching Method Between NLMS and FTF Algorithms

Two adaptation algorithms are switched based on the situations, that is the double-talk, the unknown path change and so on. These situations are detected using $\overline{e_f^2(n)}$. The switching scheme is summarized as follows:

(1) If $\overline{e_f^2(n)} > \Theta$, then the FTF algorithm is selected.

(2) If $\overline{e_f^2(n)} \leq \Theta$, then the NLMS algorithm is selected.

Θ is a prescribed threshold, which is mainly determined by experience. In Sect. 4, how to determine Θ will be further discussed. Efficiency of this switching scheme is further discussed for typical situations in the next section.

Following the above switching method, either F_n or F_f is adjusted. Thus, the tap-weight vectors are transferred to each other as follows:

(a) $\mathbf{w}_f$ is transferred to $\mathbf{w}_n$ as the initial guess only when the FTF algorithm is switched to the NLMS algorithm and $\overline{e_f^2(n)}$ is smaller than $\overline{e_n^2(n)}$.

(b) $\mathbf{w}_n$ is transferred to $\mathbf{w}_f$ whenever the NLMS algorithm is implemented.

2.3 Adaptation Process

In this section, the proposed switching method is further discussed in typical situations under which the echo and noise cancelers are used.

(A) Initial training period:

Since $\mathbf{w}_f$ starts from zero, $\overline{e_f^2(n)}$ is larger than Θ, then the FTF algorithm is selected. After convergence,

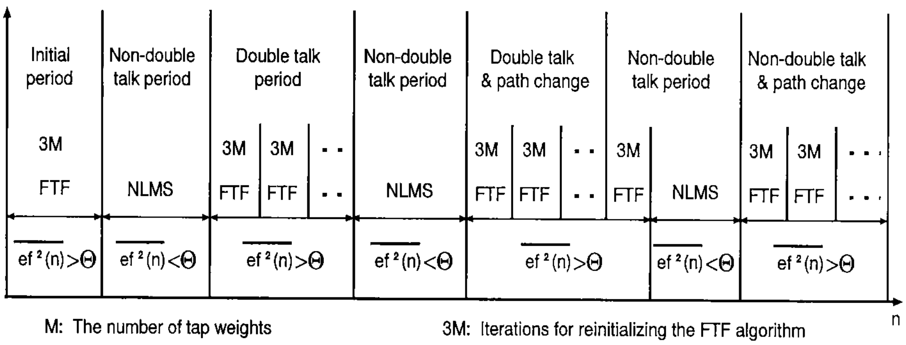


Fig. 2 Timing chart of adaptation process.

$\overline{e_f^2(n)} < \Theta$ can be satisfied. So the FTF algorithm is switched to the NLMS algorithm, and $\mathbf{w}_f$ is transferred to $\mathbf{w}_n$ as the initial guess. Since F_n is not adjusted, $\overline{e_n^2(n)}$ is larger than $\overline{e_f^2(n)}$.

(B) No double-talk and slow unknown path change:

After convergence and there is no double-talk, the NLMS algorithm is used, and $\mathbf{w}_n$ is always transferred to $\mathbf{w}_f$. Furthermore, even though the unknown path changes its characteristics, if the NLMS algorithm can track, then $\overline{e_f^2(n)}$, produced by $\mathbf{w}_f$ in F_f, can be smaller than Θ. Thus, the NLMS algorithm is used. In this period, F_f is equivalently adjusted by the NLMS algorithm.

(C) Double-talk period:

When double-talk appears, $\overline{e_f^2(n)}$ exceeds Θ. In other word, Θ is determined so as to satisfy $\overline{e_f^2(n)} > \Theta$ in the periods of the double-talk and the unknown path change. Therefore, the FTF algorithm is selected. During the double-talk period, $\overline{e_f^2(n)}$ can not be decreased under Θ even though the FTF algorithm is used. Thus, the FTF algorithm is continuously used in this period. In the main filter F_n, the tap-weight vector is fixed to the values just before the FTF is selected. Like this, F_n is not affected by the double-talk.

After the double-talk finishes, the FTF algorithm can converge quickly, and $\overline{e_f^2(n)}$ becomes smaller than Θ, then the FTF algorithm is switched to the NLMS algorithm. In this case, however, $\overline{e_f^2(n)}$ is not always smaller than $\overline{e_n^2(n)}$ because F_f is affected by the double-talk. If $\overline{e_n^2(n)} < \overline{e_f^2(n)}$, then $\mathbf{w}_f$ is not transferred to $\mathbf{w}_n$, and the previous $\mathbf{w}_n$, which is fixed in the double-talk period, is used as the initial guess.

In this case, the auxiliary adaptive filter F_f is required to converge very quickly, because the non-double-talk period is not always guaranteed to be long enough for adaptation.

(D) Unknown path change:

When the unknown path changes, $\overline{e_f^2(n)}$ also exceeds Θ, and the FTF algorithm is selected. In this case, however, the FTF algorithm can converge. After $\overline{e_f^2(n)}$ is decreased under Θ, the FTF algorithm is switched to the NLMS algorithm. At the same time, since $\mathbf{w}_n$ is not adjusted, $\overline{e_f^2(n)}$ can be smaller than $\overline{e_n^2(n)}$, and $\mathbf{w}_f$ is transferred to $\mathbf{w}_n$ as the initial guess. When the unknown path changes and the FTF algorithm is selected, the tap-weight vector of F_n is fixed, and its error is increased. Therefore, the fast convergence algorithm is required in the auxiliary adaptive filter.

(E) Double-talk and unknown path change:

The FTF algorithm is also selected when both the double-talk and the unknown path change occur. This situation is a combination of (C) and (D). $\overline{e_f^2(n)}$ can not decrease under Θ like in (C). However, the FTF algorithm can quickly track the unknown path change after the double-talk finishes. In this situation, the fast convergence is also very important. The following operations are the same as in (C) and (D).

Figure 2 shows the timing chart for some adaptation cases. From the figure, we can see that in order to overcome the instability problem that produced by the accumulation of round-off errors as well as the double-talk, the FTF algorithm is reinitialized at every $3M$ implementations, where M denotes the number of tap weights and $3M$ is the number of iterations for obtaining the convergence. The reason for choosing $3M$ will be explained in the following section.

We note that the use of the FTF algorithm in the auxiliary filter F_f makes it possible to provide a near-optimum tap-weight vector for the main filter F_n whenever there is a non-double-talk period. Thus the proposed structure is capable of detecting the double-talk precisely and distinguishing between the double-talk and the path change. This feature is, however, difficult to be realized by using only the NLMS algorithm because it may not be possible to converge during the non-double-talk period due to its slow convergence rate.

This problem becomes more serious if the unknown path changes continuously.

3. Properties of Proposed Method

3.1 Convergence Performance

The convergence performance of the RLS algorithm in the non-double-talk period can be shown as [8]

$$\xi(n) = E[\epsilon^2(n)] \approx \sigma_m^2 + \frac{M\sigma_m^2}{n} \tag{1}$$

where $\xi(n)$ denotes the mean square value of the residual error $\epsilon(n)$ and σ_m^2 denotes the variance of the measurement noise.

We note from Eq. (1) that the RLS algorithm converges in the mean square in about $2M$ iterations. In the case of using the FTF algorithm, reinitialization is needed to avoid the instability problem, this produces a transient period that needs additional M iterations [6]. So the FTF algorithm converges in about $3M$ iterations. The bias $\mathbf{b}(n)$ of the estimated tap-weight vector produced by a finite implementation of the FTF algorithm can be written as (see Appendix for the derivation)

$$\begin{aligned}\mathbf{b}(n) &= \mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n)] \\ &= \delta\lambda^n\Phi^{-1}(n)\mathbf{w}_{opt} \quad (n \geqq 1)\end{aligned} \tag{2}$$

where $\mathbf{w}_{opt}$ and $\hat{\mathbf{w}}(n)$ denote the optimum and the estimated tap-weight vectors, respectively. $\Phi^{-1}(n)$ denotes the inverse of the input correlation matrix. δ denotes the initial parameter and λ denotes the forgetting factor.

In practical applications, δ and λ should be chosen in order to achieve a stable performance for $3M$ iterations of the FTF algorithm. Experiment shows that the choice of δ that equals the power of the input can effectively prevent the instability in the initialization period (large-order effects) [7]. Then, under the condition of a 32-bit floating point arithmetic, the $3M$ iterations of the FTF algorithm is stable by using $\lambda = 0.9, 0.98, 0.99$ for $M = 50, 300, 500$, respectively. In the case of the white noise input, $\delta\Phi^{-1}(n)$ becomes an identity matrix. So by using λ and M described above and after $3M$ implementation of the FTF algorithm, the bias becomes

$$\mathbf{b}(3M) \approx \lambda^{3M}\mathbf{w}_{opt} \approx (10^{-6} \sim 10^{-7})\mathbf{w}_{opt} \tag{3}$$

This estimation is also valid for the speech signal input. So we can say that the FTF algorithm converges within $3M$ implementation.

3.2 Residual Error

In stationary state, the relation between the step size $\tilde{\mu}$ and the residual MSE ξ in the NLMS algorithm can be written as [6],[9]

$$\xi = E[\epsilon^2(\infty)] = \sigma_m^2 + \frac{\tilde{\mu}}{2-\tilde{\mu}}(\sigma_s^2 + \sigma_m^2) \tag{4}$$

where σ_s^2 is the variance of the near-end signal.

Since σ_s^2 is usually much greater than σ_m^2, the adjustment of the adaptive filter in the double-talk period will cause a large ξ. In the proposed structure, however, the NLMS algorithm is implemented only in the non-double-talk period so that $\sigma_s^2 \approx 0$. Therefore, ξ becomes very small.

In the noise cancellation problem, we recall that uncorrelation between signal and noise sources is one of the basic assumptions. Based on this assumption, MSE of the canceler output e_n in Fig. 1 can be written as

$$\begin{aligned}E[e_n^2] &= \sigma_m^2 + E[s^2] - 2E[s(v-o_n)] + E[(v-o_n)^2] \\ &\approx \sigma_m^2 + E[s^2] + E[(v-o_n)^2]\end{aligned} \tag{5}$$

Since $\sigma_m^2 + E[s^2]$ is a constant, minimizing the total power $E[e_n^2]$ is equivalent to minimizing the output noise power $E[(v-o_n)^2]$. However, when signal is correlated with noise, the term $2E[s(v-o_n)]$ in Eq. (5) will not be zero, which results in a degraded performance. In the new method, the coming of the near-end signal is detected and the adjustment of the main adaptive filter F_n in the double-talk period is avoided. So the minimization of the MSE process satisfies

$$E[e_n^2] \approx \sigma_m^2 + E[(v-o_n)^2] \tag{6}$$

This means that the minimum MSE solution is always achievable.

3.3 Computational Requirement

Computational cost increase by using the new method is an additional M computations per iteration for parallel implementation of the two adaptive filters. So for every iteration, $8M$ computations are required when the FTF algorithm is selected and $3M$ computations are required when the NLMS algorithm is selected.

4. Computer Simulations

The new structure can be applied to both noise and echo cancellation problems. In this section, we will do some simulations on a noise cancellation problem. The results of applying the new structure to an echo cancellation problem has been shown in Ref. [10].

The simulations are divided into two cases. In the first case, we suppose a speech signal corrupted by (1) white noise (2) colored noise (3) another speech. In the second case, we discuss the change of the noise path.

4.1 Conditions of Simulation

The speech signal used for the simulation consists of a segment of a male voice, sampled at 10 KHz as shown in Fig. 3. The variance of the speech signal σ_s^2 is normalized to unity. The noise is mixed with the signal

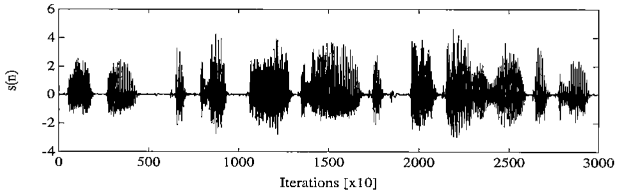


Fig. 3 Speech signal used for simulation.

after passing through a noise path, which is assumed to be a second-order all pole filter. The transfer function can be written as

$$H(z) = \frac{1}{1 - 2r_p cos(\theta) z^{-1} + r_p^2 z^{-2}} \quad (7)$$

In the simulation, $r_p = 0.5$ and $\theta = \frac{\pi}{4}$ are used. The number of tap weights used for both the main and the auxiliary filters is 20.

4.2 Choice of Threshold

Since the main filter F_n, which is adjusted by the NLMS algorithm, is adapted only when $\overline{e_f^2(n)} \leq \Theta$, from Eq. (4), the residual MSE ξ can be written as

$$\xi = E[\epsilon^2(\infty)] = \xi_{min} + \xi_{ex} \approx \sigma_m^2 + \frac{\tilde{\mu}}{2-\tilde{\mu}}\Theta \quad (8)$$

where $\xi_{min} = \sigma_m^2$ is the minimum residual MSE and ξ_{ex} is the extra residual MSE shown by

$$\xi_{ex} \approx \frac{\tilde{\mu}}{2-\tilde{\mu}}\Theta \quad (9)$$

Equation (9) gives the relations among ξ_{ex}, the step size $\tilde{\mu}$ and the threshold Θ.

In practical applications, ξ_{ex} can be chosen between $0.1\sigma_m^2$ and σ_m^2, which represents a misadjustment from 10% to 100%. Then by choosing an appropriate step size $\tilde{\mu}$, we can determine the threshold Θ, and vice versa. Experiments show that a choice of $\tilde{\mu}$, that makes Θ in the range from $10\sigma_m^2$ to $100\sigma_m^2$ can provide satisfactory results. For example, in the simulations that follows, suppose $\sigma_m^2 = 0.001$ and $\xi_{ex} = 0.1\sigma_m^2 = 10^{-4}$. If we choose a threshold $\Theta = 0.01$ (or $\Theta = 0.1$), then from Eq. (9), the step size can be chosen as $\tilde{\mu} = 0.02$ (or $\tilde{\mu} = 0.002$). Satisfactory results are obtained for both selections of the threshold, which will be shown in Sect. 4.3.1.

The other parameters are chosen as: $\tilde{\mu} = 0.02$ for the NLMS algorithm, $\lambda = 0.95$ and $\delta = 1$ for the FTF algorithm. From Eq. (3), the bias after $3M$ implementation is about $4.6\%\mathbf{w}_{opt}$.

4.3 Simulation Results and Discussions

4.3.1 Change of Noise Source

Three properties are investigated in this simulation, those are improved SNR after cancellation, sensitivity to the eigenvalue spread of the noise source and the correlation between signal and noise sources.

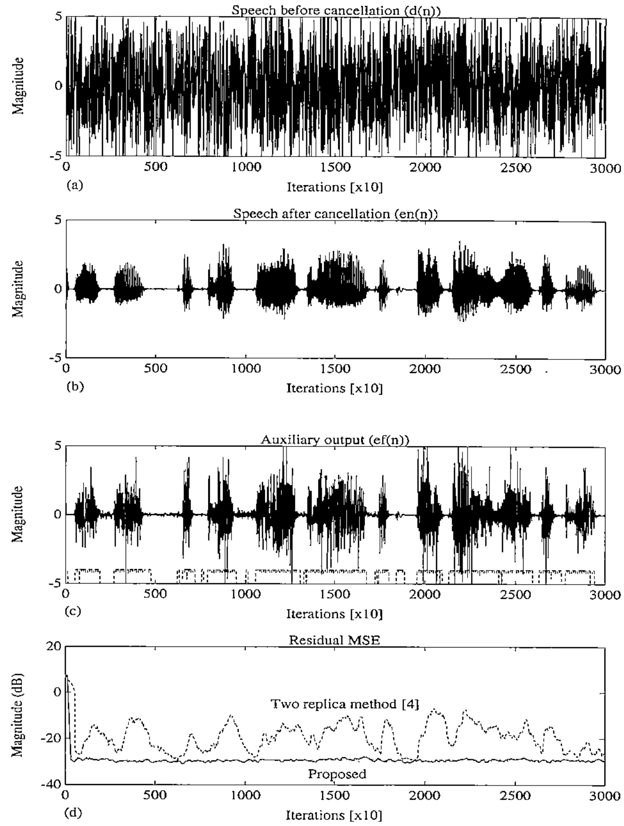


Fig. 4 Speech corrupted by white noise. (a) Speech before cancellation, (b) Speech after cancellation, (c) Auxiliary output (solid line) and interval of implementing the FTF algorithm (dashed line) and (d) Residual MSE: $\Theta = 0.01$, $\tilde{\mu} = 0.02$ (solid line), $\Theta = 0.1$, $\tilde{\mu} = 0.002$ (dotted line).

A workstation noise is used as a colored noise. The eigenvalue spread of this noise is about 100. When the noise is assumed to be another speech signal, the variance of both signal and noise sources are normalized to unity.

The simulation results are shown in Figs. 4 through 6, which include the results of using the two replica method proposed in Ref. [4] (Step size in model 1: $\tilde{\mu}_1 = 0.02$; Step size in model 2: $\tilde{\mu}_2 = 1 - \tilde{\mu}_1 = 0.98$. The other conditions are the same as in Ref. [4]). The SNR before and after cancellation is shown in Table 1.

From these results, we make the following observations

- Fast convergence rates are obtained for the white noise, the colored noise and the speech signal inputs. The proposed method is insensitive to the eigenvalue spread of the noise source due to the property of the RLS algorithm. The two replica method is sensitive to the eigenvalue spread of the noise source since only the NLMS algorithm is used.

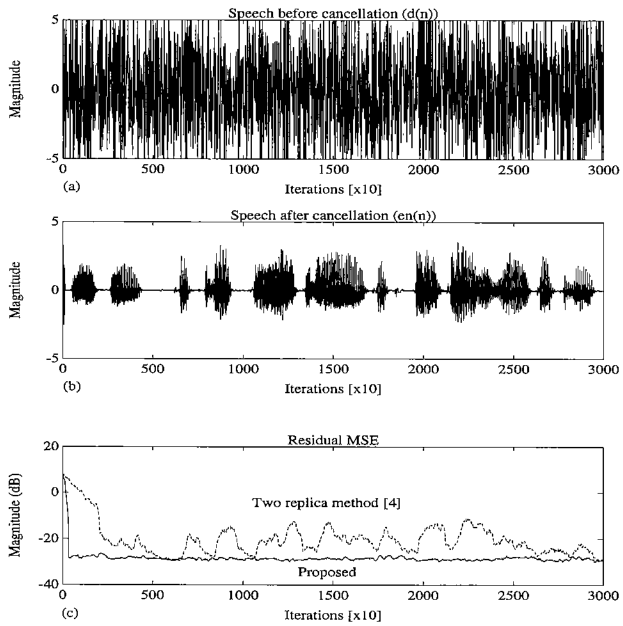


Fig. 5 Speech corrupted by colored noise. (a) Speech before cancellation, (b) Speech after cancellation and (c) Residual MSE.

- Whenever the speech signal appears, the FTF algorithm is selected and the adjustment of the tap weights in the NLMS algorithm is stopped (see Fig. 4 (c)). So there is almost no degradation of the performance during the double-talk period. The result of using the two replica method is not satisfactory due to the fail in detecting the double-talk. This problem will be further discussed in Sect. 4.4.
- Figure 4 (d) also shows that satisfactory results are achieved for both selections of the threshold ($\Theta = 0.01$ and 0.1), which confirms the analysis described in Sect. 4.2.
- The new structure is readily applicable to an adaptive echo canceler. The echo cancellation can be considered as an example of noise cancellation, in which the noise is an echo that is, in fact, another speech signal.

4.3.2 Change of Noise Path

The change of the noise path is realized by changing the phase of pole θ in Eq. (7) as shown in Fig. 7. The noise path has a sudden change at 5,000 and 25,000 iterations, a slow change from 10,000 to 15,000 and a fast change from 15,000 to 20,000 iterations.

The simulation results are shown in Fig. 8. From these results, we make the following observations:

- When the noise path changes slowly, the NLMS algorithm can track this change. When the noise path changes suddenly in the non-double-talk period, a very fast convergence rate is achieved. When the speech signal appears, the convergence is delayed until the next non-speech period (see Figs. 8 (c) and (d)). This is because $\mathbf{w}_f$ adjusted by the FTF algorithm will not be transferred to $\mathbf{w}_n$ in the double-talk period.

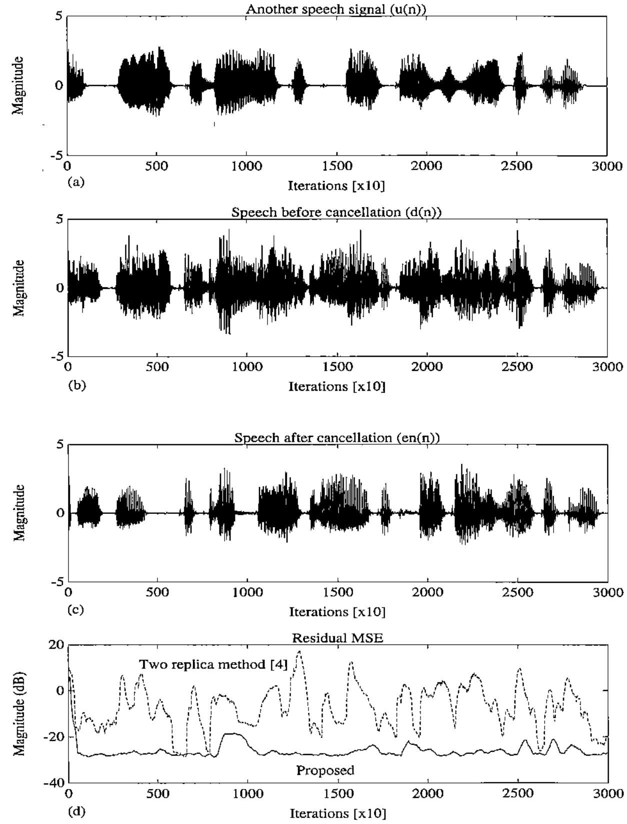


Fig. 6 Speech corrupted by another speech. (a) Another speech signal, (b) Speech before cancellation, (c) Speech after cancellation and (d) Residual MSE.

Table 1 SNR before and after cancellation.

Signal source	Noise source	Before cancellation	After cancellation
Speech #1	White noise	-10dB	29dB
Speech #1	Colored noise	-10dB	29dB
Speech #1	Speech #2	0dB	26dB

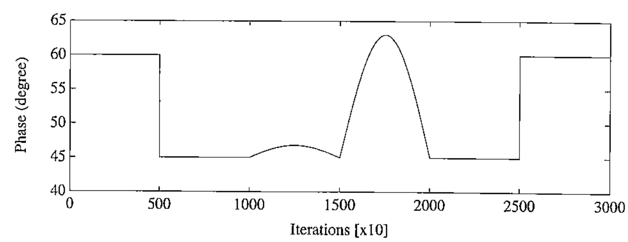


Fig. 7 Change of the pole phase in the noise path.

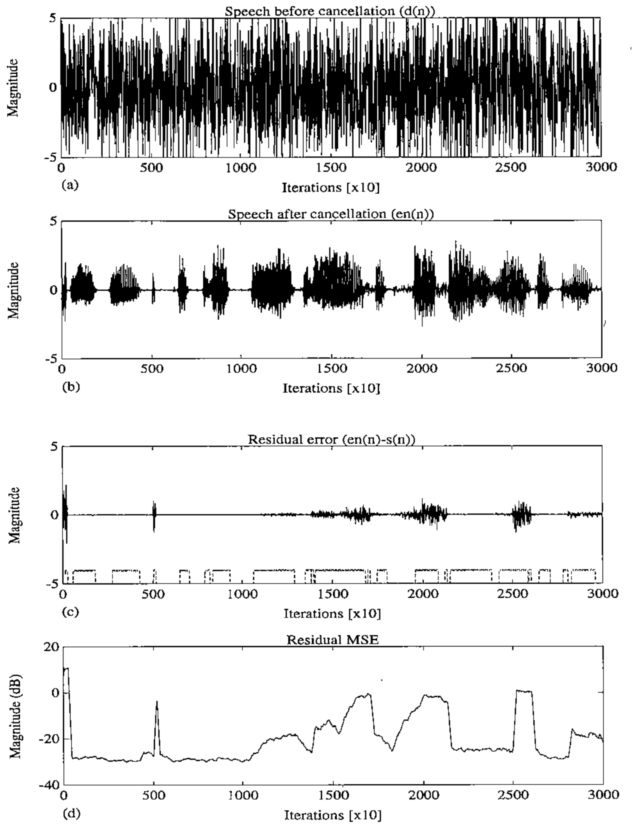


Fig. 8 Simulation results when the noise path changes. (a) Speech signal before cancellation, (b) Speech signal after cancellation, (c) Noise after cancellation (solid line) and interval of implementing the FTF algorithm (dashed line) and (d) Residual MSE.

- When the noise path changes fast, fast tracking is achieved by reciprocally adjusting F_f and F_n. The tap-weight vector of F_f is transferred to F_n whenever the transfer conditions described in Sect. 2.2 are satisfied.

The simulation results shown in Fig. 8 demonstrate that the proposed method is capable of distinguishing the noise path change from the near-end signal and quickly tracking this change.

4.4 Comparison with Conventional Method

The use of the double adaptive filters adjusted by the NLMS algorithm for echo cancellation problem has been proposed [3],[4]. One problem of these methods is that it is difficult to give the condition for detecting the double-talk. Another problem is the slow convergence rate especially when the training signal is the speech signal. In this section, we will discuss the two replica method proposed in Ref. [4].

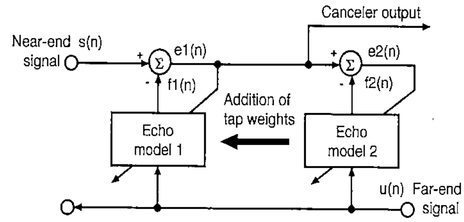


Fig. 9 Echo canceler with two quasi-echo estimators.

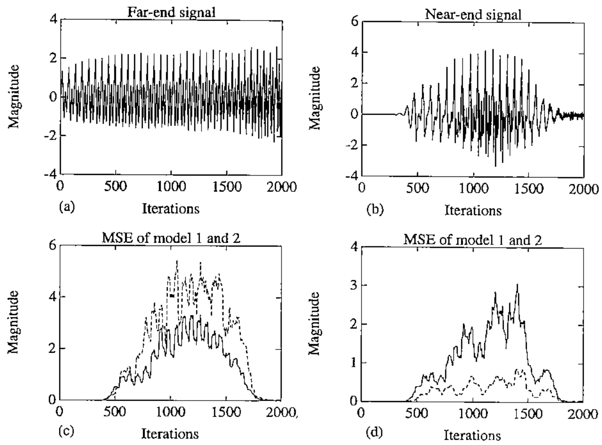


Fig. 10 Simulation results on double-talk detection, (a) Far-end signal (speech), (b) Near-end signal (speech), (c) $\overline{e_1^2(n)}$ (solid line) and $\overline{e_2^2(n)}$ (dashed line) when the far-end signal is white noise, (d) $\overline{e_1^2(n)}$ (solid line) and $\overline{e_2^2(n)}$ (dashed line) when the far-end signal is the speech signal shown in (a).

4.4.1 Double-Talk Detection

Figure 9 shows the structure of the echo canceler proposed in Ref. [4]. The condition used for detecting the double-talk and stopping the adaptation of the echo model 1 (main filter) is given by

$$\frac{\overline{e_2^2(n)}}{\overline{e_1^2(n)}} > 10\,\mathrm{dB} \quad \text{(continuous over 10 times)} \qquad (10)$$

Simulation shows, however, that this condition is not always valid. Figure 10 gives the simulation results that uses the unknown path and the number of tap weights as shown in Sect. 4.1. The step sizes used for the echo model 1 and 2 (auxiliary filter) are 0.1 and 0.9, respectively. Since we want to investigate the operation of this method after convergence, both models are assumed to have converged at the beginning of the iterations. As shown in Fig. 10 (c), when the white noise is used for the far-end (not shown in the figure) and the speech is used for the near-end signals, $\overline{e_2^2(n)}$ is greater than $\overline{e_1^2(n)}$. However, when the speech is used for both the far-end and the near-end signals, $\overline{e_2^2(n)}$ is smaller than $\overline{e_1^2(n)}$.

The reason for this phenomenon can be explained as follows. Since convergence is assumed and a small step size is chosen for the echo model 1, when the double-talk comes, we have $e_1(n) \approx s(n)$ and $e_2(n) \approx s(n) - f_2(n)$. The MSE of these errors can be written as

$$\overline{e_1^2(n)} \approx \overline{s^2(n)} \tag{11}$$

$$\begin{aligned}\overline{e_2^2(n)} &\approx \overline{(s(n) - f_2(n))^2} \\ &= \overline{s^2(n)} - 2\overline{s(n)f_2(n)} + \overline{f_2^2(n)}\end{aligned} \tag{12}$$

In the former case, the white noise is considered as uncorrelated with the speech. From Eq. (12), we have $2\overline{s(n)f_2(n)} \approx 0$. So $\overline{e_2^2(n)}$ will be greater than $\overline{e_1^2(n)}$. In the latter case, the speech signals have some correlations. As described in Sect. 3.1, the term $2\overline{s(n)f_2(n)}$ will not be zero. Furthermore, the sign of this term is indefinite. So we can not say that $\overline{e_2^2(n)}$ is always greater than $\overline{e_1^2(n)}$ during the double-talk period. Thus, it is rather difficult to detect the double-talk by using Eq. (10). This fact has been demonstrated by the simulation results shown in Figs. 4–6. In these simulations, we found that it was difficult to select a unified threshold for detecting the double-talk. In fact, the residual error produced in the stationary state is mainly due to the fail in detecting the double-talk.

4.4.2 Convergence Performance

The simulation results on convergence performance without the near-end signal are shown in Fig. 11. In the case of the white noise input, both the proposed and the two replica methods give a fast convergence rate and a small residual error. However, in the case of the speech signal input, the deterioration of the convergence rate by using the two replica method is significant. This is

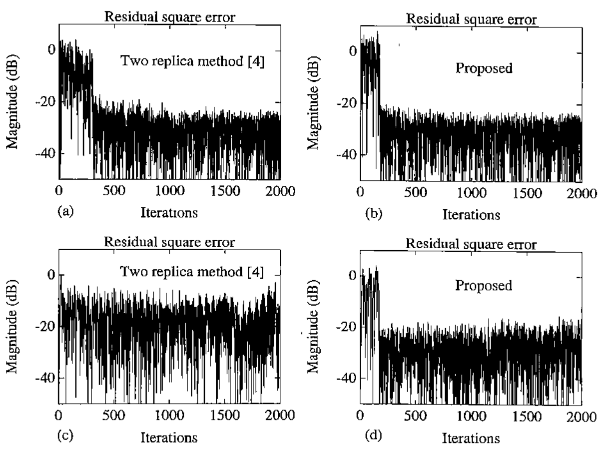


Fig. 11 Comparison of convergence performance. Far-end signal: White noise and speech signal (shown in Fig. 10 (a)); Near-end signal: None; Number of tap weights: $M = 50$; The two replica method: $\tilde{\mu}_1 = 0.1$, $\tilde{\mu}_2 = 0.9$; Proposed: $\delta = 1$, $\lambda = 0.95$ (FTF), $\tilde{\mu} = 0.1$ (NLMS), (a) The two replica method (white noise input), (b) Proposed (white noise input), (c) The two replica method (speech input), (d) Proposed (speech input).

because the NLMS algorithm is essentially sensitive to the eigenvalue spread of the input signal. No big deterioration occurs by using the proposed method.

5. Conclusion

A new structure for noise and echo cancelers based on the combined fast adaptive algorithm has been proposed. The main adaptive filter adjusted by the NLMS algorithm and the auxiliary adaptive filter adjusted by the FTF algorithm are combined. These algorithms are switched based on the mean square errors of both adaptive filters. Through computer simulation on the noise cancellation problem, the following features have been confirmed. Performance of the adaptive filter is not affected by the eigenvalue spread and the correlation between the signal and noise sources. The near-end signal is precisely detected and distinguished from the noise path change. Thus, the proposed method can provide a good and stable performance for noise and echo canceler applications.

References

[1] Sugiyama, A., et al., "A fast convergence algorithm for adaptive FIR filters," *Proc. ICASSP*, Glasgow, Scotland, pp.892–895, 1989.
[2] Fujii, K. and Ohga, J., "A fast adaptive algorithm suitable for acoustic echo canceler," *Trans. IEICE*, vol.E75-A, no.11, pp.1509–1515, Nov. 1992.
[3] Ochiai, K., et al., "Echo canceler with two echo path models," *IEEE Trans. Commun.*, vol.COM-25, pp.589–595, Jun. 1977.
[4] Sugaya, F., et al., "An echo canceler with two quasi-echo estimators," *'91 National Spring Conference*, vol.3, p.310, Mar. 1991.
[5] Tsujii, S., *Technology of Echo Canceler* (in Japanese), Japan Industrial Technology Center, 1986.
[6] Wang, Y. and Nakayama, K., "A combined fast adaptive filter algorithm with an automatic switching method," *Trans. IEICE*, vol.E77-A, no.1, pp.247–256, Jan. 1994.
[7] Cioffi, J.M. and Kailath, T., "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol.ASSP-32, no.2, pp.304–337, Apr. 1984.
[8] Haykin, S., *Adaptive Filter Theory*, Second Edition, Prentice-Hall, 1991.
[9] Widrow, B. and Stearns, D., *Adaptive Signal Processing*, Prentice-Hall, 1985.
[10] Wang, Y., Nakayama, K. and Ma, Z., "A combined fast adaptive filter algorithm with automatic switching method applied to echo cancellation," *IEICE Technical Report*, CS93-155, DSP93-79, Jan. 1994.

Appendix: Derivation of Eq. (2)

From the RLS algorithm, we know that [8]

$$\begin{aligned}e(n) &= d(n) - \hat{\mathbf{w}}^T(n-1)\mathbf{u}(n) \\ &= \left(\mathbf{w}_{opt}^T - \hat{\mathbf{w}}^T(n-1)\right)\mathbf{u}(n) + e_m\end{aligned} \tag{A$\cdot$1}$$

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{\Phi}^{-1}(n)\mathbf{u}(n)e(n) \tag{A$\cdot$2}$$

where $e(n)$ is the estimation error. e_m is assumed to be the measurement error and $E[e_m] = 0$.

Replacing Eq. (A· 1) into Eq. (A· 2), we get

$$\begin{aligned}\hat{\mathbf{w}}(n) &= \hat{\mathbf{w}}(n-1) + \mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\big(d(n) \\ &\quad -\mathbf{u}^T(n)\hat{\mathbf{w}}(n-1)\big) \\ &= \hat{\mathbf{w}}(n-1)+\mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\big(\mathbf{u}^T(n)\mathbf{w}_{opt}+e_m\big) \\ &\quad -\mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\hat{\mathbf{w}}(n-1) \qquad (A\cdot 3)\end{aligned}$$

The bias $\mathbf{b}(n)$ is defined as

$$\begin{aligned}\mathbf{b}(n) &= \mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n)] \\ &= \mathbf{w}_{opt} - \Big(E[\hat{\mathbf{w}}(n-1)] \\ &\quad +\mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}_{opt} \\ &\quad +\mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)E[e_m] \\ &\quad -\mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)E[\hat{\mathbf{w}}(n-1)]\Big) \\ &= \Big(\mathbf{I} - \mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\Big) \\ &\quad \cdot(\mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n-1)]) \qquad (A\cdot 4)\end{aligned}$$

Since

$$\mathbf{\Phi}(n) = \lambda\mathbf{\Phi}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n) \qquad (A\cdot 5)$$

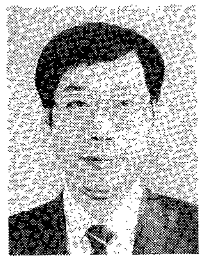
So

$$\begin{aligned}&\mathbf{I} - \mathbf{\Phi}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n) \\ &\quad = \mathbf{I} - \mathbf{\Phi}^{-1}(n)\big(\mathbf{\Phi}(n) - \lambda\mathbf{\Phi}(n-1)\big) \\ &\quad = \lambda\mathbf{\Phi}^{-1}(n)\mathbf{\Phi}(n-1) \qquad (A\cdot 6)\end{aligned}$$

Replacing these results into Eq. (A· 4), yields

$$\begin{aligned}\mathbf{b}(n) &= \mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n)] \\ &= \lambda\mathbf{\Phi}^{-1}(n)\mathbf{\Phi}(n-1)(\mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n-1)] \\ &= \lambda\mathbf{\Phi}^{-1}(n)\mathbf{\Phi}(n-1)\big(\lambda\mathbf{\Phi}^{-1}(n-1)\mathbf{\Phi}(n-2)\big) \\ &\quad \cdot\big(\mathbf{w}_{opt} - E[\hat{\mathbf{w}}(n-2)]\big) \\ &= \lambda^n\mathbf{\Phi}^{-1}(n)\mathbf{\Phi}(0)\big(\mathbf{w}_{opt} - E[\hat{\mathbf{w}}(0)]\big) \qquad (A\cdot 7)\end{aligned}$$

Setting $\mathbf{\Phi}(0) = \delta\mathbf{I}$ and $\hat{\mathbf{w}}(0) = \mathbf{0}$, we get the final result of Eq. (2).



Youhua Wang received the M.S. degree in 1988 from Shanghai University of Technology, Shanghai, China, and the Ph.D. degree in 1995 from Kanazawa University, Kanazawa, Japan. From 1988 to 1990, he was a Research Assistant at the Advanced Technology Research Center, Shenzhen University, Shenzhen, China. He is presently working with Matsushita Communication Kanazawa R&D Laboratories Co. Ltd., Kanazawa, Japan. His current research interests include digital signal processing, adaptive filtering, speech coding, and their applications to communications.

Kenji Nakayama received the B.E. and Dr. degrees in electronics engineering from Tokyo Institute of Technology (TIT), Tokyo, Japan, in 1971 and 1983, respectively. From 1971 to 1972 he was engaged in the research on classical network theory in TIT. He was involved in NEC Corporation from 1972 to 1988, where his research subjects were filter design methodology and signal processing algorithms. He joined the Department of Electrical and Computer Engineering at Kanazawa University, in Aug. 1988, where he is currently a Professor. His current research interests include neural networks and adaptive filters.

Zhiqiang Ma received the B.S., M.S. degrees from Dept. of Automatic Control, Central South University of Technology, China, in 1975 and 1982, respectively. She received the Ph.D. degree in 1991, from Dept. of Natural Science, Kanazawa University, Japan. She was with the Hunan Microwave Communication Company (1975–79), Central South University of Technology (1982–86), Fukui Institute of Technology (1986–87), and C & C Systems Research Laboratories, NEC Corp (1991–92). Since 1992, she has been a lecturer in Dept. of Electrical and Computer Engineering, Kanazawa University. Her research interests include analog, digital and adaptive signal processing. She is a member of IEEE.