

WANG Zhen-li, ZHANG Xiong-wei, YANG Ji-bin, CHEN Gong

Study of a new fast adaptive filtering algorithm

© Higher Education Press and Springer-Verlag 2006

Abstract A new fast adaptive filtering algorithm was presented by using the correlations between the signal's former and latter sampling times. The proof of the new algorithm was also presented, which showed that its optimal weight vector was the solution of generalized Wiener equation. The new algorithm was of simple structure, fast convergence, and less stable maladjustment. It can handle many signals including both uncorrelated signal and strong correlation signal. However, its computational complexity was comparable to that of the normalized least-mean-square (NLMS) algorithm. Simulation results show that for uncorrelated signals, the stable maladjustment of the proposed algorithm is less than that of the VS-NLMS algorithm, and its convergence is comparable to that of the algorithm proposed in references but faster than that of L.E-LMS algorithm. For strong correlation signal, its performance is superior to those of the NLMS algorithm and DCR-LMS algorithm.

Keywords adaptive filter, NLMS algorithm, Wiener equation, correlation

1 Introduction

Adaptive filtering technique has been widely used in echo cancellation, system identification, spectrum analysis and beamforming etc.. Figure 1 shows the principle of adaptive filter. Where $v(n)$ denotes a disturbed noise vector. The idea behind adaptive filter is to give an estimation $\hat{y}(n)$ of the output signal $y(n)$ of an unknown system. And the error signal $e(n)$ between $\hat{y}(n)$ and $y(n)$ in accordance with certain algorithm is used to adjust the parameters of

the adaptive filter so as to approach the parameters of the unknown system. In the past least mean square (LMS) algorithm based on the steepest descent method was widely employed for its simple structure and easy implementation etc.. However, its convergence speed is slow and its performance is reduced by gradient noise.

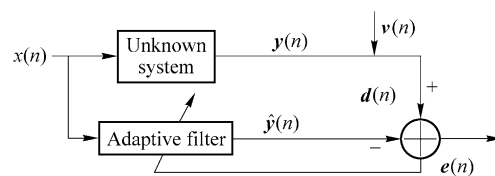


Fig. 1 The principle of adaptive filter

Moreover, it performs badly for processing high correlation signal. To overcome these limitations, some researchers propose NLMS algorithm [1], affine projection algorithm (APA) [2–5] and variable step size LMS algorithms [6–9]. APA, which is of fast convergence speed and high computation precision, has good performance for processing strong correlation signal. Unfortunately, its application is limited for large computation complexity. NLMS algorithm is for small computation complexity but large stable misadjustment. A variable step size algorithm with gradual reduced step size $\mu(n)$ versus increasing iteration times is presented in Ref. [6], which has very little stable error for non-time-variant system. However, it can't run in accordance with time-variant systems. The step parameter μ of the variable step size algorithm proposed in Ref. [7] is directly proportional to the estimation of the cross-function between $e(n)$ and $X(n)$. It still has good performance when disturbed noise increases. However, large computation complexity limits its application. By processing the error signal $e(n)$ non-linear, the variable step L.E-LMS algorithm with moderate computation complexity and little stable misadjustment but not enough convergence speed is presented in Ref. [8]. The variable step algorithm with non-linear functional relationship between step parameter μ and error signal $e(n)$ is presented in Ref. [9], which has better performance for processing uncorrelated signal

Translated from *Journal on Communications*, 2005, 26(11): 1–6 (in Chinese)

WANG Zhen-li (✉), ZHANG Xiong-wei, YANG Ji-bin, CHEN Gong
Institute of Communications Engineering, People's Liberation Army
University of Science and Technology, Nanjing 210007, China
E-mail: dongwen3619@163.com

but rapid reduced performance for processing strong correlation signal. A decorrelation DCR-LMS algorithm with very good performance for processing strong correlation signal is presented in Ref. [11]. However, it performs as LMS algorithm for uncorrelated signal inputs [12]. A variable step size decorrelation VSDLMS algorithm by injecting pseudorandom code is presented in Ref. [12], which has good performance versus increasing system cost.

In short, all the above algorithms do not have perfect performances. Is there an algorithm that can process both uncorrelated signal and correlation signal with good performance? In this paper, a new adaptive filtering algorithm based on the correlation of signals is presented. This new algorithm has very good performance for processing not only uncorrelated signal but also strong correlation signal.

It is also of simple structure, fast convergence speed, and comparable computation complexity with the NLMS algorithm.

2 The structure and convergence analysis of the new algorithm

The input vector of the adaptive filter is defined as $\mathbf{X}(n) = [x(n), x(n-1), \dots, x(n-M+1)]^H$ at sampling time n , where M and H denote the filter order and conjugate transpose of a vector or matrix, respectively. Its corresponding weight vector is defined as $\hat{\mathbf{W}}(n) = [w(n), w(n-1), \dots, w(n-M+1)]^T$, and the error signal is defined as:

$$\boldsymbol{\varepsilon}(n) = \mathbf{d}(n) - \hat{\mathbf{W}}^H(n) \mathbf{X}(n) \quad (1)$$

where $\mathbf{d}(n)$ denotes the desired signal. The correlation coefficient $\gamma(n)$ of the vector $\mathbf{X}(n)$ between sampling time n and sampling time $n-1$ is defined in Ref. [10]:

$$\gamma(n) = \frac{\mathbf{X}^H(n) \mathbf{X}(n-1)}{\mathbf{X}^H(n-1) \mathbf{X}(n-1)} \quad (2)$$

The correlation components between $\mathbf{X}(n)$ and $\mathbf{X}(n-1)$ are subtracted from $\mathbf{X}(n)$. A decorrelation vector is then defined as the Refs. [11, 12]

$$\mathbf{U}(n) = \mathbf{X}(n) - \gamma(n) \mathbf{X}(n-1) \quad (3)$$

From Eqs. (2)–(3), we can obtain the following expression:

$$\begin{aligned} & \mathbf{X}^H(n-1) \mathbf{U}(n) \\ &= \mathbf{X}^H(n-1) \mathbf{X}(n) - \mathbf{X}^H(n-1) \frac{\mathbf{X}(n-1) \mathbf{X}^H(n-1)}{\mathbf{X}^H(n-1) \mathbf{X}(n-1)} \mathbf{X}(n) \quad (4) \\ &= \mathbf{X}^H(n-1) \mathbf{X}(n) - \mathbf{X}^H(n-1) \mathbf{X}(n) = 0 \end{aligned}$$

Equation (4) indicates that the vector $\mathbf{U}(n)$ is orthogonal to the signal vector $\mathbf{X}(n-1)$ at sampling time $n-1$. Thus, the convergence speed of the proposed algorithm is increased. Considering that $\mathbf{U}(n)$ denotes the uncorrelated components of the signal vector $\mathbf{X}(n)$ with the

former instant vector, we define the energy of the signal at current sampling time as the following expression:

$$\begin{aligned} \beta(n) &= \mathbf{X}^H(n) \mathbf{U}(n) \\ &= \mathbf{U}^H(n) \mathbf{U}(n) + \gamma(n) \mathbf{X}^H(n-1) \mathbf{U}(n) \\ &= \|\mathbf{U}(n)\|^2 \end{aligned} \quad (5)$$

where $\beta(n)$ is also named as normalized energy parameter. Assuming that the decorrelation vector $\mathbf{U}(n)$ is regarded as the input vector of the filter at sampling time n , the updating weights equation of the proposed algorithm can be written as

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\|\mathbf{U}(n)\|^2 + \delta} \mathbf{U}(n) \mathbf{e}^*(n) \quad (6)$$

where, $\tilde{\mu}$ is a real number which can be adjusted; δ ($\delta > 0$) is a real constant used to control the magnitude of stable misadjustment and the convergence speed of this algorithm. It is followed that the function of δ is shown via uncorrelated white Gaussian noise. We assume that the order of adaptive filter equals to 5. The weight of the unknown system is defined as $\mathbf{W} = [0.25, 0.75, 1, 0.75, 0.25]^T$. Both the reference input signal $\mathbf{X}(n)$ and the vector $\mathbf{V}(n)$ are zero-mean white Gaussian randoms, with variance 1 and 0.01, respectively, and these two vectors are uncorrelated. Statistical mean times and samples are 200 and 600, respectively. The real number is set to $\tilde{\mu} = 1.5$.

Figure 2 shows that the new algorithm has the fastest convergence speed and the least stable maladjustment when $\delta = 1$, $\delta = 10$, respectively. In its application, δ can be obtained by considering the tradeoff between convergence speed and stable maladjustment. The proposed algorithm reduces to DCR-LMS algorithm when $\delta = 0$ and $\|\mathbf{U}(n)\|^2$ is substituted by $\mathbf{X}^H(n) \mathbf{U}(n)$ [12]. However, the proof of its algorithm is not presented in Ref. [12].

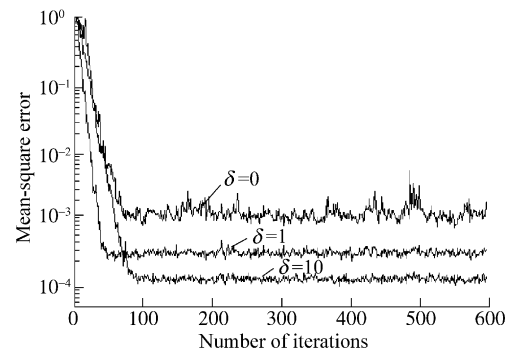


Fig. 2 Convergence curves

In the following, we will show that the mean value of the filtering weight is close to the solution of generalized Wiener equation when iterative times increases infinitely. By substituting Eq. (1) into Eq. (6), the updating weights expression can be obtained as

$$\begin{aligned}
\hat{\mathbf{W}}(n+1) &= \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\|\mathbf{U}(n)\|^2 + \delta} \mathbf{U}(n) \mathbf{e}^*(n) \\
&= \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\beta(n) + \delta} \left[\mathbf{d}^*(n) \mathbf{U}(n) - \hat{\mathbf{W}}^T(n) \mathbf{X}^*(n) \mathbf{U}(n) \right] \\
&= \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\beta(n) + \delta} \left[\mathbf{d}^*(n) \mathbf{U}(n) - \hat{\mathbf{W}}^T(n) \gamma(n) \mathbf{X}^*(n-1) \mathbf{U}(n) - \hat{\mathbf{W}}^T(n) \mathbf{U}^*(n) \mathbf{U}(n) \right] \\
&= \hat{\mathbf{W}}(n) - \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{U}^*(n) \mathbf{U}^T(n) \hat{\mathbf{W}}(n) - \frac{\tilde{\mu}}{\beta(n) + \delta} \gamma(n) \mathbf{U}(n) \mathbf{X}^H(n-1) \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{d}^*(n) \mathbf{U}(n) \\
&= \hat{\mathbf{W}}(n) - \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{U}(n) \mathbf{U}^H(n) \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{d}^*(n) \mathbf{U}(n) \\
&= \left[\mathbf{I} - \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{R}_U(n) \right] \hat{\mathbf{W}}(n) + \frac{\tilde{\mu}}{\beta(n) + \delta} \mathbf{P}(n) \tag{7}
\end{aligned}$$

where matrix $\mathbf{R}_U(n) = \mathbf{U}(n) \mathbf{U}^H(n)$, which is defined as the autocorrelation matrix of $\mathbf{U}(n)$. Vector $\mathbf{P}(n) = \mathbf{d}^*(n) \mathbf{U}(n)$, which is defined as the cross-correlation vector between $\mathbf{U}(n)$ and $\mathbf{d}(n)$. Considering $\beta(n) = \mathbf{U}^H(n) \mathbf{U}(n) = \text{tr} \{ \mathbf{R}_U(n) \}$, the normalized energy autocorrelation matrix of $\mathbf{R}_U(n)$ and the normalized energy cross-correlation vector of $\mathbf{P}(n)$ are defined as $\bar{\mathbf{R}}_U(n) = \mathbf{R}_U(n) \tilde{\mu} / [\beta(n) + \delta]$, $\bar{\mathbf{P}}(n) = \mathbf{P}(n) \tilde{\mu} / [\beta(n) + \delta]$, respectively. Assuming that $\mathbf{X}(n)$ and $\mathbf{d}(n)$ are both wide-sense stationary, we can know that $\mathbf{U}(n)$ is also wide-sense stationary. Thus, $\bar{\mathbf{R}}_U = E \{ \bar{\mathbf{R}}_U(n) \}$, $\bar{\mathbf{P}} = E \{ \bar{\mathbf{P}}(n) \}$.

We assume that $\bar{\mathbf{R}}_U = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T$ is the eigenvalue decomposition of $\bar{\mathbf{R}}_U$, which is a positive definite symmetry matrix. Where, $\mathbf{Q}^{-1} = \mathbf{Q}^T$, $\mathbf{\Lambda} = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_M \}$, $0 \leq \lambda_i \leq 1$, $1 \leq i \leq M$. By applying $\bar{\mathbf{R}}_U$ to Eq. (8), a derivation can be given as the following

$$\begin{aligned}
E \{ \hat{\mathbf{W}}(n+1) \} &= E \{ \mathbf{I} - \bar{\mathbf{R}}_U(n) \} E \{ \hat{\mathbf{W}}(n) \} + E \{ \bar{\mathbf{P}}(n) \} \\
&= [\mathbf{I} - \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T] E \{ \hat{\mathbf{W}}(n) \} + \bar{\mathbf{P}} \\
&= \mathbf{Q} (\mathbf{I} - \mathbf{\Lambda}) \mathbf{Q}^T E \{ \hat{\mathbf{W}}(n) \} + \bar{\mathbf{P}} \tag{8}
\end{aligned}$$

From Eq. (8), we can obtain

$$E \{ \hat{\mathbf{W}}(n) \} = \mathbf{Q} (\mathbf{I} - \mathbf{\Lambda})^n \mathbf{Q}^T E \{ \hat{\mathbf{W}}(0) \} + \bar{\mathbf{P}} \sum_{i=0}^{n-1} \mathbf{Q} (\mathbf{I} - \mathbf{\Lambda})^i \mathbf{Q}^T \tag{9}$$

where $\mathbf{I} - \mathbf{\Lambda} = \text{diag} (1 - \lambda_1, 1 - \lambda_2, \dots, 1 - \lambda_M)$. Since all of its diagonal components are no greater than 1, we can get

$$\begin{aligned}
\lim_{n \rightarrow \infty} (\mathbf{I} - \mathbf{\Lambda})^n &= \lim_{n \rightarrow \infty} \text{diag} [(1 - \lambda_1)^n, (1 - \lambda_2)^n, \dots, (1 - \lambda_M)^n] \\
&= \mathbf{0}_{M \times M} \tag{10}
\end{aligned}$$

So

$$\begin{aligned}
\lim_{n \rightarrow \infty} E \{ \hat{\mathbf{W}}(n) \} &= \mathbf{Q} \lim_{n \rightarrow \infty} (\mathbf{I} - \mathbf{\Lambda})^n \mathbf{Q}^T E \{ \hat{\mathbf{W}}(0) \} \\
&\quad + \bar{\mathbf{P}} \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \mathbf{Q} (\mathbf{I} - \mathbf{\Lambda})^i \mathbf{Q}^T = \bar{\mathbf{P}} \mathbf{Q} \sum_{i=0}^{\infty} (\mathbf{I} - \mathbf{\Lambda})^i \mathbf{Q}^T \\
&= \bar{\mathbf{P}} \mathbf{Q} \mathbf{\Lambda}^{-1} \mathbf{Q}^T = \bar{\mathbf{P}} (\bar{\mathbf{R}}_U)^{-1} = \hat{\mathbf{W}}_{\text{opt}} \tag{11}
\end{aligned}$$

Equation (11) indicates that the weight solution of the proposed algorithm is the optimal solution of Wiener equation. Adaptive constant $\tilde{\mu}$ can be concluded

$$\tilde{\mu}_{\text{opt}} = \frac{\text{Re} \{ E [\xi_U(n) \varepsilon^*(n) / \|\mathbf{U}(n)\|^2] \}}{E [|\varepsilon(n)|^2 / \|\mathbf{U}(n)\|^2]} \tag{12}$$

where $\xi_U(n) = [\mathbf{W} - \hat{\mathbf{W}}(n)]^H \mathbf{U}(n) = \mathbf{e}^H(n) \mathbf{U}(n)$, $\mathbf{e}(n) = \mathbf{W} - \hat{\mathbf{W}}(n)$, the vector \mathbf{W} denotes the weight vector of the reference channel. The deduction of $\tilde{\mu}$ is similar to the solution of NLMS's optimal step size parameter [10].

3 Performance comparisons of the convergence curves

3.1 Comparison between the new algorithm and VS-NLMS algorithm

The step size parameter μ of the variable step VS-NLMS algorithm in Ref. [7] is directly proportional to the estimation of the cross-function between $\mathbf{e}(n)$ and $\mathbf{X}(n)$. However, this algorithm has large computation complexity for the updated weight. The assumption of Ref. [7] is used in computer simulation:

1) The order of adaptive filter is set to $M = 10$.

2) The FIR weight of the unknown system is defined as $\mathbf{W} = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.5, 0.4, 0.3, 0.2]^T$.

3) The reference input signal $X(n)$ is zero-mean white Gaussian random with variance $\sigma^2 = 1$.

4) $V(n)$ is white Gaussian noise, which is uncorrelated with $X(n)$.

5) $\alpha = 1, \lambda = 0.997$, statistical mean times and samples are 50 and 3 000, respectively. The parameters of the proposed algorithm are set to $\tilde{\mu} = 0.5, \delta = 10^{-6}$, respectively. Figure 3 shows the comparison of the convergence curves between the new algorithm and VS-NLMS algorithm. And both of these algorithms have computation complexities $o(7M)$. From Fig. 3 we can obtain that the new algorithm has less stable misadjustment compared with VS-NLMS algorithm at the same convergence speed.

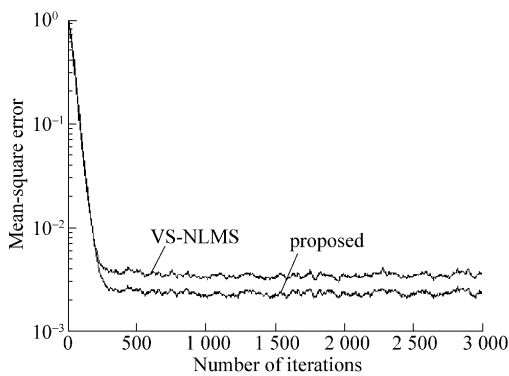


Fig. 3 Comparison of the convergence curves between the new algorithm and VS-NLMS algorithm

3.2 Comparison between the new algorithm and L.E-LMS algorithm

By optimizing gradient estimation, i.e., non-linear processing the error signal $e(n)$, the variable step L.E-LMS algorithm is presented in Ref. [8]. The assumption of this reference is used in the following computer simulation:

1) The order of adaptive filter is set to $M = 5$.

2) The FIR weight of the unknown system is defined as $W = [0.33, 0.67, 1, 0.67, 0.33]^T$.

3) The reference input signal $X(n)$ is zero-mean white Gaussian random with variance $\sigma^2 = 1$.

4) $V(n)$ is white Gaussian noise with variance equaling to 0.01, which is uncorrelated with $X(n)$.

5) $\mu_{\max} = 0.1, \mu_{\min} = 0.01$, statistical mean times and samples are 100 and 600, respectively. The parameters of the new algorithm are set to $\tilde{\mu} = 0.3, \delta = 0.1$, respectively. Figure 4 shows the comparison of the convergence curves between the new algorithm and L.E-LMS algorithm with computation complexities $o(4M)$. It is known in this figure that the proposed algorithm has a faster convergence speed than L.E-LMS algorithm at the same stable misadjustment.

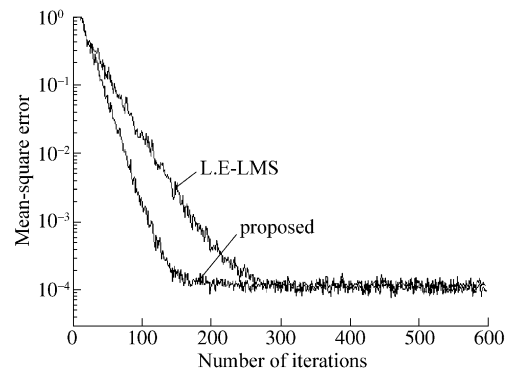


Fig. 4 Comparison of the convergence curves between the new algorithm and L.E-LMS algorithm

3.3 Comparison between the new algorithm and the algorithm in Ref. [9]

The variable step algorithm in Ref. [9] is presented by establishing another non-linear functional relationship $\mu(n) = \beta[1 - \exp(-\alpha |e(n)|^2)]$ between μ and $e(n)$. The assumption of Ref. [9] is adopted in the following computer simulation:

1) The order of adaptive filter is set to $M = 2$.

2) The FIR weight of the unknown system is defined as $W = [0.8, 0.5]^T$. And this unknown system changes at the sampling time of number 500 with the new weight $W = [0.4, 0.2]^T$.

3) The reference input signal $X(n)$ is zero-mean white Gaussian random with variance equal to 1.

4) $V(n)$ is white Gaussian noise with variance $\sigma^2 = 0.04$, which is uncorrelated with $X(n)$. Statistical mean times and samples are 200 and 1 000, respectively.

5) $\alpha = 300, \beta = 0.1$. The parameters of the new algorithm are set to $\tilde{\mu} = 0.7, \delta = 0.95$, respectively.

The comparison of the convergence curves between the new algorithm and the algorithm in Ref. [9] with computation complexities $o(3M)$ is given in Fig. 5. This figure indicates that the two convergence curves almost overlap

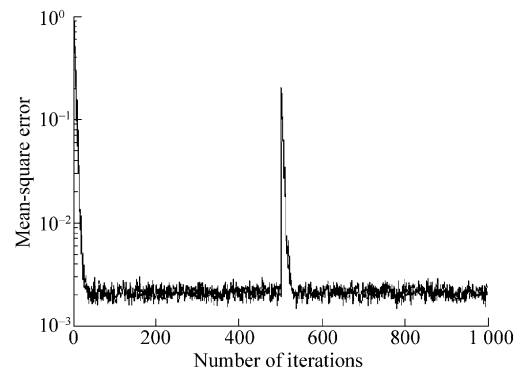


Fig. 5 Comparison of the convergence curves between the new algorithm and the algorithm in Ref. [9]

each other. So the conclusion is that these two algorithms almost have the same convergence speed at the comparative stable misadjustment.

4 Computer simulation results

In speech communication, the mismatching impedances of a hybrid transformer winding for connecting a two-wire circuit to a four-wire circuit give rise to an echo. To overcome the presence of echoes, the common practice is to use an adaptive echo canceller. The idea behind using an adaptive filter, the idea behind it is to produce a synthesized echo, which is similar to the desired signal i.e. the echo signal. Where the tap input vector and the error signal of the filter act on the adaptation and step-size controller to adjust the tap weights of the transversal filter so as to minimize the mean-square value of the error signal in accordance with a certain designated algorithm. To test the better performance of the proposed algorithm for processing high correlation signal, we apply it to adaptive echo cancellation to be compared with NLMS algorithm and DCR-LMS algorithm at the comparative computation complexities. The computation complexities of NLMS algorithm and DCR-LMS algorithm are $o(3M)$, $o(7M)$, respectively. In our test, the speech-sampling rate is 8 kHz. The order of echo cancellation is set to $M = 512$. The time delay of echo is assumed to be $\tau = 62.5$ ms. The adaptive step-size parameters of NLMS algorithm, DCR-LMS algorithm, and the proposed algorithm are set to $\tilde{\mu} = 1$, $\tilde{\mu} = 0.95$, and $\tilde{\mu} = 0.95$, respectively. The real constant δ of our algorithm is 0.1. Echo return loss enhancement (ERLE) is used to show performance comparisons, which is defined as [12]

$$\text{ERLE} = 10 \log_{10} \frac{E[y(n)^2]}{E[e(n)^2]} \quad (13)$$

where $E[\cdot]$ denotes mathematics expectations, which is substituted by averaging 1 200 samples at the sample instant n fore-and-aft in practical measure. $y(n)$ and $e(n)$ denote the clean speech echo and the remained speech echo of cancellation fore-and-aft at the sample instant n , respectively.

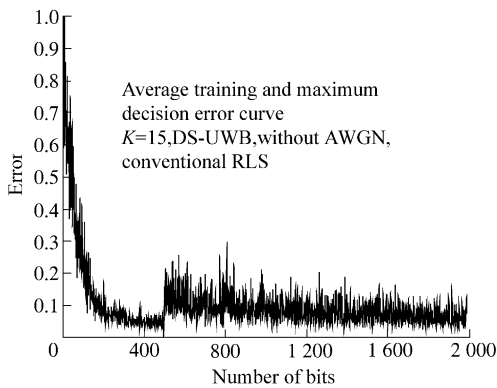


Fig. 6 Speech echo and its cancellation results by three algorithms. (a) Speech echo; (b) NLMS algorithm; (c) DCR-LMS algorithm; (d) The proposed algorithm

From Figs. 6 and 7, we can conclude that the stable misadjustment of the proposed algorithm is remarkably less than that of the NLMS algorithm. And its convergence speed is comparable to that of DCR-LMS algorithm but faster than that of NLMS algorithm. Its stable misadjustment is remarkably less than that of DCR-LMS algorithm when iterative times run more than 10^4 . Figure 8 shows that the proposed algorithm performs best compared with the NLMS algorithm and the DCR-LMS algorithm, especially for the highest energy of echo. It can be explained that the correlation of voiced enhances when the energy of speech echo increases. By the decorrelation processing, the new algorithm almost keeps stable performance for the increased correlation of the input signals. In other words, its

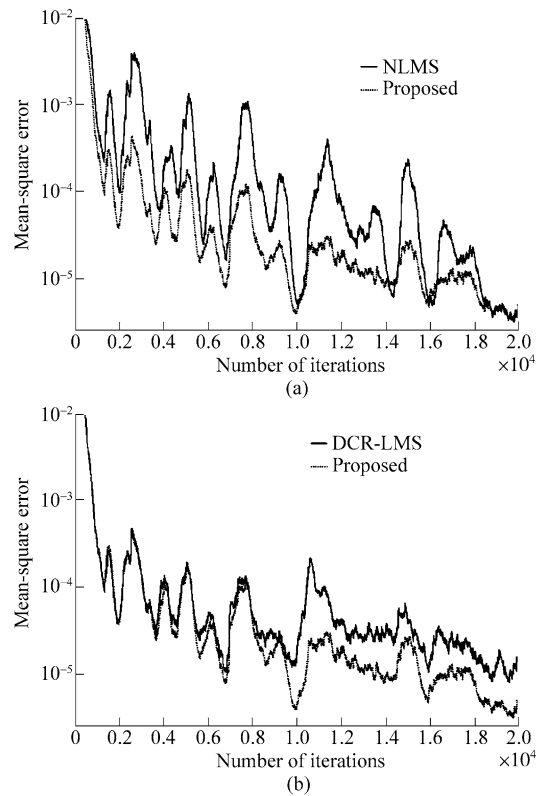


Fig. 7 Convergence curves of the algorithms

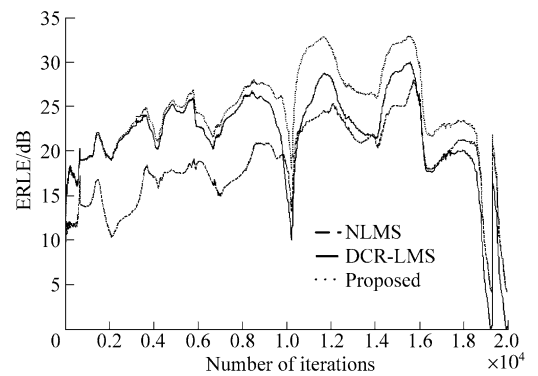


Fig. 8 ERLE comparison of the algorithms

good performance of suppressing echo signal can be obtained.

5 Conclusions

By defining the correlation coefficient between the signal's former and latter sampling instants and establishing the new input vector of the filter, a new adaptive filter algorithm has been proposed in this paper. It is of simple structure, fast convergence speed and comparable computation complexity with the NLMS algorithm. It can process both uncorrelated signal and high correlation signal. For processing uncorrelated signal, the presented algorithm performs best compared with VS-LMS algorithm and L.E-LMS algorithm. Its convergence speed is comparable to that of the algorithm in Ref. [9]. For processing high correlation signal, the new algorithm is used in adaptive echo cancellation. It has least stable misadjustment compared with the NLMS algorithm and DCR-LMS algorithm. Research results indicate that the proposed algorithm can be widely used in engineering application for its advantages.

References

1. Bitmead R., Anderson B., Performance of adaptive estimation algorithms in dependent random environments, *IEEE Trans. Automatic Control*, 1980(b), AC-25(4): 788–794
2. Ozeki K., Umeda T., An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties, *Electronics and Communications in Japan* 1984, 67-A(5): 19–27
3. Steven L., Sanjeev T., The fast affine projection algorithm, *Proc ICASSP'95*, 1995: 3023–3026
4. Benesty J., Duhamel D., Grenier Y., A multi-channel affine projection algorithm with applications to multi-channel acoustic echo cancellation, *IEEE Signal Processing Letters*, 1996, 3(2): 35–37
5. Makino S., Strauss K., Shimauchi S. et al., Subband stereo echo canceller using the projection algorithm with fast convergence to the true echo path, *ICASSP'1997*, 1997: 299–302
6. Gitlin R., Weinstein S., The effects of large interference on the tracking capability of digitally implemented echo cancellers. *IEEE Trans on COM*, 1978(6): 833–839
7. Ye Hua, Wu Bo-xiu, Study of variable step size LMS adaptive algorithm, *Acta Electronica Sinica*, 1990, 18(4): 63–69 (in Chinese)
8. Wu Guang-bi, Zhu Lin-yu, A LMS adaptive filtering algorithm with variable step algorithm, *Acta Electronica Sinica*, 1994, 22(1): 55–60 (in Chinese)
9. Gao Ying, Xie Sheng-li, A variable step size LMS adaptive filtering algorithm and its analysis, *Acta Electronica Sinica*, 2001, 29(8): 1094–1097 (in Chinese)
10. Simon H.-K., *Adaptive filter theory*, Fourth Edition, Prentice Hall, 2001
11. Yasukawa H., Shimada S., An acoustic echo canceller using subband sampling and decorrelation methods, *IEEE Trans Signal Processing*, 1993, 41(2): 926–930
12. Doherty J., Porayath R, A robust echo canceller for acoustic environments, *IEEE Trans CAS: Analog and Digital Signal Processing*, 1997, 44(5): 389–396