A MODIFIED NLMS ALGORITHM FOR ADAPTIVE NOISE CANCELLATION

Lin Bai, Qinye Yin

Ministry of Education Key Lab for Intelligent Networks and Network Security,
Xi'an Jiaotong University, Xi'an, China 710049 Email: ber_ling@msn.com

ABSTRACT

Adaptive noise cancellation (ANC) is used widely to reduce noise from a noisy speech sound. However the least-mean-square (LMS) algorithm and its variants, such as the normalized (N)LMS, the modified (M)-LMS and the constrained stability (CS)-LMS algorithms do not perform well in ANC since the desired speech signal has a bad effect on the convergence rate and steady state misadjustment of these algorithms. Thus, we propose a new adaptive algorithm that further relaxes the constraint in the CS-LMS algorithm. The new algorithm attempts to minimize the estimation error of the a posteriori error and the estimation is obtained using the concept of Taylor's expansion. The analysis and simulation results show that the proposed new algorithm outperforms the NLMS and CS-LMS algorithms.

***Index Terms*—** Adaptive noise cancellation, NLMS algorithm, Taylor's expansion

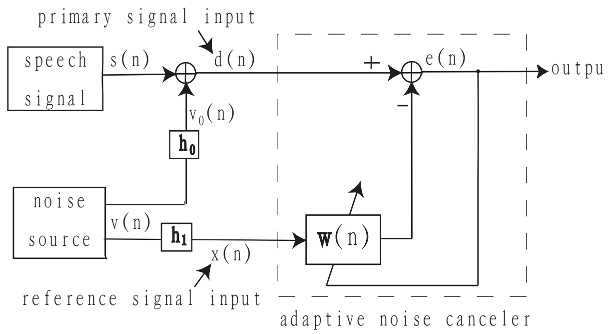

Fig. 1. Adaptive noise canceler.

1. INTRODUCTION

The LMS algorithm is popular due to its low computational cost and ease of implementation and it is widely used in many applications, including equalization, channel estimation, speech recognition and echo cancellation [1]-[3]. The typical adaptive noise canceler for filtering speech sounds has two input signals as shown in Fig. 1. The primary signal input $d(n)$ is the desired speech signal $s(n)$ corrupted by the noise $v_0(n)$, and the reference input signal $x(n)$ is correlated with $v_0(n)$. Due to this correlation, the adaptive canceler can reduce the noise from the primary signal and obtain the output error signal $e(n)$ equal to the speech signal $s(n)$. To validate the canceler, the LMS algorithm with the following equations can be used to adjust the weight vector.

$$e(n) = d(n) - \boldsymbol{w}^H(n)\boldsymbol{x}(n) \tag{1}$$

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu e^*(n)\boldsymbol{x}(n) \tag{2}$$

where μ is the step size for the weight vector adjustment and $\boldsymbol{x}(n) = [x(n), x(n-1), \cdots, x(n-N+1)]^T$ is a collum vector. N is the length of the adaptive filter. The subscript H in (1) denotes the complex conjugate transpose and $*$ in (2) gets the complex conjugate of its argument. Although the LMS algorithm has many good properties, it does not perform well with strong target signal $s(n)$ in ANC.

Several variable step size LMS algorithms have been proposed in these years to improve the performance for LMS. The NLMS algorithm was proposed by Nagumo and Noda [4], and Albert and Gardner [5] independently. Later, a simplified version of the least-perturbation property for the NLMS algorithm was studied by Goodwin and Sin [6]. Kwong and Johnston proposed the variable-step-size (VSS)-LMS algorithm [7] with the current step size depending partly on the previous one. The normalized data nonlinearity (NDN)-LMS algorithm [8] proposed by Douglas and Meng applies nonlinearity to the input data vector. All these algorithms have better performance than LMS on both the convergence speed and steady state misadjustment in many cases. However, they still can not achieve good performance in the ANC with strong target signal $s(n)$. The M-LMS and CS-LMS algorithms proposed in [9] and [10] separately are more concentrated on the ANC problem. CS-LMS does not need a priori information that M-LMS needs and it can achieve better performance than NLMS by relaxing the restrictive constraint in NLMS. The CS-LMS algorithm enforces the change of the a posteriori error to be as small as possible instead of enforcing the a posteriori error itself to vanish. However, the constraint in CS-LMS is also considered to be restrictve as the speech signal is characterized by its changes. Thus, we propose a new adaptive algorithm in this paper that further relaxes the constraint of CS-LMS and performance improvement for CS-LMS can be achieved.

In Section II, the new adaptive algorithm using the concept of Taylor's expansion is proposed. Thereafter, the performance analysis of this new algorithm is detailed in Section III. Section IV gives some simulation results comparing the new algorithm with the NLMS and CS-LMS algorithms and Section V draws the conclusions.

2. NEW ADAPTIVE ALGORITHM

Let us first define the a posteriori error

$$r_n(i) \triangleq d(i) - \boldsymbol{w}^H(n+1)\boldsymbol{x}(i),\ i \leqslant n \tag{3}$$

and the a priori error

$$e_n(i) \triangleq d(i) - \boldsymbol{w}^H(n)\boldsymbol{x}(i),\ i \leqslant n \tag{4}$$

From (4), we known $e_n(n) = e(n)$. As illustrated in [1], the NLMS algorithm with weight vector update formula

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \frac{\mu e^*(n)\boldsymbol{x}(n)}{\| \boldsymbol{x}(n) \|^2} \tag{5}$$

can also be regarded as the exact solution to a local optimization problem of determining $\boldsymbol{w}(n+1)$ by solving the constrained optimization problem:

$$\min_{\boldsymbol{w}(n+1)} \| \boldsymbol{w}(n+1) - \boldsymbol{w}(n) \|^2 \text{ subject to } r_n(n) = (1-\mu)e_n(n) \tag{6}$$

As the constraint in NLMS is too restrictive, the CS-LMS algorithm relaxes it and arrives at another constrained optimization problem:

$$\min_{\boldsymbol{w}(n+1)} \| \boldsymbol{w}(n+1) - \boldsymbol{w}(n) \|^2 \text{ subject to } \delta r_n(n) = (1-\mu)\delta e_n(n) \tag{7}$$

where $\delta[\cdot](n) = [\cdot](n) - [\cdot](n-1)$. The solution to this optimization problem is

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \frac{\mu \delta e^*(n)\delta\boldsymbol{x}(n)}{\| \delta\boldsymbol{x}(n) \|^2} \tag{8}$$

Comparing (7) with (6), we see the CS-LMS algorithm forces the sequence of $\{r_n(i)\}$ to be more smooth than the sequence of $\{e_n(i)\}$ instead of forcing $r_n(n)$ itself to be small and it can achieve better performance than NLMS by using this less restrictive constraint. However, the constraint in CS-LMS is still considered to be over restrictive as it attempts to force the a posteriori error to be unchanged, which can not be achieved due to the changes of the speech signal $s(n)$.

Thus, we further relax the constraint in CS-LMS and receive the new constrained optimization problem as follows

$$\min_{\boldsymbol{w}(n+1)} \| \boldsymbol{w}(n+1) - \boldsymbol{w}(n) \|^2 \text{ subject to } \tilde{r}_n(n) = (1-\mu)\tilde{e}_n(n) \tag{9}$$

where $\tilde{r}_n(n)$ and $\tilde{e}_n(n)$ are differences given by

$$\tilde{r}_n(n) = r_n(n-k+1) - \hat{r}_n(n-k+1) \tag{10}$$

$$\tilde{e}_n(n) = e_n(n-k+1) - \hat{e}_n(n-k+1) \tag{11}$$

$\hat{r}_n(n-k+1)$ and $\hat{e}_n(n-k+1)$ are the estimation of $r_n(n-k+1)$ and $e_n(n-k+1)$ respectively and $\tilde{r}_n(n)$ and $\tilde{e}_n(n)$ denotes the corresponding estimation errors for the a posteriori and a priori errors. Thus the new algorithm attempts to make the estimation error of the a posteriori error be as small as possible. This makes the new algorithm allow the a posteriori error to change, which is more reasonable than previous algorithms as speech signals are characterized by their changes.

In order to achieve accurate estimates without priori information, we use the sum of the lower order derivative terms of the Taylor's expansion for our estimation. $k \geqslant 1$ in (10) and (11) represents the highest order of derivatives we use. Considering the form of the Taylor's expansion, we derive

$$\hat{e}_n(n-k+1) = e_n(n-k) + \sum_{m=1}^{k} \frac{\hat{\delta}^{(m)} e_n(n-k)}{m!} \tag{12}$$

where $\hat{\delta}^{(m)} e_n(n-k), m = 1, \cdots, k$ can be seen as the the derivatives of $e_n(i)$ at instant $n-k$. As $e_n(i)$ is discrete and its derivatives do not exist, $\hat{\delta}^{(m)} e_n(n-k), m = 1, \cdots, k$ in (12) are actually approximated by the differences as follows

$$\hat{\delta}^{(1)} e_n(n-i) = \frac{e_n(n-i+1) - e_n(n-i-1)}{2}, \quad i = 1, 2, \cdots, 2k-1 \tag{13}$$

$$\hat{\delta}^{(m+1)} e_n(n-i) = \frac{\hat{\delta}^{(m)} e_n(n-i+1) - \hat{\delta}^{(m)} e_n(n-i-1)}{2}, \quad i = m+1, m+2, \cdots, 2k-m-1 \tag{14}$$

From (13) and (14), we can see that $2k+1$ samples: $e_n(i), i = n-2k, n-2k+1, \cdots, n$ are used to obtain the first to k-th order derivatives. Considering (11) to (14) and (4), we can arrive at

$$\tilde{e}_n(n) = \tilde{d}(n) - \boldsymbol{w}^H(n)\tilde{\boldsymbol{x}}(n) \tag{15}$$

with

$$\tilde{\boldsymbol{x}}(n) = \boldsymbol{x}(n-k+1) - \left[\boldsymbol{x}(n-k) + \sum_{m=1}^{k} \frac{\hat{\delta}^{(m)}\boldsymbol{x}(n-k)}{m!}\right] \tag{16}$$

Replacing $\boldsymbol{x}$ by d in the above equation, we can get $\tilde{d}(n)$. The derivative estimates $\hat{\delta}^{(m)}\boldsymbol{x}(n-k), m = 1, \cdots, k$ in (16) are also obtained using the method shown in (13) and (14). As $\hat{r}_n(n-k+1)$ is obtained by substituting r for e in (12) to (14), we can also derive

$$\tilde{r}_n(n) = \tilde{d}(n) - \boldsymbol{w}^H(n+1)\tilde{\boldsymbol{x}}(n) \tag{17}$$

Substituting (15) and (17) into the constraint in (9) and using the Lagrangian method, we can solve this optimization problem and the received solution is

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \frac{\mu \tilde{e}_n^*(n)\tilde{\boldsymbol{x}}(n)}{\| \tilde{\boldsymbol{x}}(n) \|^2} \tag{18}$$

Similar to the NLMS and CS-LMS algorithm, we can add a small positive value ε to the denominator in (18) to make our new algorithm more robust. In the new adaptive algorithm, properly increasing the order k can make the Taylor's expansion based estimation become more accurate and performance improvement can be achieved. However, k should not be too large as the estimation of the high order derivatives become inaccurate with the method shown in (13) and (14). Until now, we consider $k \geqslant 1$. If $k = 0$, which means only zero order derivative term is used, $\tilde{e}_n(n)$ and $\tilde{\boldsymbol{x}}(n)$ will have the form: $\tilde{e}_n(n) = e_n(n) - e_n(n-1)$ and $\tilde{\boldsymbol{x}}(n) = \boldsymbol{x}(n) - \boldsymbol{x}(n-1)$. In this case, the new algorithm shown in (18) is the same as the CS-LMS algorithm, which means the CS-LMS algorithm can be seen as a special case of our new algorithm with $k = 0$.

3. PERFORMANCE ANALYSIS

In this section, we analyze the performance of the new adaptive algorithm. The constraint on the step size μ is determined to make the weight vector converge to the optimal solution (Wiener filter) $\boldsymbol{w}_o$. The excess mean square error (EMSE) of the new algorithm is also analyzed. The performance of the CS-LMS algorithm is the same as that of the new algorithm by setting $k = 0$.

From Fig. 1, we can see the primary signal

$$d(n) = s(n) + v_0(n) \tag{19}$$

where the noise $v_0(n)$ that corrupts the desired speech signal $s(n)$ can be assumed to have the form

$$v_0(n) = \boldsymbol{w}_o^H \boldsymbol{x}(n) + v_e(n) \tag{20}$$

where $v_e(n)$ is the small component of $v_0(n)$ that can not be canceled by the optimal Wiener filter and it is assumed to be independent of $\boldsymbol{x}(n)$ and $s(n)$. Considering (19) and (20), we can arrive at

$$e_n(n) = s(n) + v_e(n) + \tilde{\boldsymbol{w}}^H(n)\boldsymbol{x}(n) \tag{21}$$

$$\tilde{e}_n(n) = \tilde{s}(n) + \tilde{v}_e(n) + \tilde{\boldsymbol{w}}^H(n)\tilde{\boldsymbol{x}}(n) \tag{22}$$

where

$$\tilde{\boldsymbol{w}}(n) = \boldsymbol{w}_o - \boldsymbol{w}(n) \tag{23}$$

and $\tilde{s}(n)$ and $\tilde{v}_e(n)$ are obtained using the same form of equation as is shown in (16) for $\tilde{\boldsymbol{x}}(n)$.

From (21) and (22), and (5) and (18), we can see the new algorithm and NLMS have the following corresponding relationship.

$$e(n) \leftrightarrow \tilde{e}_n(n) \quad \boldsymbol{x}(n) \leftrightarrow \tilde{\boldsymbol{x}}(n) \quad s(n) \leftrightarrow \tilde{s}(n) \quad v_e(n) \leftrightarrow \tilde{v}_e(n) \tag{24}$$

Replacing the variables of NLMS on the left handside of the arrows by the variables of the new algorithm on the right handside of the arrows, we can analyze the performance of the new algorithm with the performance analysis method for NLMS [1]. The received results are as follows.

1) $\mu < 2$ is a sufficient condition to make the new algorithm be convergent in the mean and be mean square stable.

2)

$$\lim_{n\to\infty} \mathrm{E}\left\{\tilde{\boldsymbol{w}}^H(n)\tilde{\boldsymbol{x}}(n)\tilde{\boldsymbol{x}}^H(n)\tilde{\boldsymbol{w}}(n)\right\} = \frac{\mu\sigma_{new}^2}{2-\mu} \tag{25}$$

with

$$\sigma_{new}^2 = \mathrm{E}\left\{|\tilde{s}(n)|^2\right\} + \mathrm{E}\left\{|\tilde{v}_e(n)|^2\right\} \tag{26}$$

The EMSE of the new algorithm can be achieved from (25). To achieve that EMSE, we first derive an expanded form of $\tilde{\boldsymbol{x}}(n)$. After some algebra, the following equation can be obtained from (16).

$$\tilde{\boldsymbol{x}}(n) = \boldsymbol{x}(n-k+1) - \sum_{l_1=-k}^{k} \alpha_{l_1}\boldsymbol{x}(n-k+l_1) \tag{27}$$

with

$$\alpha_{l_1} = \begin{cases} \displaystyle\sum_{l_2=\lfloor l_1/2\rfloor}^{\lfloor k/2\rfloor} \frac{(-1)^{l_2-l_1/2}C_{2l_2}^{l_2-l_1/2}}{2^{2l_2}(2l_2)!}, & l_1 \text{ is even} \\ \displaystyle\sum_{l_2=b}^{\lfloor\frac{k-1}{2}\rfloor} \frac{(-1)^{l_2-\frac{l_1-1}{2}}C_{2l_2+1}^{l_2-\frac{l_1-1}{2}}}{2^{2l_2+1}(2l_2+1)!}, & l_1 \text{ is odd} \end{cases} \tag{28}$$

where $\lfloor\nu\rfloor$ rounds ν to the nearest integer less than ν and $C_m^l = m!/[l!(m-l)!]$. b in (28) is

$$b = \begin{cases} \dfrac{l_1-1}{2}, & l_1 \geqslant 1 \\ \dfrac{-l_1-1}{2}, & l_1 < 1 \end{cases} \tag{29}$$

If we let $\tilde{\alpha}_1 = 1-\alpha_1$ and $\tilde{\alpha}_{l_1} = -\alpha_{l_1}$ for $l_1 \neq 1$, we can further simplify the form of (27).

$$\tilde{\boldsymbol{x}}(n) = \sum_{l_1=-k}^{k} \tilde{\alpha}_{l_1}\boldsymbol{x}(n-k+l_1) \tag{30}$$

The EMSE ζ^{new} of the new algorithm can be derived as

$$\begin{aligned} \zeta^{new} &= \lim_{n\to\infty} \mathrm{E}\left\{\tilde{\boldsymbol{w}}^H(n)\boldsymbol{x}(n)\boldsymbol{x}^H(n)\tilde{\boldsymbol{w}}(n)\right\} \\ &\simeq \lim_{n\to\infty} \mathrm{E}\left\{\tilde{\boldsymbol{w}}^H(n)\boldsymbol{R}_x\boldsymbol{w}(n)\right\} \end{aligned} \tag{31}$$

where the correlation matrix $\boldsymbol{R}_x = \mathrm{E}\{\boldsymbol{x}(n)\boldsymbol{x}^H(n)\}$. Considering the widely used independence assumptions that assume the sequence $\{\boldsymbol{x}(n)\}$ to be independent and identically distributed (i.i.d.), we can get the following equation from (25) and (30).

$$\begin{aligned} \frac{\mu\sigma_{new}^2}{2-\mu} &\simeq \lim_{n\to\infty} \mathrm{E}\left\{\tilde{\boldsymbol{w}}^H(n)\left[\sum_{l_1=-k}^{k}\tilde{\alpha}_{l_1}^2\right]\boldsymbol{R}_x\tilde{\boldsymbol{w}}(n)\right\} \\ &= \left[\sum_{l_1=-k}^{k}\tilde{\alpha}_{l_1}^2\right]\lim_{n\to\infty} \mathrm{E}\left\{\tilde{\boldsymbol{w}}^H(n)\boldsymbol{R}_x\tilde{\boldsymbol{w}}(n)\right\} \end{aligned} \tag{32}$$

Considering both (31) and (32), the EMSE of the new algorithm can be obtained as follows

$$\zeta^{new} = \frac{\mu\sigma_{new}^2}{(2-\mu)\sum_{l_1=-k}^{k}\tilde{\alpha}_{l_1}^2} = \frac{\mu\left[\mathrm{E}\left\{|\tilde{s}(n)|^2\right\} + \mathrm{E}\left\{|\tilde{v}_e(n)|^2\right\}\right]}{(2-\mu)\sum_{l_1=-k}^{k}\tilde{\alpha}_{l_1}^2} \tag{33}$$

As given in [1], the EMSE of the NLMS algorithm is

$$\zeta^{NLMS} = \frac{\mu\left[\mathrm{E}\{|s(n)|^2\} + \mathrm{E}\{|v_e(n)|^2\}\right]}{2-\mu} \tag{34}$$

Comparing (33) with (34), since $v_e(n)$ and $\tilde{v}_e(n)$ are often very small, the EMSE of the new algorithm and the EMSE of the NLMS algorithm are dominated by $\mathrm{E}\{|\tilde{s}(n)|^2\}$ and $\mathrm{E}\{|s(n)|^2\}$ respectively. With the strongly correlated desired speech signal $s(n)$, the power of $\tilde{s}(n)$ is much smaller than that of $s(n)$. Moreover, it can be proved that $\sum_{l_1=-k}^{k}\tilde{\alpha}_{l_1}^2 > 1$ in (33). Thus, the new algorithm can achieve much smaller EMSE than the NLMS algorithm.

4. SIMULATION RESULTS

Simulations are implemented to compare the performance of the new algorithm with the performance of NLMS and CS-LMS. In the simulations, we assume that the noise source $v(n)$ is a white Gaussian noise with variance 1. The filters $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are modeled as

$$H_0(z) = 0.8 - 0.3z^{-1} + 0.1z^{-2}, \quad H_1(z) = 1 \tag{35}$$

The length of the adaptive filter is 3 (i.e. $N=3$). The small positive value ε is set to be 10^{-4} for all the concerned algorithms. We resort to the following mean square error MSED [1] with respect to the desired signal $s(n)$ for the performance analysis. MSED can represent the EMSE performance.

$$\mathrm{MSED}(n) = \frac{1}{J}\sum_{m=0}^{J-1}|e(n-m) - s(n-m)|^2 \tag{36}$$

4.1. AR Model Signal

In this simulation, the desired signal $s(n)$ is a sum of a AR(2) model process with the form: $s(n) = 1.95s(n-1) - 0.99s(n-2) + \gamma(n)$ and a zero mean white Gaussian noise with variance 10^{-4}. $\gamma(n)$ is a white Gaussian noise with variance 0.001. The step size μ is

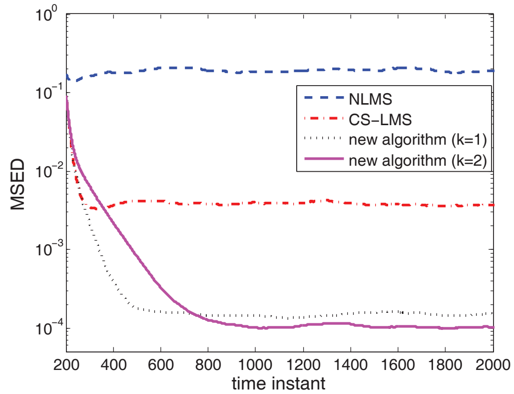

Fig. 2. MSED comparison of NLMS, CS-LMS and the new algorithm with $k = 1, 2$ in a stationary environment. ($\mu = 0.1$)

set to be 0.1. We average the results of 500 realizations. A new desired signal $s(n)$ and a new noise source $v(n)$ are generated in each realization. J in (36) is selected to be 200. As depicted in Fig. 2, the new algorithm with $k = 1, 2$ can clearly reduce the MSED for NLMS and CS-LMS and increasing the order k from 1 to 2 can make the new algorithm achieve even smaller MSED.

4.2. Speech Signal

The desired signal $s(n)$ is a real speech signal as shown in Fig. 3. The step size μ is set to be 1 and $J = 200$. Also 500 realizations are implemented for average. In this simulation, only $v(n)$ changes in different realizations. Thus, $s(n)$ in this simulation is a nonstationary speech signal. Fig. 4 shows that the MSED of the NLMS algorithm is large. The CS-LMS algorithm (i.e. the new algorithm with $k = 0$) has a smaller MSED. However, its MSED still can not be very small when the amplitude of the speech signal is large. Properly increasing the order k of the new algorithm can obtain even smaller MSED, which means properly higher order k makes the output $e(n)$ of the new algorithm track the speech signal better.

5. CONCLUSION

A new adaptive algorithm for ANC to filter a noisy speech signal is proposed in this paper. The new algorithm further relaxes the constraint in the CS-LMS algorithm. Instead of forcing the a posteriori error to be as smooth as possible, the new algorithm attempts to minimize the estimation error of the a posteriori error and the estimation is obtained using the concept of Taylor's expansion. CS-LMS can be seen as a special case of the new algorithm with $k = 0$. Analysis and simulation results show that the new algorithm can obtain better EMSE performance than NLMS and the new algorithm can get performance improvement by properly increasing the order k.

6. REFERENCES

[1] Ali H. Sayed, *Adaptive Filters*, New York: John Wiley & sons, 2008.

[2] B.Widrow, J. R. Glover, J. M. Mccool, J. Kaunitz, C. S. Williams, R. H. Hean, J. R. Zeidler, E. Dong and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, no. 12, pp.1692-1716, Dec. 1975.

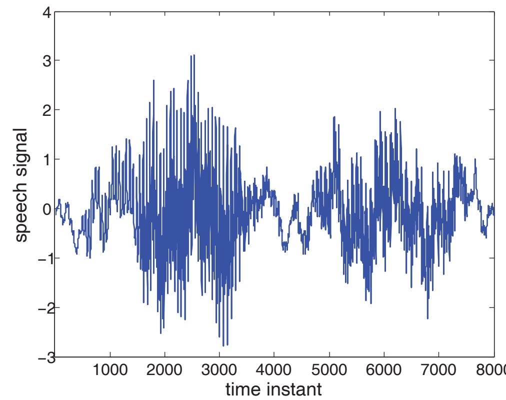

Fig. 3. the desired speech signal waveform

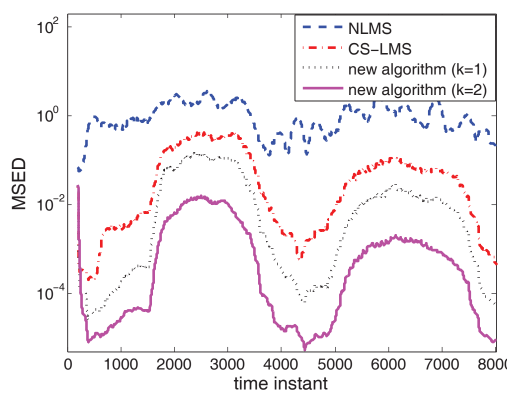

Fig. 4. MSED comparison of the three concerned algorithms with the nonstationary speech signal ($\mu = 1$)

[3] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1986.

[4] J.I. Nagumo and A. Noda, "A learning method for system identification," *IEEE Trans. Signal Process*, vol. 41, no. 6, pp. 2075-2087, June 1993.

[5] A.E. Albert and L.S. Gardner, *Stochastic Approximation and Nonlinear Regression*, Cambridge, MA: MIT Press, 1967.

[6] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[7] R.H. Kwong and E.W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1633-1642, July 1992.

[8] S. C. Douglas and T. H. Y. Meng, "Normalized data nonlinearities for LMS adaptation," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1352-1354, Jun. 1994.

[9] J.E. Greenberge, "Modified LMS algorithms for speech processing with an adaptive noise canceler," *IEEE Trans. Speech Audio Process*, vol. 6, no. 4, pp. 338-351, Jul. 1998.

[10] J. M. Gorriz, Javier Ramirez, S. Cruces-Alvarez, Carlos G. Puntonent, Elmar W. Lang and Deniz Erdogmus, "A Novel LMS algorithm Applied to Adaptive Noise Cancellation," *IEEE Signal Processing Letters*, vol. 6, no. 1, pp. 34-37, Jan. 2009.