Modified LMS and NLMS Algorithms with a New Variable Step Size

Zhao Shengkui, Man Zhihong, Khoo Suiyang
School of Computer Engineering
Nanyang Technological University
Nanyang Avenue, Singapore 639798
zhao0024@ntu.edu.sg, aszhman@ntu.edu.sg, khoo0032@ntu.edu.sg

***Abstract*—In this paper, the modified LMS and NLMS algorithms with variable step-size are presented. It is shown that the variable step size is computed using a ratio of the sums of weighted energy of the output error with two exponential factors α and β, thus the fast error convergence of the modified LMS and NLMS algorithms can then be achieved. Also, by properly choosing the values of α and β, the misadjustment can be further improved. A few simulation results are presented in support of the good performance of the proposed algorithms by comparing with other LMS-type algorithms.**

***Keywords*—LMS algorithm, NLMS algorithm, variable step-size**

I. Introduction

The standard least mean square (LMS) algorithm in [1] is one of the mostly popular adaptive algorithms, which has been widely used for system identification [2, 3], channel equalization [4], linear predication [5, 6], and noise cancellation [7, 8], due to its simplicity and robustness. Recently, many researches have been carried out to improve the tradeoff between the convergence rate and misadjustment. It is known that a large step size may lead to fast convergence but big misadjustment, and a small step-size may provide small misadjustment but slow convergence [1]. To achieve both fast convergence and small misadjustment, many algorithms with variable step-size have been proposed in the literature [9]-[13]. In [9], the author proposed a variable step-size LMS algorithm which adjusts the step-size based on the ratio of absolute value of the output error and absolute value of the desired output. In [10], a variable step-size LMS algorithm using error-data normalization is proposed. Another approach to adjust the time-varying step-size based on the square of the time-averaged estimate of autocorrelation of $e(n)$ and $e(n-1)$ is proposed in [11]. As seen in [9]-[13], a large step-size is usually used at the beginning of the adaptation process to obtain a fast convergence, and a smaller step-size is then used to reduce the level of residual error in steady state.

The normalized LMS (NLMS) algorithm is proven to provide faster convergence than the LMS algorithm [11]. In the NLMS algorithm, the fastest convergence is usually achieved by setting the step size equal to one. Therefore, when the noise level is high at the beginning of the process, a step-size equal or close to one appears to be the best choice. While in the steady state, the step-size of NLMS algorithm is expected to be small value to obtain smaller misadjustment. In [13], a new criterion using the square of the output error is proposed, where the step-size value is chosen from a big value α_2 and a smaller one α_1 to yield both faster tracking speed and smaller misadjustment. Also, a new type of algorithms based on error normalization is provided in [14] and [15].

In this paper, we further investigate LMS (VS-LMS) algorithm and NLMS (VS-NLMS) algorithm with a proposed variable step-size. It will be shown that the new variable step-size is updating by a ratio of the sums of the weighted energy of the output error. By properly choosing the exponential factors of α and β, large step sizes are obtained to provide fast convergence at the beginning of the adaptation process, and then the step-size gradually decreases to ensure small misadjustment in the steady state.

II. The Proposed VS-LMS and VS-NLMS Algorithms

The weight update equation of the standard LMS algorithm is given by [1]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n), \tag{1}$$

where $\mathbf{w}(n)$ is the weight vector of the adaptive filter, $e(n)$ is the output error between the reference signal and the filter output , $\mathbf{x}(n)$ is the input vector. The parameter μ is a fixed step-size. In this paper, the tap inputs, tap coefficients and reference signals are assume to be real.

It has been shown in [17] that a sufficient condition for mean-square error (MSE) convergence of the LMS algorithm is given by

$$0 < \mu \leq \frac{2}{3Tr(\mathbf{R})}, \tag{2}$$

where $Tr(\mathbf{R})$ denotes the trace of the autocorrelation matrix $\mathbf{R}$ of the inputs.

In [11], the authors proposed a variable step size for the modified weight update equation

1-4244-0342-1/06/$20.00 ©2006 IEEE

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu_n e(n)\mathbf{x}(n). \tag{3}$$

where μ_n is a variable step-size, which is adjusted based on the square of the time-averaged estimate of autocorrelation of $e(n)$ and $e(n-1)$ described by

$$p(n) = \beta p(n-1) + (1-\beta)e(n)e(n-1), \tag{4}$$

and

$$\mu_{n+1} = \alpha\mu_n + \gamma p(n)^2, \tag{5}$$

where $0 < \beta < 1$, $0 < \alpha < 1$, $0 < \gamma$, and $\mu_n \in [\mu_{\min}, \mu_{\max}]$. To guarantee the stability of the MSE, the following condition is derived by the authors in [11]:

$$0 < \frac{\gamma\xi_{\min}^2(1-\beta)}{1-\alpha^2} \le \frac{1}{3Tr(R)}. \tag{6}$$

In this paper, the proposed variable step-size is given by

$$\mu_n = \frac{\sum_{i=0}^{n} \alpha^i e^2(n-i)}{\sum_{j=0}^{n} \beta^j e^2(n-j)}, \tag{7}$$

where $0 < \alpha < \beta < 1$, $\mu_n \in [\mu_{\min}, \mu_{\max}]$, and $e^2(n)$ is the energy of the instantaneous output error.

It is obvious that the following inequality is true.

$$\sum_{i=0}^{n} \alpha^i e^2(n-i) \le \sum_{j=0}^{n} \beta^j e^2(n-j). \tag{8}$$

Remark 2.1: Intuitively speaking, the variable step-size in (7) provides large values at the initial adaptation process to obtain fast convergence speed. Then it decreases gradually to ensure low level of residual errors in steady state. In addition, the variable step-size can also increase to large values when abrupt change occurs in the system.

Remark 2.2: The sum of weighted energy of the output error $\varepsilon_i(n) = \sum_{i=0}^{n} \alpha^i e^2(n-i)$ can be reformulated as

$$\begin{aligned}\varepsilon_i(n) &= \sum_{i=0}^{n} \alpha^{n-i} e^2(i) = \alpha \sum_{i=0}^{n-1} \alpha^{n-1-i} e^2(i) + e^2(n) \\ &= \alpha\varepsilon_i(n-1) + e^2(n),\end{aligned} \tag{9}$$

Notice that this is a recursive form for updating the sum of the weighted energy of the output error.

Remark 2.3: In a similar way, the recursive form of the term $\varepsilon_j(n) = \sum_{j=1}^{n} \beta^{n-j} e^2(j)$ is then given by

$$\varepsilon_j(n) = \beta\varepsilon_j(n-1) + e^2(n). \tag{10}$$

Therefore, the proposed variable step-size in (7) can be simplified as

$$\mu_n = \frac{\varepsilon_i(n)}{\varepsilon_j(n)}, \tag{11}$$

where $\varepsilon_i(n)$ and $\varepsilon_j(n)$ are updated in the way given in (9) and (10) respectively. Therefore, the weight update equation of the proposed VS-LMS algorithm can be reformulated as:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\varepsilon_i(n)}{\varepsilon_j(n)} e(n)\mathbf{x}(n). \tag{12}$$

The sufficient condition to guarantee the mean weight vector convergence of the proposed VS-LMS algorithm is easily derived as follows according to [12]:

$$0 < E[\mu_n] < \frac{2}{\lambda_{\max}}, \tag{13}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}$.

For the NLMS algorithm [1], the weight update equation is given by:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\lambda + \|\mathbf{x}(n)\|^2} e(n)\mathbf{x}(n), \tag{14}$$

where λ is a small positive constant used to avoid the singularity of the input norm $\|\mathbf{x}(n)\|^2 = \sum_{k=0}^{M-1} x^2(n-k)$, where M is the coefficient length of the adaptive filter. The step-size μ ($0 < \mu < 2$) controls the convergence speed and misadjustment of the NLMS algorithm.

In [16], the authors have analyzed a class of nonlinearly-modified stochastic gradient algorithms and generalized the modified NLMS algorithm as follows:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + e(n)\frac{\mathbf{x}(n)}{1/\mu + \|\mathbf{x}(n)\|^2}. \tag{15}$$

The proposed VS-NLMS algorithm in this paper has the weight update equation given by

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\varepsilon_i(n)}{\varepsilon_j(n)} \frac{e(n)\mathbf{x}(n)}{\lambda + \|\mathbf{x}(n)\|^2}, \tag{16}$$

where $\varepsilon_i(n)$ and $\varepsilon_j(n)$ are updated in the way given in (9) and (10) respectively.

III. MSE Convergence and Performance Analysis of the Proposed VS-LMS Algorithm

In this section, the MSE convergence and the steady-state performance of the new VS-LMS algorithm are to be analyzed. It is assumed that the exact modeling of the unknown system is considered, where the measurement noise $v(n)$ is zero mean white Gaussian with variance σ_v^2, which is independent of the zero mean white Gaussian input sequence $x(n)$. It is also assumed that the input signal $x(n)$ and the desired output $d(n)$ are jointly stationary and independent. The desired output $d(n)$ can then be expressed as $d(n) = \mathbf{w}_o^T \mathbf{x}(n) + v(n)$, where the parameter $\mathbf{w}_o$ is the optimal weight vector to be identified.

To study the MSE convergence, the steady-state mean behavior of the variable step-size is discussed first. We assume the algorithm is close to the steady state where the MSE is described as [1]:

$$\lim_{n\to\infty} E[e^2(n)] = \xi_{\min} + \xi_{VS\_excess} = \sigma_v^2 + \xi_{VS\_excess}, \tag{17}$$

where $\xi_{\min} = \sigma_v^2$ denotes the minimum MSE, and $\xi_{VS\_excess} = (\mathbf{w}(\infty) - \mathbf{w}_o)^T \mathbf{R}(\mathbf{w}(\infty) - \mathbf{w}_o)$ represents the excess MSE produced by the VS-LMS algorithm.

Multiplying both sides of (7) by $\sum_{j=0}^{n} \beta^j e^2(n-j)$, we obtain:

$$\mu_n \sum_{j=0}^{n} \beta^j e^2(n-j) = \sum_{i=0}^{n} \alpha^i e^2(n-i). \tag{18}$$

Taking expectations on both sides of (18), we have:

$$E\left[\mu_n \sum_{j=0}^{n} \beta^j e^2(n-j)\right] = E\left[\sum_{i=0}^{n} \alpha^i e^2(n-i)\right]. \tag{19}$$

Since the algorithm is assumed to close to the steady-state, the variable step-size is expected to be slow varying, which justifies the following approximation

$$E\left[\lim_{n\to\infty} \mu_n \sum_{j=0}^{n} \beta^j e^2(n-j)\right] \approx E[\mu_\infty] E\left[\lim_{n\to\infty} \sum_{j=0}^{n} \beta^j e^2(n-j)\right]. \tag{20}$$

Using (17), (19) and (20), we obtain the following approximation for the mean behavior of the step-size in the steady-state:

$$E[\mu_\infty] \approx \frac{\lim_{n\to\infty} \sum_{i=0}^{n} \alpha^i E\left[e^2(n-i)\right]}{\lim_{n\to\infty} \sum_{j=0}^{n} \beta^j E\left[e^2(n-j)\right]} \approx \frac{\lim_{n\to\infty} \sum_{i=0}^{n} \alpha^i \left(\sigma_v^2 + \xi_{VS\_excess}\right)}{\lim_{n\to\infty} \sum_{j=0}^{n} \beta^j \left(\sigma_v^2 + \xi_{VS\_excess}\right)} = \frac{1-\beta}{1-\alpha}. \tag{21}$$

Notice that the following term is approximated in [12] as:

$$\lim_{n\to\infty} E[e^4(n)] \approx 3\left(\xi_{\min} + \xi_{VS\_excess}\right)^2. \tag{22}$$

Using (7), the same argument of (20), and the result of (22), we have

$$E\left[\mu_\infty^2\right] \approx \frac{\alpha(1+\beta)(1-\beta)^2}{\beta(1+\alpha)(1-\alpha)^2}. \tag{23}$$

According to [12], the sufficient condition for the MSE convergence is described as

$$0 < \frac{E\left[\mu_\infty^2\right]}{E\left[\mu_\infty\right]} \le \frac{2}{3Tr(\mathbf{R})}. \tag{24}$$

Substituting the results of (21) and (23) into (24), the sufficient condition for the MSE convergence of the proposed VS-LMS is given by

$$0 < \frac{\alpha\left(1-\beta^2\right)}{\beta\left(1-\alpha^2\right)} \le \frac{2}{3Tr(\mathbf{R})}. \tag{25}$$

According to [11], the excess MSE of the VS-LMS is described as

$$\xi_{VS\_excess} \approx \frac{E\left[\mu_\infty^2\right]}{2E\left[\mu_\infty\right]} \xi_{\min} Tr(\mathbf{R}). \tag{26}$$

Using (21) and (23), the excess MSE of the proposed VS-LMS algorithm is approximated by

$$\xi_{VS\_excess} \approx \frac{\alpha\left(1-\beta^2\right)}{2\beta\left(1-\alpha^2\right)} \xi_{\min} Tr(\mathbf{R}). \tag{27}$$

It can be seen from (27) that the excess MSE is determined by the parameters α, β, the additive noise power, and the input signal power. Also note from (25) that the choices of α and β are important to ensure the MSE stability of the new algorithm. For stationary environment, the typical values of α and β are selected close to unity for fast convergence.

IV. SIMULATION RESULTS

In this section, the simulations are demonstrated to show a better performance of the proposed VS-LMS algorithm and VS-NLMS algorithm by comparing with the MVSS algorithm [11] and MNLMS algorithm [16] in adaptive channel equalization and system identification respectively. The standard LMS algorithm [1] and NLMS algorithm [1] are also used in the comparisons. All the simulation plots are obtained by ensemble averaging 500 independent simulation runs.

A. *Adaptive Channel Equalization*

This simulation compares the new VS-LMS algorithm with the MVSS and the standard LMS. In the experiment setup of adaptive channel equalization, the input signal applied to the channel is a random Bernoulli sequence with $r(n) = \pm 1$, which has zero mean and unit variance. The channel is described as [1]:

$$h(n) = \begin{cases} 0.5\left(1+\cos\dfrac{2\pi(n-2)}{W}\right), & \text{n=1,2,3} \\ 0, & \text{otherwise} \end{cases},$$

where the parameter W is known to control the amount of amplitude distortion. An increase in W increases the channel distortion thereby increasing the eigenvalue spread of the autocorrelation matrix $\mathbf{R}$ of the inputs. The channel is corrupted by an additive white Gaussian noise $v(n)$, which has zeros mean and variance of σ_v^2=0.001. The tap input signal $x(n)$ is described as $x(n) = h(n) * r(n) + v(n)$. The equalizer has 11 taps. Since the impulse response of the channel is symmetric about $n = 2$, it follows that the optimum tap weights of the equalizer are symmetric about the fifth tap. Accordingly, the desired response of the equalizer is provided by the channel input with seven samples delay. The simulation was performed with the parameter W equaling 2.9 and 3.3, which provides eigenvalue spread 6.0782 and 21.7132 respectively. In all the simulations, we used $\mu_{\max} = 0.0833$ and $\mu_{\min} = 10^{-5}$ for the new algorithm and the MVSS.

Fig.1 shows the MSE behaviors of the three algorithms for the small eigenvalue spread of $W = 2.9$. The parameters for the proposed VS-LMS algorithm used are $\alpha = 0.98$, and $\beta = 1-10^{-5}$. We used $\alpha = 0.97$, $\beta = 0.99$ and $\gamma = 1.0$ for the MVSS algorithm according to [11]. In addition, we used $\mu = 0.01$ for the LMS algorithm.

Fig.2 shows the MSE behaviors of the three algorithms for the high eigenvalue spread of $W = 3.3$. The MVSS algorithm has the change of $\gamma =$ 0.2. The step size of the LMS was changed to $\mu = 0.035$ .The parameters for the proposed VS-LMS remained the same as for Fig.1. Other parameters ot the MVSS algorithm remain unchanged.

As seen from Fig.1 and Fig.2, the proposed VS-LMS algorithm provides the fastest convergence speed and the smallest misadjustment among all other algorithms in both cases.

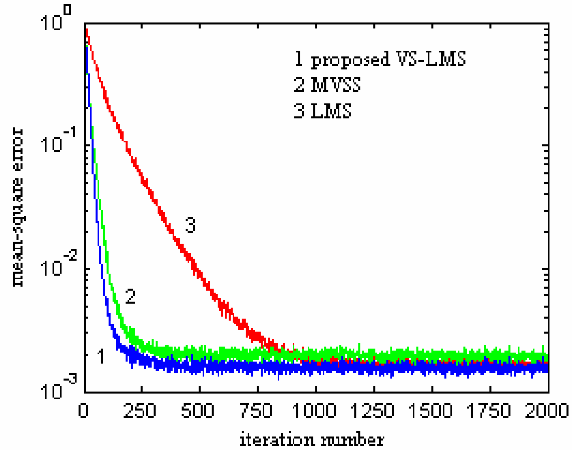

Figure 1. Comparison of MSE behaviors of all algorithms with $W = 2.9$.

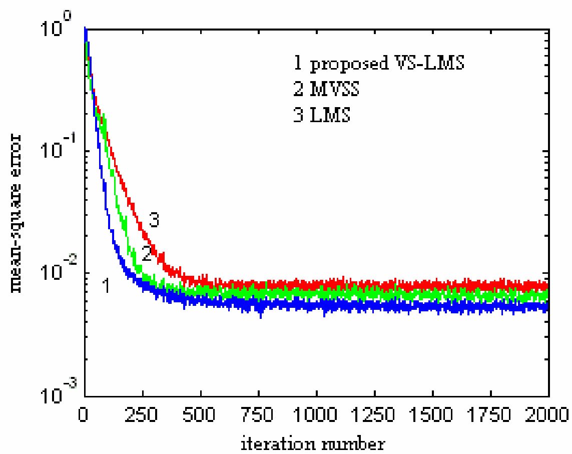

Figure 2. Comparison of MSE behaviors of all algorithms with $W = 3.3$.

B. *System Identification*

This simulation compares the proposed VS-NLMS algorithm with the MNLMS algorithm and the NLMS algorithm in the implementation of system identification. The cases of uncorrelated data and correlated data for stationary and nonstationary environments were demonstrated.

In all the simulations, the unknown system is assumed to be FIR and have a length of 10. The additive noise $v(n)$ is assumed to be white Gaussian sequence with zero mean and variance of σ_v^2=0.01 or -20dB. The optimal parameters were chosen to achieve comparable misadjustment. In all the simulations, we used $\mu_{\max} = 1.0$ and $\mu_{\min} = 10^{-5}$ for the new algorithm.

1) *White Gaussian Input for Stationary Environment*

In this case, the white Gaussian noise with zero mean and unit variance is used as input signal. The system coefficients are time invariant. To obtain comparable misadjustment, we used $\alpha = 0.98$, $\beta = 0.999$ for the proposed VS-NLMS algorithms, $\mu = 0.0041$ for the MNLMS algorithm. In addition, we used $\mu = 0.03$ for the NLMS algorithm. Fig.3

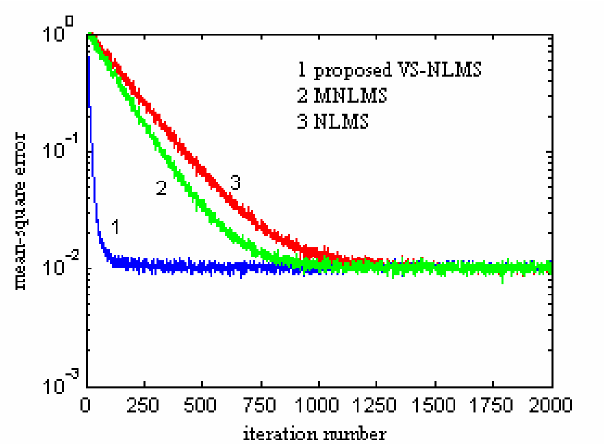

Figure 3. Comparison of MSE behaviors of all algorithms for white Gaussian input and stationary environment.

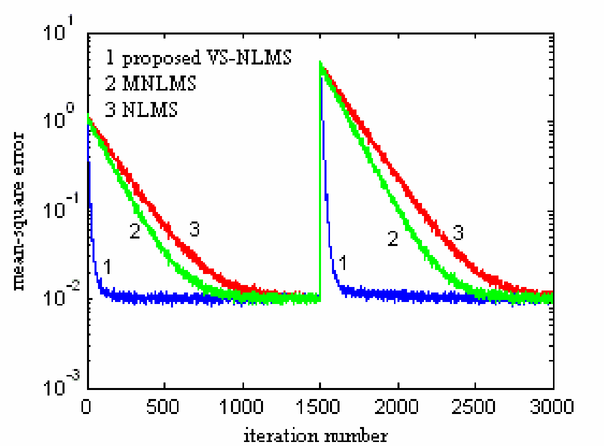

Figure 4. Comparison of MSE behavior of all algorithms with an abrupt change in system coefficients.

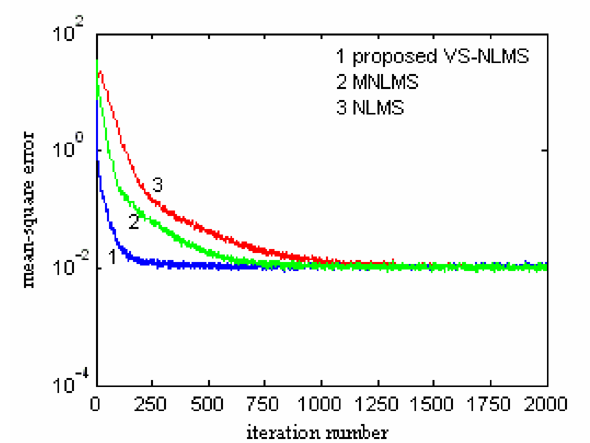

Figure 5. Comparison of MSE behavior of all algorithms with correlated input and stationary environment.

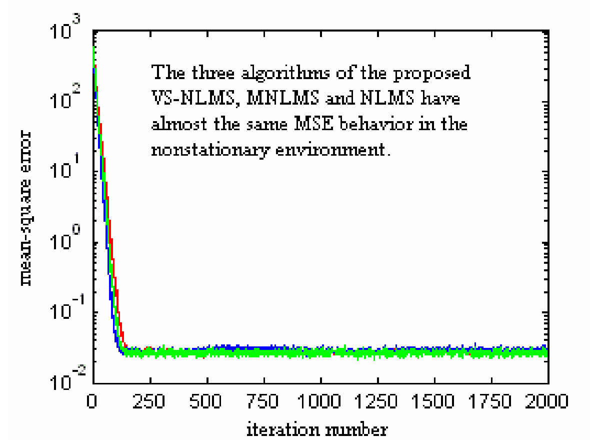

Figure 6. Comparison of MSE behavior of all algorithms with white Gaussian input and nonstationary environment.

shows that the proposed VS-NLMS algorithm provides the fastest convergence speed among all other algorithms.

2) White Gaussian Input and Abrupt Change in the System Coefficients

In this case, an abrupt change is set at the iteration of 1500 while remaining all other conditions as the same as in *1)*. Fig.4 shows the robustness of the proposed VS-NLMS algorithm, which sustains the fastest convergence speed after the abrupt change in the plant. Note that the MNLMS and the NLMS algorithms also can detect the abrupt change but provide slow tracking speed.

3) Correlated Input for Stationary Environment

In this case, the input signal is correlated, which was generated by

$$x(n) = 0.9x(n-1) + r(n),$$

where the signal $r(n)$ is a white Gaussian noise with zero mean and unit variance. The independence between $r(n)$ and the internal system noise $v(n)$ was ensured. To obtain the comparable misadjustment, the parameters used for the proposed VS-NLMS algorithm were $\alpha = 0.99$, $\beta = 0.999$, and for the MNLMS algorithm , the parameter used was $\mu = 0.0016$. In addition, we used $\mu = 0.028$ for the NLMS algorithm. Fig.5 shows the MSE behaviors of the compared algorithms. It is observed that the proposed VS-NLMS algorithm achieves the fastest convergence speed among all other algorithms.

4) White Gaussian Input for Nontationary Environment

A time-varying system is modeled in this case and its coefficients vary in the way of a random walk process [1] given by

$$h(n+1) = h(n) + c(n),$$

where $c(n)$ is a white Gaussian noise with zero mean and small variance $\sigma_c^2 = 10^{-4}$. The comparable misadjustment was obtained by setting the parameters as $\alpha = 0.99$, $\beta = 0.992$ for the proposed VS-NLMS algorithm, $\mu = 0.2$ for the MNLMS algorithm, and $\mu = 0.5$ for the NLMS algorithm. Fig.6 shows the three algorithms have almost the same MSE behaviors and the minimum level of MSE obtained by all algorithms is about -18dB

V. CONCLUSIONS

In this paper, two modified algorithms of VS-LMS and VS-NLMS with a new variable step-size are proposed to improve the convergence performance as well as the misadjustment. In the proposed algorithms, the new variable step-size is computed using a ratio of the sums of the weighted energy of the output error with two different exponential factors. Computer simulations have been demonstrated to show that the proposed VS-LMS and VS-NLMS algorithms provide better performances in the implementations of the adaptive channel equalization and system identification comparing with other existing algorithms.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, Nj: Prentice Hall, 2002.

[2] B. Farhang-Boroujeny, "On statistical efficiency of the LMS algorithm in system modeling," *IEEE Trans. Signal Process*. vol.41, pp. 1947-1951, 1993.

[3] M. Mboup, M. Bonnet, and N. Bershad, "LMS coupled adaptive prediction and system identification: A statistical model and transient mean analysis," *IEEE Trans. Signal Process.* vol. 42, pp. 2607-2615, 1994.

[4] S. Qureshi, "Adaptive equalization data transmissions," *IEEE Commun, Mag*., vol. 20, pp. 2-16, 1982.

[5] J. makhoul, "Linear predicationi: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561-580, 1975.

[6] P. Prandoni and M. Vetterli, "An FIR cascade structure for adaptive linear prediction," *IEEE Trans. Signal Process.* vol. 46, pp. 2566-2571, 1998.

[7] S. Ikeda and A. Sugiyanma, "An adaptive noise canceller with low signal distortion for speech codecs," *IEEE Trans. Signal Process.* vol. 47, no. 3, pp. 665-674, March 1999.

[8] J. E. Greenberg, "Modified LMS algorithms for speech processing with an adaptive noise canceller," *IEEE Trans. Speech Audio Process.* vol. 6, no 4, pp. 338-358, July 1998.

[9] D.W. Kim, J.H.Choi, Y.S. Choi, C.H. H. Y. Ko, "A VS-LMS algorithm using normalized absolute estimation error," *Proceedings. IEEE TENCON. Digital Signal Processing Applications.* vol. 2, pp. 692-697. Nov. 1996.

[10] Z. Ramadan, A. Poularikas, "A robust variable step-size LMS algorithm using error-data normalization," *Proc. of the IEEE SoutheastCon*, April 2005, pp. 219 – 224.

[11] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: Analysis and simulations," *IEEE Trans. Signal Process.* vol. 45, no. 3 pp. 631-639, March 1997.

[12] R. H. Kwong and E. W. Hohnston, "A variable step-size LMS algorithm," *IEEE Trans. Signal Process.* vol. 40, pp. 1633-1642, July 1992.

[13] F. Casco, H. Perez, R. Marcelin, M. Lopez, "A two step size NLMS adaptive filter algorithm,"*Conference Proceedings on Singapore ICCS*. vol. 2, Nov. 1994 pp.814 – 819.

[14] Z. Ramadan and A. Poularikas, "A variable step-size adaptive noise canceller using signal to noise ratio as the controlling factor," *Proc. of the 36th IEEE Southeaster Symposium on System Theory (SSST),* Atlanta, Georgia, March 2004, pp. 456-461.

[15] Z. Ramadan and A. Poularikas, "Performance analysis of a new variable step-size LMS algorithm with error nonlinearities," *Proc. of the 36th IEEE Southeaster Symposium on System Theory (SSST),* Atlanta, Georgia, March 2004, pp. 384-388.

[16] S. C. Douglas and T. H. Meng, "Normalized data nonlinearities for LMS adaptation," *IEEE Trans. Signal Process.* vol. 42, no.6, pp. 1352-1365, June 1994.

[17] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Process.* vol. ASSP-33, no.1, pp. 222-229, Feb 1985.