Proceedings of the World Congress on Engineering and Computer Science 2008
WCECS 2008, October 22 - 24, 2008, San Francisco, USA

The Generalized Square-Error-Regularized LMS Algorithm

Junghsi Lee, Hsu-Chang Huang, and Yung Ning Yang

***Abstract*—The purpose of a variable step-size normalized LMS filter is to solve the problem of fast convergence or low mis-adjustment. While most VSS-NLMS algorithms need to tune several parameters for better performance, we presented a tuning-free square-error-based VSS-NLMS algorithm recently [4]. In this paper, we propose a more general structure of that of [4], the generalized square-error-regularized LMS (GSER-LMS) algorithm. Extensive simulation results demonstrate that our GSER-LMS outperforms existing schemes in speed of convergence, tracking ability, and low mis-adjustment.**

***Index Terms*—Adaptive filters, normalized least mean square (NLMS), variable step-size NLMS, regularization parameter.**

I. Introduction

Adaptive filtering algorithms have been widely employed in many signal processing applications. The normalized least mean square (NLMS) adaptive filter is very popular because of its simplicity and robustness. It is well-know that the stability of NLMS is controlled by a step-size parameter μ, which also controls the speed of convergence, tracking ability and steady-state mis-adjustment of the filter. In practice, the NLMS is implemented by dividing the step-size parameter by the squared norm of the input vector plus a small positive constant ε called the regularization parameter. The inclusion of ε alleviates the problem when the squared norm getting too close to zero in certain applications. Since the overall step-size affects the performance of the NLMS, this regularization parameter ε has an effect on the convergence properties and mis-adjustment as well.

There are conflicting objectives between fast convergence and low mis-adjustment for fixed regularized NLMS algorithms. In the past two decades, many variable step-size NLMS (VSS-NMS) algorithms have been proposed to solve this dilemma associated with the fixed regularized NLMS [1-8]. Mandic [6] presented a generalized normalized gradient descent (GNGD) algorithm which used a time-varying regularization parameter $\varepsilon(n)$. Mandic claimed that the GNGD adapts its learning rate according to the dynamics of the input signals, and its performance is bounded from below by the performance of the NLMS. Very recently, Mandic introduced another scheme with hybrid filters structure to further improve the steady-state mis-adjustment of the GNGD [5]. Choi, Shin, and Song [3] introduced a normalized gradient in the update process for the regularization parameter and proposed a modified GNGD which improves the steady-state performance.

While most VSS-NLMS algorithms need to tune several parameters for better performance, we presented a tuning-free square-error-based VSS-NLMS algorithm [4] recently. Our new regularized NLMS algorithm outperforms existing schemes in convergence, tracking, and mis-adjustment. In this paper, we propose a more general structure of that of [4], the generalized square-error-regularized LMS (GSER-LMS) algorithm.

II. Generalized Square-error- regularized LMS Algorithm

In this section, we summarize several algorithms including NLMS, GNGD algorithm [6], and Choi's regularized NLMS [3]. We then present the generalized square-error-regularized LMS.

Let $d(n)$ be the desired response signal of the adaptive filter

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n) + v(n)\,, \tag{1}$$

where $\mathbf{h}(n)$ denotes the coefficient vector of the unknown system with length N,

$$\mathbf{h}(n) = [h_0(n), h_1(n), \ldots, h_{N-1}(n)]^T\,, \tag{2}$$

$\mathbf{x}(n)$ is the input vector

$$\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T\,, \tag{3}$$

and $v(n)$ is the additive noise that is independent of $x(n)$.

A. Conventional Regularization for NLMS (ε-NLMS)

Let the adaptive filter have the same structure and same order as that of the unknown system. Denoting the coefficient vector of the filter at iteration n as $\mathbf{w}(n)$. We may express the *a priori* estimation error as

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n)\,. \tag{4}$$

The ε-NLMS algorithm updates $\mathbf{w}(n)$ as follows

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\|\mathbf{x}(n)\|_2^2 + \varepsilon} e(n)\mathbf{x}(n)\,, \tag{5}$$

where μ is the fixed positive step-size and ε is a fixed small positive constant called regularization parameter. For signals that have great spectrum, the overall effective step-size $\mu\big/\left(\|\mathbf{x}(n)\|_2^2 + \varepsilon\right)$ might vary radically, and the

Manuscript received August 8, 2008. This work was supported in part by the TAIWAN National Science Council under grant NSC 96-2221-E-155-008, and by the Yuan-Ze University.

J. Lee, H. C. Huang, and Y. N. Yang are with the Department of Electrical Engineering, Yuan-Ze University, Chung-Li, Taoyuan, 32026, TAIWAN (Phone+886-3-4638800,Ext.7116;Fax+886-3-4639355;
E-Mail:eejlee@saturn.yzu.edu.tw)

value of ε may notably affect the convergence and tracking performance of the ε-NLMS adaptive filter.

B. *GNGD algorithm* [6]

The GNGD algorithm uses a time-varying regularization parameter $\varepsilon(n)$ calculated as

$$\varepsilon(n) = \varepsilon(n-1) - \rho\mu \frac{e(n)e(n-1)\mathbf{x}^T(n)\mathbf{x}(n-1)}{(\|\mathbf{x}(n-1)\|_2^2 + \varepsilon(n-1))^2}, \quad (6)$$

where ρ is an adaptation parameter needs tuning, and the initial value $\varepsilon(0)$ has to be set as well.

C. *Choi's Regularized NLMS Algorithm* [3]

To improve the steady-state performance of GNGD, Choi proposed a time-varying regularization parameter as

$$\varepsilon'(n) = \varepsilon(n-1) - \rho\,\mathrm{sgn}\left[e(n)e(n-1)\mathbf{x}^T(n)\mathbf{x}(n-1)\right]$$
$$\varepsilon(n) = \begin{cases} \varepsilon'(n), & if\ \varepsilon'(n) \ge \varepsilon_{\min} \\ \varepsilon_{\min}, & if\ \varepsilon'(n) < \varepsilon_{\min} \end{cases}, \quad (7)$$

where $\mathrm{sgn}(x)$ represents the sign function, and $\varepsilon_{\min}$, a minimum allowable value of $\varepsilon(n)$, is a parameter needs tuning.

D. *Proposed GSER-LMS Algorithm*

For the conventional ε-NLMS algorithm, the role of ε is to prevent the associated denominator from getting too close to zero, so as to keep the filter from divergence. However, in applications of speech signals, a too small ε may make the denominator very close to zero while a too big ε will slow down the adaptation of the filter. In this paper, we propose a generalized square-error-regularized LMS which employs the inverse of the weighted square-error as the time-varying regularization parameter. The proposed GSER-LMS algorithm updates $\mathbf{w}(n)$ as follows

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\|\mathbf{x}(n)\|_2^2 + \alpha\delta^{-1}(n)} e(n)\mathbf{x}(n), \quad (8)$$

where α is a positive parameter that makes the filter more general, and the error signal power $\delta(n)$ can be estimated as

$$\delta(n) = (1-\lambda)\delta(n-1) + \lambda e^2(n). \quad (9)$$

The positive constant λ is less than, and usually close to, unity.

The idea behind GSER-LMS is that when error signal power $\delta(n)$ gets bigger, the regularization parameter $\alpha\delta^{-1}(n)$ gets smaller, and the effective step-size becomes relatively large. Consequently, the filter can make bigger adaptation at this mode. When the estimation error is small, our $\alpha\delta^{-1}(n)$ gets larger, and it gives smaller effective step-size. At this stage, the adaptive filter makes small adjustment. A good property of our GSER-LMS is that, in practice, neither $\delta(n)$ nor $\delta^{-1}(n)$ gets too close to zero. Therefore, we do not need to set an $\varepsilon_{\min}$ as that of Choi's algorithm [3].

III. Simulation Results

In this section, we present the results of several experiments that compare the performance of ε-NLMS, GNGD, Choi's NLMS, and our GSER-LMS. The adaptive filter was used to identify a 512-tap acoustic echo system. The acoustic echo impulse response was set to be time-varying from seconds 1.9 to 5.1. The evolution of coefficients is described by

$$\mathbf{h}(n) = \mathbf{h}_o + g(n), \quad (10)$$

where $\|\mathbf{h}_o\| = 1$ and $g(n)$ is a white Gaussian noise with variance 10^{-4}. We have used the normalized squared coefficient error (NSCE) to evaluate the performance of the algorithms. The NSCE is defined as

$$NSCE(n) = 10\log_{10} \frac{\|\mathbf{h}(n) - \mathbf{w}(n)\|^2}{\|\mathbf{h}(n)\|^2} \quad (11)$$

where $\mathbf{w}(n)$ is the filter coefficient vector. We have run extensive simulations. The results are pretty consistent. In this section, we show simulation results with the following setup: $\varepsilon_{\min} = 0.001$, $\rho = 0.15$, and $\mu = 1$.

A. *AR processes*

We have used AR processes as the reference input signals. The power of each AR process is approximately 1. The NSCE curves shown here are results of ensemble averages over 20 independent runs. Figures 1 and 2 demonstrate the results of AR(1) as reference input signal and the additive noise is a zero mean white Gaussian with variance 10^{-2} and 10^{-4}, respectively. The results showed that ε-NLMS and GNGD had similar performance. Choi's algorithm performed reasonably well. Our GSER-LMS ($\alpha = 10$) outperformed others in the case shown in Fig. 1, and GSER-LMS ($\alpha = 0.1$) was the best in the case shown in Fig. 2.

B. *Speech Signals*

In this experiment, the excitations are 8-second-long Chinese speech signals. Speech I was given in Figure 3, and results of additive white Gaussian noise with variance 10^{-2} and 10^{-4} were illustrated in Figures 4 and 5, respectively. The results showed that ε-NLMS performed badly. GNGD and Choi's NLMS did not perform consistently. It is obvious to see that our filters performed consistently well.

Simulation results of Speech II and Speech III were illustrated in Figures 6-11. Notice that Choi's filter had a pretty good performance in one case (shown in Fig. 8.) However, just like that observed earlier, GNGD and Choi's algorithm had unsatisfactory performance in some cases. It is clear that our GSER-LMS outperformed other competing algorithms.

IV. Conclusions

This paper presented a general square-error-regularized LMS algorithm. Extensive simulation results demonstrated that our GSER-LMS algorithm performed consistently well. However, we observed that the parameter α may affect the performance of GSER-LMS. We are currently investigating this issue.

REFERENCES

[1] M. T. Akhtar, M. Abe, and M. Kawamata, "A new variable step size LMS algorithm-based method for improved online secondary path modeling in active noise control systems," *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 2, pp. 720-726, March 2006.

[2] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Processing Letters*, Vol. 13, No. 10, pp 581-584, Oct. 2006.

[3] Y. S. Choi, H. C. Shin, and W. J. Song, "Robust regularization for normalized LMS algorithms," *IEEE Transactions on Circuits and Systems II, Express Briefs*, Vol. 53, No. 8, pp. 627–631, Aug. 2006.

[4] J. Lee, H. C. Huang, Y. N. Yang, and S. Q. Huang, "A square-error-based regularization for normalized LMS algorithms," Proceedings of IMCES 2008, pp. 1399-1402, March, 2008.

[5] D. P. Mandic, P. Vayanos, C. Boukis, B. Jelfs, S. L. Goh, T. Gautama, and T. Rutkowski; "Collaborative adaptive learning using hybrid filters," *Proceedings of 2007 IEEE ICASSP*, pp. III 921–924, April 2007.

[6] D. P. Mandic; "A generalized normalized gradient descent algorithm," *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp. 115–118, Feb. 2004.

[7] H. C. Shin, A. H. Sayed, and W. J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp 132-135, Feb. 2004.

[8] J. M. Valin and I. B. Collings, "Interference-normalized least mean square algorithm," *IEEE Signal Processing Letters*, Vol. 14, No. 12, pp. 988-991, Dec. 2007.

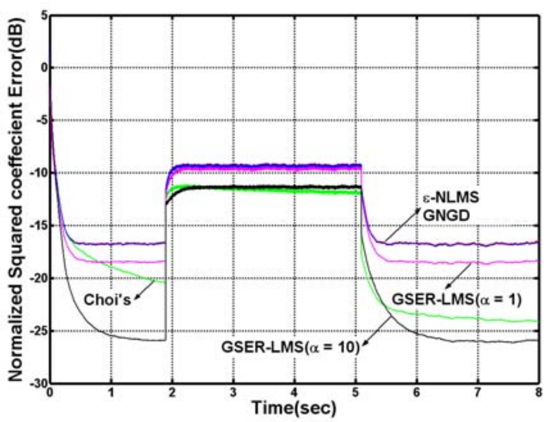

Fig. 1, NSCE curves of ε-NLMS, GNGD [6], Choi's [3], and our GSER-LMS. Input signal is AR(1). Additive noise is zero mean white Gaussian with variance 0.01 .

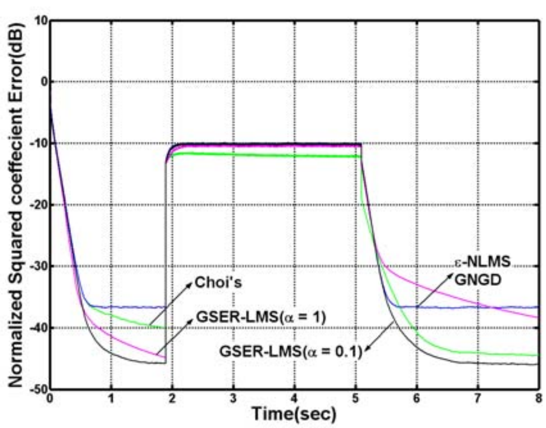

Fig. 2, NSCE curves of ε-NLMS, GNGD [6], Choi's [3], and our GSER-LMS. Input signal is AR(1). Additive noise is zero mean white Gaussian with variance 0.0001 .

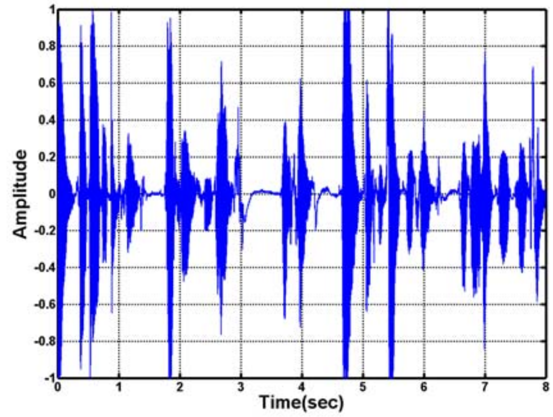

Fig. 3, Speech I.

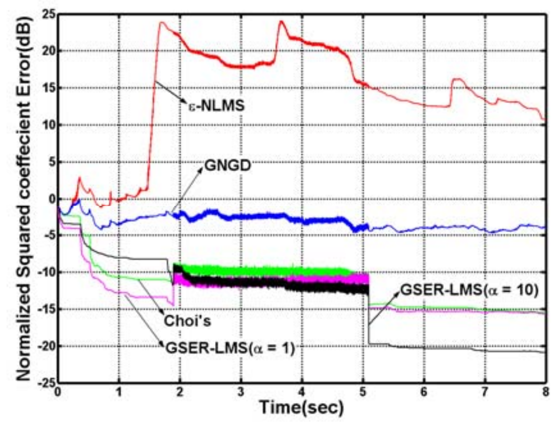

Fig. 4, NSCE curves of ε-NLMS, GNGD [6], Choi's [3], and our GSER-LMS. Input signal is Speech I. Additive noise is zero mean white Gaussian with variance 0.01 .

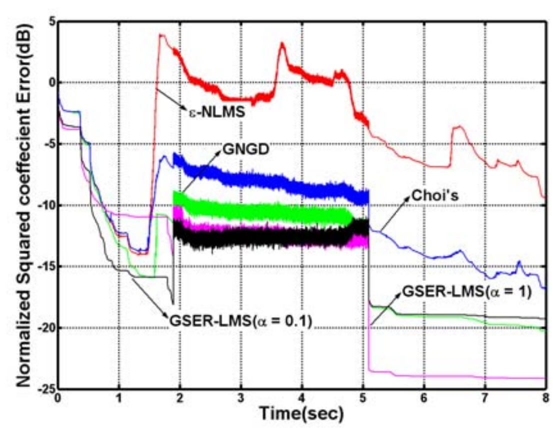

Fig. 5, NSCE curves of ε-NLMS, GNGD [6], Choi's [3], and our GSER-LMS. Input signal is Speech I. Additive noise is zero mean white Gaussian with variance 0.0001 .

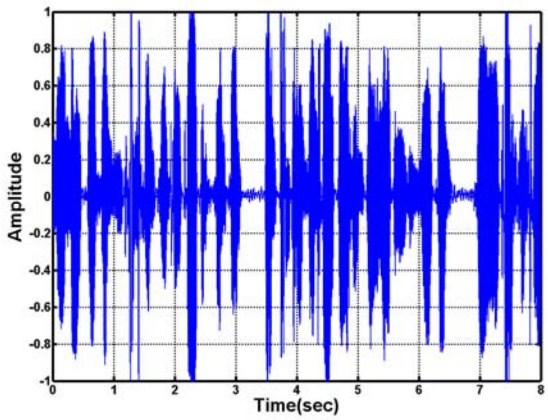

Fig. 6, Speech II.

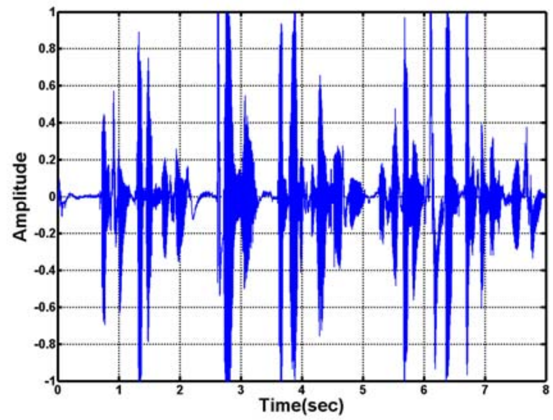

Fig. 9, Speech III.

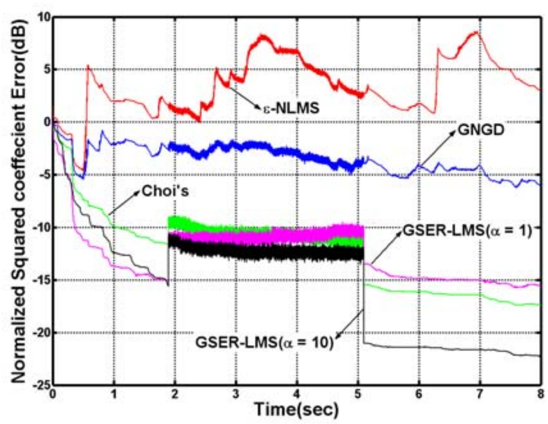

Fig. 7, NSCE curves of ε-NLMS, GNGD [6], Choi’s [3], and our GSER-LMS. Input signal is Speech II. Additive noise is zero mean white Gaussian with variance 0.01.

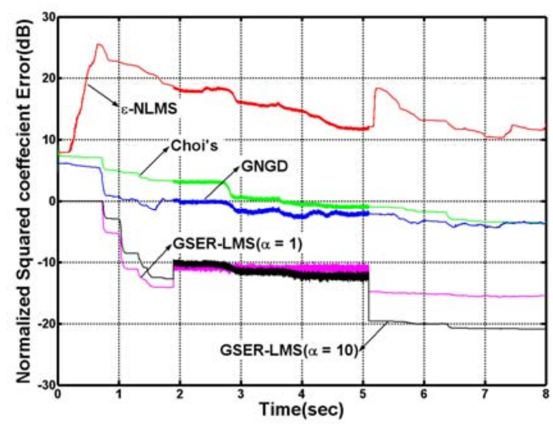

Fig. 10, NSCE curves of ε-NLMS, GNGD [6], Choi’s [3], and our GSER-LMS. Input signal is Speech III. Additive noise is zero mean white Gaussian with variance 0.01.

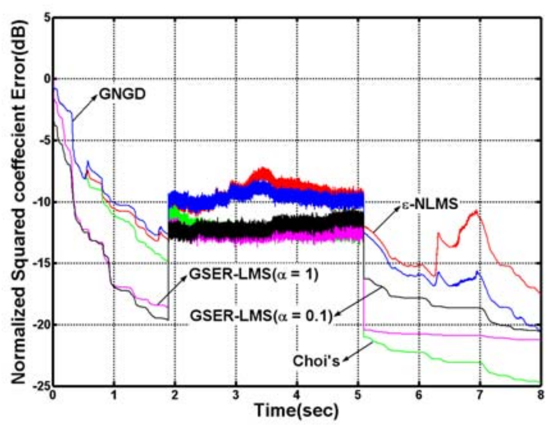

Fig. 8, NSCE curves of ε-NLMS, GNGD [6], Choi’s [3], and our GSER-LMS. Input signal is Speech II. Additive noise is zero mean white Gaussian with variance 0.0001

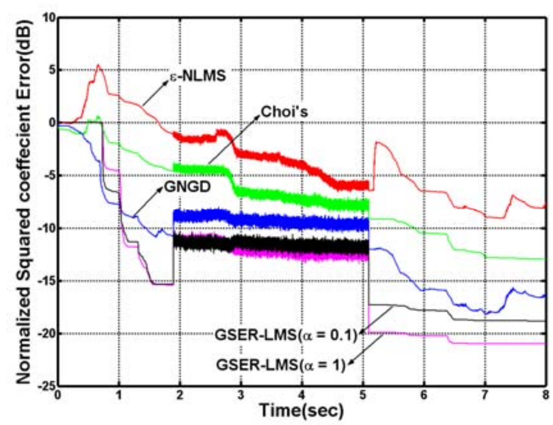

Fig. 11, NSCE curves of ε-NLMS, GNGD [6], Choi’s [3], and our GSER-LMS. Input signal is Speech III. Additive noise is zero mean white Gaussian with variance 0.0001

ISBN: 978-988-98671-0-2