Sensors & Transducers

Improved Variable Step-Size and Variable Parameters LMS Adaptive Filtering Algorithm

Huamin GE, Weiwei ZHAO
School of Information Science & Control, Nanjing University of Information Science & Technology,
Nanjing 210044, China
Tel.: 86-025-83466421, fax: 86-025-83466421
E-mail: ghm@nuist.edu.cn

Received: 1 September 2013 /Accepted: 25 October 2013 /Published: 30 November 2013

Abstract: In order to achieve the best filtering performance and make the filter automatically adjust its coefficients with the changing of its working environment, we need to use adaptive filter technique. To overcome the contradiction of traditional LMS algorithm between convergence speed, convergence accuracy and the selection of step-size, a novel variable step-size & variable parameters adaptive filter algorithm based on modified hyperbolic tangent function was proposed, whose theory and process were worked out and the design principle of adjusting factors were given. Theory and simulation results verified the convergence properties and superiorities of this algorithm under different circumstances. *Copyright © 2013 IFSA.*

Keywords: Adaptive filter, Variable step-size LMS algorithm, LabVIEW.

1. Introduction

The least mean square (LMS) adaptive algorithm has been used in numerous applications for its outstanding advantages. But the fixed step-size LMS has a tradeoff between the rate of convergence and the steady-state mean square error. To overcome this contradiction, numerous variable step-size LMS algorithms were proposed in papers [1-10]. These algorithms were all based on Sigmoid function and corresponding to the principle of variable step-size (VSS), they can achieve a better convergence performance and a faster tracking behavior when compared to other algorithms whereas their stabilities were bad and computational costs were too big. So in this paper, we proposed a more general VSSLMS based on Hyperbolic tangent function and the combination of fixed and variable step-size. The step-size of this algorithm can vary with the instantaneous estimation error and the energy of input signal for it was adjusted by the mean value of the sum of gradient vector and the self-correlation of the instantaneous estimation error. It can eliminate the influences of the measurement noises; overcome the defects that the step-size would decline quickly and the initial convergence speed was too slow in the above algorithms. Theory and simulation results verified that our algorithm has lower steady-state MSE or misadjustment, faster convergence speed, lesser computational cost, better stability, simpler structure and better tracking ability.

2. Algorithm Formulation and its Performance Analysis

Based on the theories of LMS and the principle of VSS, we address the conflict with the following step-size adjustment scheme:

$$\mu(n) = \beta(n)\tanh(\alpha(n)^k \|F(n)\|^m) \quad (1)$$

where $\| \|$ denotes abs. $\alpha(n)$ controls the shape of step-size, $\beta(n)$ controls the range of step-size, they are all functions vary with $e(n)$. To figure out $F(n)$, $\alpha(n)$ and $\beta(n)$, we must first introduce the theories of LMS as follows:

$$y(n) = \boldsymbol{w}^H(n)\boldsymbol{x}(n) \quad (2)$$

$$e(n) = d(n) - y(n) \quad (3)$$

$$d(n) = \boldsymbol{w}_o^H(n)\boldsymbol{x}(n) + v(n) \quad (4)$$

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu(n)e^*(n)\boldsymbol{x}(n) \quad (5)$$

The schematic diagram of LMS was shown in Fig. 1 where $\boldsymbol{x}(n) = [x(n),...,x(n-N+1)]^H$ is the tap-delayed input vector at n moment. $\boldsymbol{w}(n) = [w_0(n), w_1(n),..., w_{N-1}(n)]^H$ is the corresponding weight vector consisting of the coefficients of the adaptive filter. $d(n)$ represents the desired response consisting of system output plus the measurement noise. $y(n)$ is the output signal. $e(n)$ is the instantaneous estimation error. $v(n)$ denotes measurement noise statistically independent of $\boldsymbol{x}(n)$. $\boldsymbol{w}_o(n)$ is the optimal Wiener solution. All signals above can be complex. N denotes the length of adaptive filter. H denotes conjugate transpose. $*$denotes complex conjugate. Letters written in bold means that the signal is a vector.

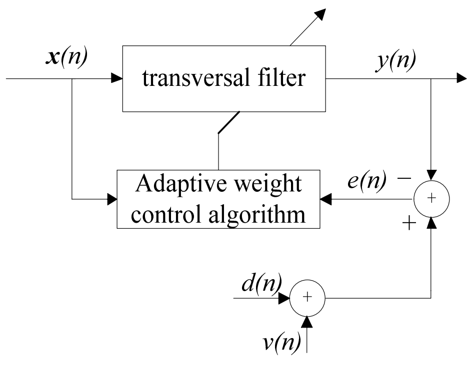

Fig. 1. Schematic diagram of adaptive filter.

From (2),(3) and (4) we have:

$$e(n) = v(n) - [\boldsymbol{w}^H(n) - \boldsymbol{w}_o^H(n)]\boldsymbol{x}(n) \quad (6)$$

Considering $\boldsymbol{\rho}(n) = \boldsymbol{w}^H(n) - \boldsymbol{w}_o^H(n)$ and σ_v^2 as the power of measurement noise, we get:

$$\begin{aligned} &E[\boldsymbol{x}(n)v(m)] = 0 \\ &E[v(n)v^*(m)] = \sigma_v^2\delta(n-m) \end{aligned} \quad (7)$$

Then:

$$\begin{aligned} &E[e^2(n)] = E[\boldsymbol{\rho}^H(n)\boldsymbol{x}(n)\boldsymbol{x}^H(n)\boldsymbol{\rho}(n)] + \sigma_v^2 \\ &E[e(n)e(n-1)] = E[\boldsymbol{\rho}^H(n)\boldsymbol{x}(n)\boldsymbol{x}^H(n-1)\boldsymbol{\rho}(n-1)] \\ &E[e(n)e(n-1) + e^2(n)] = E[e(n)e(n-1)] + E[e^2(n)] \\ &E[e(n)\boldsymbol{x}(n)] = E[-\boldsymbol{\rho}^H(n)\boldsymbol{x}(n)\boldsymbol{x}^H(n)] \end{aligned} \quad (8)$$

The convergence in mean sense of a system can be represented by $E[\mu(n)]$. So according to Eq. (8), if we introduce $e(n)$ or the positive integer power of $e(n)$ in $\mu(n)$, the system would be very sensitive to measurement noise $v(n)$. If we introduce $e(n)e(n-1)$ only, the system would not be so sensitive to $v(n)$, but the self-correlation of $e(n)$ would getting smaller as the length of filter getting longer, this would surely smaller the step-size, slow down the convergence speed and deteriorate the performance of the algorithm in limited time.

If we introduce $e(n)e(n-1) + e^2(n)$, the performance in higher order system would be improved, but it would be still affected by $v(n)$. From paper [11] we know that the steepest drop direction of adaptive LMS algorithm is $-\nabla J(n) = 2e(n)\boldsymbol{x}(n)$, make $\nabla J(n) = 0$, we can get the optimal solution of $\boldsymbol{w}(n)$. Furthermore, the optimal Wiener solution is $e(n)\boldsymbol{x}(n)$, from Eq. (8)we know $E[e(n)\boldsymbol{x}(n)]$ is insensitive to $v(n)$. So we introduce $e(n)\boldsymbol{x}(n)$ to adjust step-size in this paper. Actually, the cross-correlation of $\boldsymbol{x}(n)$ and $e(n)$ is stronger than the self-correlation of $e(n)$, hence, the step-size would not decrease sharply in high order system. Besides, this expression can stronger the tracking ability for it considers input signal in variable step-size formula. In this paper, we adopt the average value of the sum of Gradient vector to update step-size. This formula can avoid the weight vector from fluctuating when the amplitude of input signal is too high or fluctuate significantly, so it can improve the performance of the algorithm.

Based on the analyses above, we can get the formula of $\alpha(n)$, $\beta(n)$ and $F(n)$ in Eq. (1) as follows: $\alpha(n)$ uses the quotient form of $e(n)$, this is mainly because it is proportional to $e(n)$ and we should ensure smaller $\alpha(n)$ as far as possible to avoid the step-size oversizing along with $F(n)$. $\beta(n)$ uses the recursive form of the self-correlation of $e(n)$ to restrain the effect of independent measurement noise further. Besides, when the algorithm converging to optimal value, the self-correlation of $e(n)$ will getting smaller to smaller the steady-state error:

$$F(n) = \frac{\sum_{i=0}^{N-1} e(n)x(n-i)}{N} \quad (9)$$

$$\alpha(n) = \left\| \frac{e(n)}{e(n-1)} \right\| \qquad (10)$$
$$\beta(n) = r\beta(n-1) + (1-r)\|e(n)e(n-1)\|$$

When $e(n-1)$ in (10) equals zero, $\alpha(n)$ approach to ∞, this would enforce $\tanh(\infty)$ approach to 1. So in this case, the algorithm would not unsteady but could reflect the real state of the system more accurately. r is a constant approach to 1, it controls $\beta(n)$, $0 < r < 1$, in this paper, $r = 0.98$.

Considering the design principle of the two adjusting factors m and k in (10). m controls the shape of the function, we can make the step-size vary slowly in steady-state by changing its value. When $\|F(n)\| < 1$, the step-size would get larger as the value of m getting smaller. So, if the value of m is too large, the step-size might equals zero even if the system have not steady, this would surely bigger the steady-state error. If the value of m is too small, the step-size would be too big even if $e(n)$ is tiny, still, this would bigger the steady-state maladjustment error and interfere the convergence performance of the system. When $\|F(n)\| > 1$, the functions of m is opposite to $\|F(n)\| < 1$. Factor k controls the varying speed of the step-size via adjusting $\alpha(n)$. The functions of k when $\alpha(n) < 1$ and $\alpha(n) > 1$ are corresponding to the functions of m when $\|F(n)\| < 1$ and $\|F(n)\| > 1$. So, the choosing of m and k should considering the following two occasions: Bigger m and k if we need higher convergence speed; Smaller m and k if we need higher convergence accuracy.

To overcome the defect that the initial convergence speed was too slow, we use the combination of fixed and variable step-size, that is, at the first l sampling points, we make the step-size a constant between 0.1 to 1, then at the following sampling points use Eq. (1) to update step-size.

The algorithm of this paper is combined by (1), (9) and (10). Theoretically, the new algorithm uses fixed step-size at the first l sampling points, it can achieve a faster initial convergence speed and smaller initial estimation error; The algorithm has little computational cost and simple structure for it was adjusted by Hyperbolic tangent function; $e(n)$ and the average value of Gradients vector make the step-size of the algorithm getting smaller along with the convergence progress and would not decrease sharply in high order system, this is corresponding to the principle of variable step-size: speed up the rate of convergence using large step-size at the initial stages, and reduce the estimation error with small step-size in steady-state; The algorithm is able to reduce its sensitivity to the measurement noise $v(n)$ and keep in good performances in low SNR environment for it was adjusted by factors independent to $v(n)$; The algorithm is able to response the variation of the input signals and the whole system no matter in stationary or nonstationary environment for it was adjusted by Gradients vector.

3. Simulation Results in LabVIEW

This section presents four examples to corroborate the analytic results derived in this paper. The algorithm was applied to noise cancelation and it was simulated in LabVIEW for real-time applications. In the examples, algorithms in paper [5], paper [12] and NLMS were compared to demonstrate that our algorithm can lead to better performances no matter in Low SNR, abrupt system variation or abrupt input signal variation environments. Besides, the design principles of adjusting factors m and k are also given below.

To begin with, let us describe the experimental setup and parameter settings: $N = 10$, $\beta(0) = 0$, $r = 0.98$; Initial weight vector $\boldsymbol{w}(0) = [0, 0, ..., 0]^H$; Input signal $\boldsymbol{x}(n) = \sin(n) + a(n)$, where $a(n)$ is zero-mean, white Gaussian noise to be removed and $\sin(n)$ is the standard unit sine signal; Desired response $d(n) = \sin(n)$ where $\sin(n)$ is corresponding to the $\sin(n)$ in $\boldsymbol{x}(n)$; The results were obtained by averaging 200 ensemble trials with $n = 1000$ each trial. Our algorithm is called VVLMS here. The experimental parameter settings of paper [5] and [12] are the same with paper [5] and [12], the step-size in NLMS here is 0.002.

Example 1. Design. principle of adjusting factors m and k.

Under the simulation environment above, $\|F(n)\|$ is at a level of 10^{-3}, so the step-size would decrease following the increase of m, but if m is too large, the step-size might equals zero even if the system have not steady, this would surely bigger the steady-state error as the curve show in Fig. 2 when $m = 15$. Too small m would interfere the fluctuation of the weight vector, then interfere the performance of the system as the curve show in Fig. 3 when $m = 0.05$. These are consistent with our analytic results. Here we take the optimal $m = 2$ considering both the fluctuation of the weight vector and the MSE.

The functions of different k to the system can be neglected under the environment above when m is fixed (figure not shown). Here we take $k = 0.05$ to give consideration to both little and large $\alpha(n)$.

Example 2. The performances of the algorithms with an abrupt input signal variation.

Fig. 4 shows the learning curves of our algorithm and the three other algorithms with an abrupt input signal variation at iteration 500 that the input signal changes to a triangle wave with 0.2 as its amplitude. The variance of $a(n)$ is 0.2. Here we take $l = 50$, then when $n < 50$ in our algorithm, we make $\mu(n) = 0.1$. It can be seen that our algorithm already converge at iteration 50. When the input signal changes at iteration 500, the algorithms in Fig. 4 can all adaptive converge with the variance of the environment, and the tracking speed of our algorithm is obviously faster than the other algorithms at iteration 500. Fig. 5 shows the average of the weight vector of the algorithms, the more evident the fluctuation of the weight vector, the worse performances of the algorithm. It can be easily seen that our weight vector is the most smooth whereas the weight vector of paper [5] fluctuates significantly and the weight vector of NLMS can't converge to optimal solution. Combine Fig. 4 and Fig. 5 we notice that our algorithm can achieve a better convergence performance when compared with other algorithms.

Example 3. The performances of the algorithms with an abrupt system variation

Fig. 6 shows the learning curves of our algorithm and the three other algorithms with an abrupt system variation at iteration 500 that the weight vector that has already identified turn back to zero again. Fig. 7 shows the average of the weight vector of the algorithms. Combine Fig. 6 and Fig. 7, following the analyze procedure in example 2 we know that our algorithm can achieve a stronger robustness and a better convergence performance when compared with other algorithms under these simulation conditions.

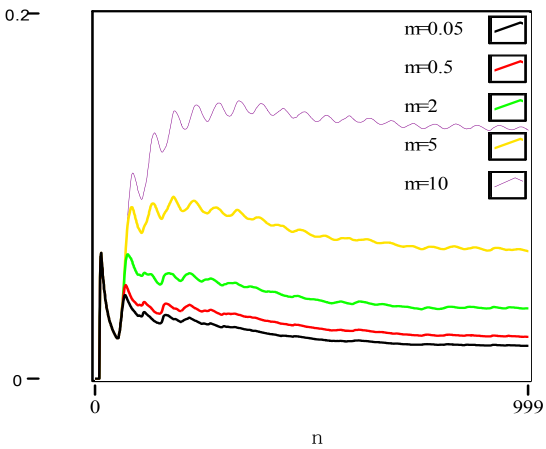

Fig. 2. Comparison of MSE with differernt m .

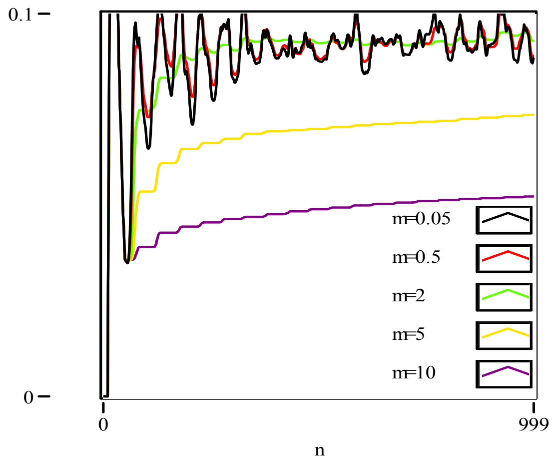

Fig. 3. Comparison of weight vector with differernt m .

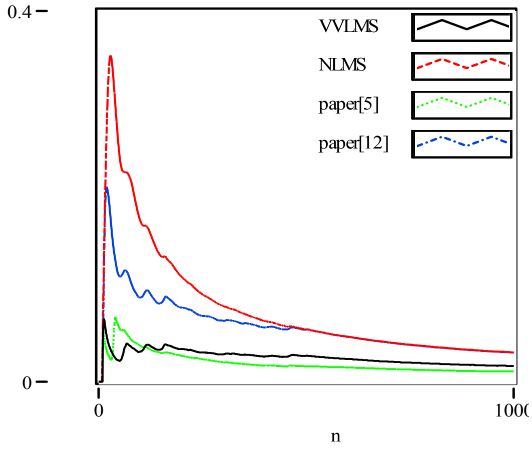

Fig. 4. Comparison of MSE with an abrupt input signal variation.

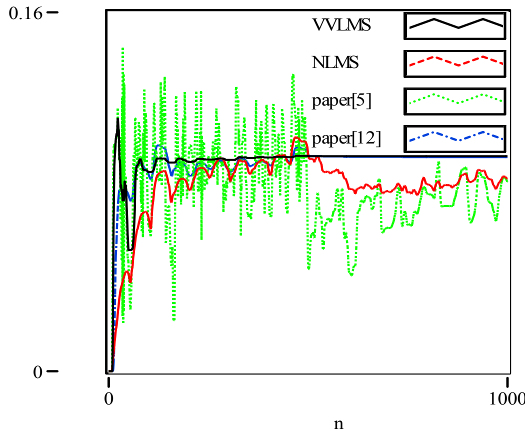

Fig. 5. Comparison of weight vector with an abrupt input signal variation.

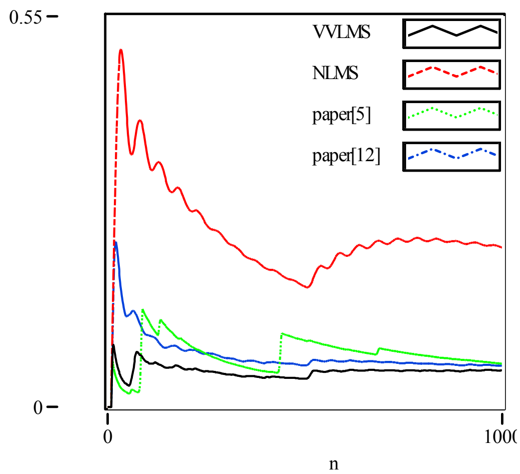

Fig. 6. Comparison of MSE with an abrupt system variation.

Example 4. The performances of the algorithms with Low SNR.

The variance of $a(n)$ is 0.8 here. Other simulation conditions remain unchanged. Fig. 8 shows the filtering results of the algorithms, it can be easily seen that our algorithm can still converge to the desired response approximately whereas other algorithms may diverge and can't converge to the desired response. Keep lowing SNR, our algorithm still have incomparable convergence performance.

In practical applications, we can filter the input signals repeatly if their amplitude is too big.

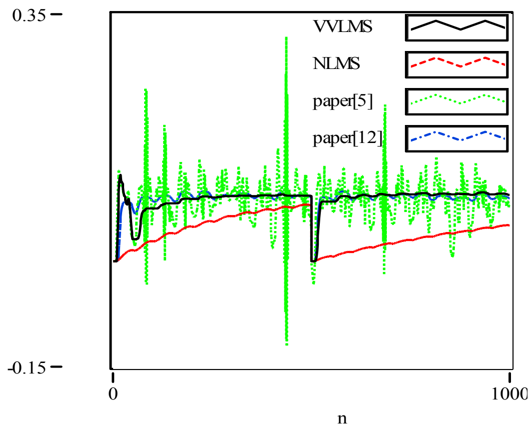

Fig. 7. Comparison of weight vector with an abrupt system variation.

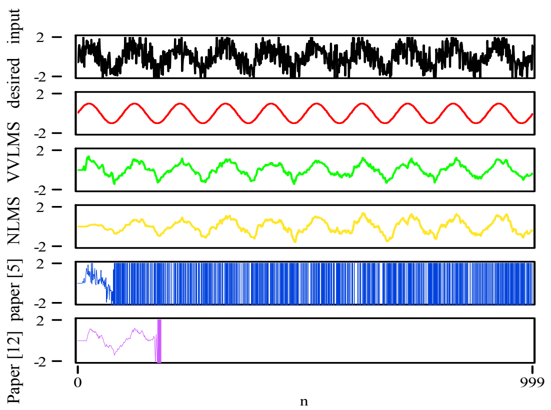

Fig. 8. Comparison of the filtering results with Low SNR.

4. Conclusions

This work presented a new variable step-size&variable parameter adaptive filter algorithm based on the nonlinear relation between the step-size and the average value of the sum of gradient vector. Unlike many existing approaches, we adjust the variable step-size by Hyperbolic tangent function and the mean value of the sum of gradient vector, then the step-size would adaptively decrease following the convergence progress and can abruptly increase when there is an abrupt system or input signal variation. The algorithm is very insensitive to measurement noise and has little computational cost and simple structure. The analytic and simulation results verified that our algorithm has faster initial convergence speed, lower steady-state MSE or misadjustment, faster convergence and tracking speed, lesser computational cost and stronger robustness.

References

[1]. S. B. Gelfand, Y. Wei, J. Vr. Krogmeie, The stability of variable step-size LMS algorithms, *IEEE Transactions on Signal Processing*, Vol. 47, Issue 12, 1999, pp. 3277-3288.

[2]. Hua Ye, Boxiu Wu, The study on a variable step size adaptive filtering algorithm, *Acta Electronica Sinica*, Vol. 18, No. 4, 1990, pp. 63-69.

[3]. R. Kwong, E. W. Johnston, A variable step size LMS algorithm, *IEEE Transactions on Signal Processing*, Vol. 40, Issue 7, 1992, pp. 1633–1642.

[4]. Jingfan Qin, Jingzheng Ouyang, A novel variable step size LMS adaptive filtering algorithm based on sigmoid function, *Journal of Data Acquisition & Processing*, Vol. 12, Issue 3, 1997, pp. 171-174.

[5]. Ying Gao, Sheng-Li Xie, A variable step size LMS adaptive filtering algorithm and its analysis, *Acta Electronica Sinica*, Vol. 29, No. 8, 2001, pp. 1094-1097.

[6]. M. H. Costa, J. C. M. Bermudez, A noise resilient variable step-size LMS algorithm, *Signal Processing*, Vol. 88, Issue 3, 2008, pp. 733–748.

[7]. Shengkui Zhao, Zhihong Man, et al., Variable step-size LMS algorithm with a quotient form, *Signal Processing*, Vol. 89, Issue 1, 2009, pp. 67–76.

[8]. Yonggang Zhang, Ning Li, et al., New gradient-based variable step size LMS algorithms, *EUR-ASIP Journal on Advances in Signal Processing*, 2008, Article ID 529480.

[9]. Jeng-Kuang Hwang, et al., Variable step-size LMS algorithm with a gradient-based weighted average, *IEEE Signal Processing Letters*, Vol. 16, Issue 12, 2009, pp. 1043-1046.

[10]. K. Mayyas, F. Momani, An LMS adaptive algorithm with a new step-size control equation, *Journal of the Franklin Institute*, Vol. 348, Issue 4, 2011, pp. 589–605.

[11]. S. Haykin, Adaptive filter theory, 4th ed., *Prentice-Hall*, Englewood Cliffs, NJ, 2002.

[12]. En-Chang Sun, Yu-Heng Li, Dong-Ying Zhang, Adaptive variable-step size LMS filtering algorithm and its analysis, *Journal of System Simulation*, Vol. 19, Issue 14, 2007, pp. 3172-3175.