35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

A Novel Gradient Adaptive Step Size LMS Algorithm with Dual Adaptive Filters

Yuzhong Jiao, Rex Y. P. Cheung, Winnie W. Y. Chow and Mark P. C. Mok

Abstract— **Least mean square (LMS) adaptive filter has been used to extract life signals from serious ambient noises and interferences in biomedical applications. However, a LMS adaptive filter with a fixed step size always suffers from slow convergence rate or large signal distortion due to the diversity of the application environments. An ideal adaptive filtering system should be able to adapt different environments and obtain the useful signals with low distortion. Adaptive filter with gradient adaptive step size is therefore more desirable in order to meet the demands of adaptation and convergence rate, which adjusts the step-size parameter automatically by using gradient descent technique. In this paper, a novel gradient adaptive step size LMS adaptive filter is presented. The proposed algorithm utilizes two adaptive filters to estimate gradients accurately, thus achieves good adaptation and performance. Though it uses two LMS adaptive filters, it has a low computational complexity. An active noise cancellation (ANC) system with two applications for extracting heartbeat and lung sound signals from noises is used to simulate the performance of the proposed algorithm.**

I. Introduction

Least mean square (LMS) adaptive filter is often used in biomedical applications to remove ambient noise or extract life signals from serious noises and interferences [1-6]. Good adaptation is one of the important features of an adaptive filtering system. A medical device with an adaptive filter had better be able to adjust its step size on the basis of different application environment. Moreover, the step size should be large at the early stage of iteration for fast convergence and becomes small at the stage of convergence to obtain useful signals with small distortion. Due to the diversity of the application environments, a LMS adaptive filter with a fixed step size always suffers from slow convergence, and large steady-state mean-square error (MSE) or large signal distortion, when extracting useful signals from noises or interferences. Gradient adaptive step size adaptive filter is a desirable solution to meet the demands of adaptation and convergence rate, which adjusts the step-size parameter automatically by using gradient descent technique [7-12].

* Research supported by Hong Kong Innovation and Technology Fund.

Yuzhong Jiao is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong, China (phone: +852-3406-0365; fax: +852-3406-2801; e-mail: yzjiao@astri.org).

Rex Y. P. Cheung is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong, China (e-mail: rexcheung@ astri.org).

Winnie W. Y. Chow is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong, China (e-mail: winniechow@astri.org).

Mark P. C. Mok is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong, China (e-mail: markmok@astri.org).

Benveniste et al firstly proposed gradient adaptive step size algorithm for LMS adaptive filter [7]. Without taking into account the independence assumptions, the algorithm was derived rigorously, resulting in good performance. Benveniste's algorithm actually performs time-varying low pass filtering of the noisy instantaneous gradients in the update of the step-size [8]. Then Mathews and Xie proposed a new gradient adaptive step size algorithm with lower computational complexity, which was based on several simplifying assumptions [9]. In Mathews' algorithm, only instantaneous gradient is used to estimate the new step size. Therefore, its performance is not as good as Benveniste's algorithm. For multiple step size technique, Ang and Farhang-Boroujeny simplified Benveniste's algorithm by replacing time-varying filtering of the instantaneous gradients with a fixed parameter low pass filter in the step-size update [11]. The performance of Ang's algorithm is comparable to that of Benveniste's, while the computational complexity of Ang's algorithm is lower than that of Benveniste's. However, for adaptive filters without multiple step size technique, Ang's algorithm has higher computational complexity than Benveniste's.

For gradient adaptive step size adaptive filters, the most important factor affecting the system performance is the accuracy of the gradient estimation. The more accurate the gradient is estimated, the better performance the adaptive filtering systems can achieve. The reason why Benveniste's algorithm achieves better adaptation is that the algorithm can obtain more accurate gradients. In this paper, a novel gradient adaptive step size LMS adaptive filter is proposed. The proposed algorithm utilizes two adaptive filters, which can estimate the gradients accurately, thus achieves good adaptation and performance.

The paper is organized as follows. LMS algorithm and several gradient adaptive step size algorithms are described in Section 2. Details of the proposed algorithm are given in Section 3. Performance and computational complexity are analyzed in Section 4. Section 5 presents the simulation results. The conclusion is made in the last section.

II. LMS and Gradient Step Size LMS Algorithm

The block diagram of an adaptive filter is illustrated in Fig. 1 [13]. The filter is a finite impulse response (FIR) filter with length N. The vector of tap inputs at time n is denoted by $X(n)$, which includes the tap inputs $x(n)$, $x(n\text{-}1)$,…, $x(n\text{-}N\text{+}1)$. The weight vector at time n is denoted by $W(n)$ which includes tap weights $w_0(n)$, $w_1(n)$,…, $w_{N\text{-}1}(n)$. $x(n)$ is the reference input, $d(n)$ is the desired response, and $y(n)$ is the corresponding estimate of $d(n)$ at the filter output. By comparing the desired response and its estimate, an

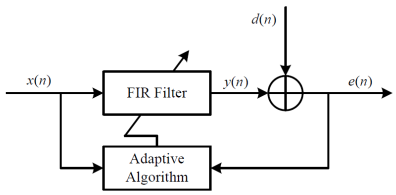

Figure 1. Block diagram of an adaptive filter

estimation error can be obtained:

$$\begin{aligned} e(n) &= d(n) - y(n) \\ &= d(n) - W^T(n)X(n) \end{aligned} \tag{1}$$

where $(\cdot)^T$ is the vector transpose operator.

A. LMS algorithm

The LMS algorithm is described by the equation [13]:

$$W(n+1) = W(n) + \mu(n)e(n)X(n) \tag{2}$$

where $\mu(n)$ is the step size of the LMS adaptive filter, which controls the convergence rate. For traditional LMS adaptive filter, $\mu(n)$ is a fixed value. The condition of the algorithm achieving convergence is $0 < \mu(n) < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the maximal eigenvalue of the input signal autocorrelation matrix.

B. Gradient adaptive step size LMS algorithms

Gradient adaptive step size adaptive filter adapts the step size sequence using a gradient descent algorithm so as to reduce the squared-estimation error at each iteration. The update equation for step size $\mu(n)$ is given as [7, 9, 11]

$$\begin{aligned} \mu(n) &= \mu(n-1) - \frac{\rho}{2}\nabla_{\mu} e^2(n) \\ &= \mu(n-1) - \frac{\rho}{2}\frac{\partial e^2(n)}{\partial \mu(n-1)} \\ &= \mu(n-1) - \frac{\rho}{2}\frac{\partial^T e^2(n)}{\partial W(n)}\frac{\partial W(n)}{\partial \mu(n-1)} \\ &= \mu(n-1) + \rho e(n)X^T(n)\varphi(n) \end{aligned} \tag{3}$$

where ρ is a small positive constant that controls the adaptive behavior of step size, $\nabla_{\mu} e^2(n)$ is the gradient of squared error at time n, and $\varphi(n)=\partial W(n)/\partial\mu(n-1)$.

As for Mathews' algorithm,

$$\varphi(n) = e(n-1)X(n-1). \tag{4}$$

Mathews' algorithm uses instantaneous gradients for step size adaptation.

As for Benveniste's algorithm,

$$\begin{aligned} \varphi(n) = {} & [I - \mu(n-1)X(n-1)X^T(n-1)] \\ & \cdot \varphi(n-1) + e(n-1)X(n-1) \end{aligned} \tag{5}$$

where I is an identity matrix. Benveniste' algorithm in fact performs time-varying low pass filtering of the noisy instantaneous gradients in the update of the step size.

As for Ang's algorithm,

$$\varphi(n) = a\varphi(n-1) + e(n-1)X(n-1) \tag{6}$$

where a is a constant smaller than but close to one. Ang's algorithm is a simplified version of Benveniste's algorithm.

III. Proposed Gradient Adaptive Step Size LMS Algorithm

The proposed algorithm as shown in Fig. 2 uses two LMS adaptive filters. One is a work filter in the solid square; the other is a reference filter in the dashed square. The outputs of the two adaptive filters are:

$$e_w(n) = d(n) - W_w^T(n)X(n) \tag{7}$$

$$e_r(n) = d(n) - W_r^T(n)X(n) \tag{8}$$

where the suffixes "w" denote the work filter; and "r" denote the reference filter. The updating functions of the weight vectors are:

$$W_w(n+1) = W_w(n) + \mu_w(n)e_w(n)X(n) \tag{9}$$

$$W_r(n+1) = W_r(n) + \mu_r(n)e_r(n)X(n) \tag{10}$$

where $\mu_w(n)=\mu_r(n)+\Delta\mu$, and $\Delta\mu$ is a value close to zero. The weight vectors of the two filters have the same initial value. Thus, we can obtain

$$\lim_{\Delta\mu\to 0}\frac{e_w^2(n) - e_r^2(n)}{\Delta\mu} = \frac{\partial e^2(n)}{\partial\mu(n-1)}. \tag{11}$$

The limit means the gradient of squared error. So we can consider $(e_w^2(n) - e_r^2(n))/\Delta\mu$ as the approximation of the gradient denoted as $\tilde{\nabla}_{\mu} e^2(n)$ when $\Delta\mu$ is very small. The update equation of the step size for the work filter is

$$\mu_w(n) = \mu_w(n-1) - \frac{\rho}{2}\tilde{\nabla}_{\mu} e^2(n). \tag{12}$$

Though the work filter and reference filter have the same coefficients at the beginning; the difference between the parameters of the two filters increases gradually due to the small $\Delta\mu$. When the difference increases to some extent, it

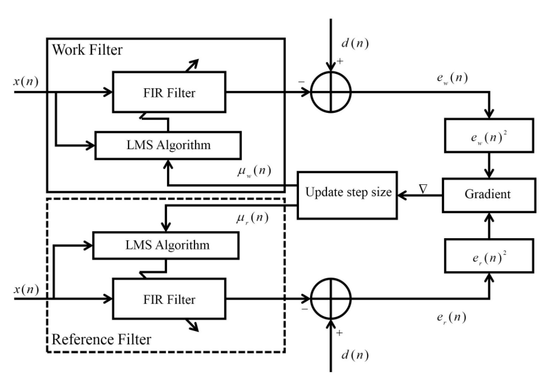

Figure 2. Block diagram of proposed gradient step size algorithm

becomes difficult to estimate the gradient of squared error. So we need to re-initialize the weight vector of the reference filter for every interval of length M

$$W_r(n+1)=\begin{cases}W_w(n)+\mu_r(n)e_w(n)X(n) & n(\operatorname{mod} M)=0\\ W_r(n)+\mu_r(n)e_r(n)X(n) & otherwise.\end{cases} \quad (13)$$

IV. Analysis of Performance and Computational Complexity

From (7-10) and (13), we can derive the approximation of the gradient of the proposed algorithm if $\Delta\mu\to 0$

$$\tilde{\nabla}_{\mu}e^2(n)\approx -2e(n)X^T\tilde{\varphi}(n) \quad (14)$$

where

$$\tilde{\varphi}(n)=\begin{cases}e(n-1)X(n-1) & n(\operatorname{mod} M)=0\\ [I-\mu(n-1)X(n-1)X^T(n-1)] \cdot\tilde{\varphi}(n-1)+e(n-1)X(n-1) & otherwise\end{cases} \quad (15)$$

$e(n)\approx e_w(n)\approx e_r(n)$ and $\mu(n)\approx\mu_w(n)\approx\mu_r(n)$.

The equations above show that the proposed algorithm is essentially the integration of Mathews' and Benveniste's algorithms. If M=1, the gradient of the proposed algorithm is almost the same as that of Mathews' algorithm. If $M=\infty$, the proposed algorithm has a similar gradient with Benveniste's algorithm. This means that the proposed algorithm can achieve a performance close to Benveniste's algorithm if the value of M is large enough.

Tab. I summarizes the computational complexities of Mathews', Benveniste's, Ang's and the proposed algorithms. In the table, MUL and MAC denote multiply operation and multiply–accumulate operation respectively. Though the proposed algorithm uses two filters, it has the same complexity as Benveniste's algorithm, lower than Ang's algorithm. One thing deserves to be mentioned is that the division in the proposed algorithm can be changed into one operation of multiplication or shift if $\Delta\mu$ at time n is known.

V. Simulations

In an active noise cancellation (ANC) system using adaptive filtering, the desired response $d(n)$ of the adaptive filter is the combination of the biomedical signals ε, and a noise derived from the reference input $x(n)$, which is the ambient noise, after passing an unknown system. Here we consider the unknown system as a five-point FIR filter with coefficients [9-10]:

$$W_o=\{0.1,\quad 0.3,\quad 0.5,\quad 0.3\quad 0.1\}. \quad (16)$$

The input signal $x(n)$ is a pseudorandom, zero-mean, and Gaussian process obtained as the output of the all-pole filter with transfer function

$$A(z)=\frac{0.44}{1-1.5z^{-1}+z^{-2}+0.25z^{-3}} \quad (17)$$

TABLE I. Computational Complexity of Gradient Algorithms

Algorithm	*Number of MUL/MAC*
Mathews'	$3N$
Benveniste's	$4N$
Ang's	$5N$
Proposed	$4N$

when the input to the filter is zero-mean, white, and pseudo-Gaussian noise with a variance of 1.

The parameters used are ρ=0.0004, a=0.95, $\Delta\mu=10^{-6}$, N=5, $\mu(0)=10^{-3}$, and μ_{max}=0.2. Fifty independent runs and 50,000 samples per run are used in the simulation. All results are the averages after 50 runs.

Here we perform two types of simulations: (1) ANC with measurement noise from electrical signal processing system, and (2) ANC with biomedical signals such as heartbeat and lung sound.

A. ANC with measurement noise

The aim of the simulation is to compare the step size behaviors of Mathews', Benveniste's and the proposed algorithms. In the simulation, the desired response signal $d(n)$ is obtained by adding the output of the system in (16) with zero-mean, 0.005 variance, white, and pseudo-Gaussian additive measurement noise . From the results shown in Fig. 3, we can learn the relationship between Mathews', Benveniste's and the proposed algorithms. When M=1, the step size of the proposed algorithm is close to that of Mathews' algorithm after tens of iterations. When $M=10^3$, the proposed algorithm almost has the same step size as Benveniste's algorithm after tens of iterations. The simulation results prove the inference in the previous section.

B. ANC with biomedical signals

In electronic stethoscopes, heartbeat and lung sound signals are extracted from ambient noises to obtain better signal quality. We use these two application cases to compare the performance of the four algorithms: Mathews', Benveniste's, Ang's, and the proposed algorithms when $M=10^3$. In the two applications, the desired response signals $d(n)$ are obtained by adding the output of the system in (16) with (1) heartbeat signal and (2) lung sound signal extracted

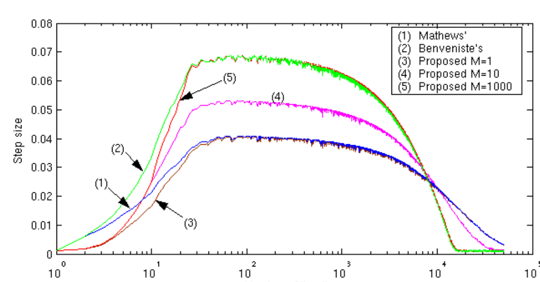

Figure 3. The comparison of Mathews', Benveniste's and the proposed algorithms in step size behaviors

from 3M Littmann stethoscope [14], respectively.

Fig. 4 shows the mean behavior of the step size of these gradient step size algorithms in the two applications. All algorithms can increase the values of the step size very rapidly from a small initial value, which is an advantage of gradient adaptive step size algorithm. In the two biomedical application cases, the proposed algorithm has the step sizes very close to those of Benveniste's algorithm. The step sizes of both algorithms decrease to smaller values than those of Ang's algorithm in the following iteration process. Mathews' algorithm does not decrease its step size in the following iterations, which is unfavorable to extract good biomedical signals.

Fig. 5 shows the MSE comparison between the four algorithms in the two biomedical applications. Benveniste's algorithm and the proposed algorithm show almost the same performance with smaller MSE, which implies less signal distortion. The MSE comparison in Fig. 5 supports the adaptive behavior of step size in Fig. 4.

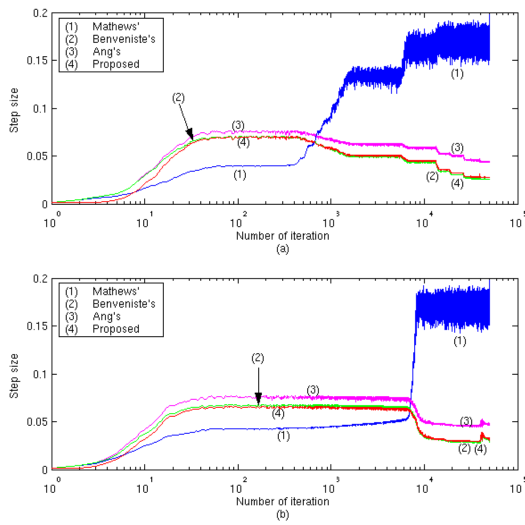

Figure 4. Step sizes for four gradient step size algorithms with (a) heartbeat signal, and (b) lung sound signal

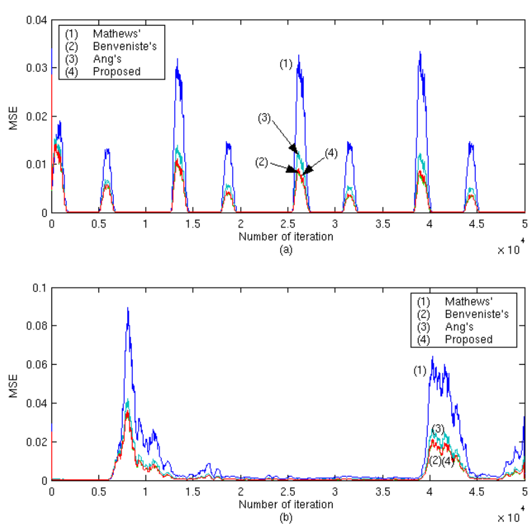

Figure 5. Mean squared error comparison between four gradient step size algorithms with (a) heartbeat signal, and (b) lung sound signal

VI. CONCLUSION

This paper presents a novel gradient adaptive step size algorithm with dual LMS adaptive filters. The algorithm is different from the traditional methods using gradient descent technique in that the gradient is measured with two LMS adaptive filters. However, the proposed algorithm can potentially be a new implementation form of Benveniste's algorithm. It almost has the same step size behavior, performance and computational complexity as Benveniste's algorithm. Simulation results demonstrate that the proposed algorithm achieves good adaptation and performance in biomedical applications.

REFERENCES

[1] C. H. –I. Kim, H. Soeleman and K. Roy, "Ultra-low-power DLMS adaptive filter for hearing aid applications", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 11, 2003, pp. 1058-1067.

[2] P. Gibbs and H. H. Asada, "Reducing motion artifact in wearable bio-sensors using MEMS accelerometers for active noise cancellation", in *Proceedings of the 2005 American Control Conference*, 2005, pp. 1581-1586.

[3] J. Gnitecki, Z. M. K. Moussavi, "Separating Heart Sounds from Lung Sounds - Accurate Diagnosis of Respiratory Disease Depends on Understanding Noises", *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, 2007, pp. 20-29.

[4] M. Rahman, R. A. Shaik and D. Reddy, "Noise Cancellation in ECG Signals using Computationally Simplified Adaptive Filtering Techniques: Application to Biotelemetry", *Signal Processing: An International Journal (SPIJ)*, vol. 3, 2009, pp. 120–131.

[5] F. Belloni, D. D. Giustina, M. Riva and M. Malcangi, "A new digital stethoscope with environmental noise cancellation" in *the 12th WSEAS International Conference on Mathematical and Computational Methods in Science and Engineering (MACMESE'10)*, 2010, pp. 169–174.

[6] Yuzhong Jiao, R. Y. P. Cheung and M. P. C. Mark, "Modified Log-LMS adaptive filter with low signal distortion for biomedical applications" in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 5210-5213.

[7] A. Benveniste, M. Metivier and P. Priouret, "Adaptive Algorithms and Stochastic Approximation", New York: Springer-Verlag, 1990.

[8] Su Lee Goh and Danilo P. Mandic, "A class of gradient-adaptive step size algorithms for complex-valued nonlinear neural adaptive filters", In *Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, 2005, pp. V/253-256.

[9] V. J. Mathews and Zhenhua Xie, "A Stochastic Gradient Adaptive Filter with Gradient Adaptive Step Size", *IEEE Transactions on Signal Processing,* vol. 41, 1993, pp: 2075-2087.

[10] Hong Chae Woo, "Improved stochastic gradient adaptive filter with gradient adaptive step size", *Electronics Letters*, vol. 34, 1998, pp. 1300-1301.

[11] Wee-Peng Ang and B. Farhang-Boroujeny, "A New Class of Gradient Adaptive Step-Size LMS Algorithms", *IEEE Transactions on Signal Processing,* vol. 49, 2001, pp: 805-810.

[12] Danilo P. Mandic, "A Generalized Normalized Gradient Descent Algorithm", *IEEE Siganl Processing Letters*, vol. 11, 2004, pp. 115-118.

[13] B. Widrow and S. D. Stearns, "Adaptive Signal Processing", New Jersey: Prentice-Hall, USA, 1985.

[14] 3M Littmann Stethoscope, "Heart and Lung Sounds", http://www.Littmann.com/wps/portal/3M/en_US/3M-Littmann/stethoscope/littmann-learning-institute/heart-lung-sounds/