Signal Processing 96 (2014) 321–324

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast communication

A variable step-size affine projection algorithm with a step-size scaler against impulsive measurement noise

Insun Song, PooGyeon Park *

Department of Electrical Engineering, Pohang University of Science and Technology, Pohang, Gyungbuk 790-784, Republic of Korea

ARTICLE INFO

Article history:
Received 31 May 2013
Received in revised form
9 September 2013
Accepted 12 September 2013
Available online 27 September 2013

Keywords:
Adaptive filters
Affine projection
Impulsive measurement noise
Step-size scaler
Robust filtering

ABSTRACT

This letter proposes a variable step-size (VSS) affine projection algorithm (APA) associated with a step-size scaler to improve the APA's robustness against impulsive measurement noise. In the proposed VSS APA, the step-size scaler is applied to the equations for updating the step size, which are developed by interpreting the behavior of the mean square deviation (MSD) of the conventional APA. To reduce the computational complexity, we also propose a simplified version of the step-size scaler, which is suitable for application in the APA. Simulations show that the proposed algorithm leads to an excellent transient and steady-state behavior with colored inputs in impulsive-noise environments.

1. Introduction

For highly correlated input data, the affine projection algorithm (APA) has improved the convergence rate as compared to the normalized least-mean-squares (NLMS) algorithm, because the APA uses a kind of self-orthogonalization process [1–3]. However, in most real environments, adaptive filtering applications can be influenced by various outliers, including impulsive measurement noise, which cause the performance of many adaptive filters to be degraded. In the conventional APA, the next weight is updated using the output error vector, which can be corrupted by impulsive measurement noise. To overcome this drawback in the APA, several APAs have been developed for application in impulsive-noise environments [4–7].

The affine projection sign algorithm (APSA) updates the weight according to the sign of the *a priori* error vector based on $\mathcal{L}_1$−norm optimization, and thus, it has the benefits of both the APA and sign algorithm: a fast convergence rate and robustness against impulsive measurement noise [4]. A variable step-size (VSS) APSA calculates the optimum step size of the APSA at each iteration by minimizing its mean square deviation (MSD) [5]. The robust VSS APA updates the filter estimate as a result of the minimization of the square norm of the *a posteriori* output error vector subject to a time-dependent constraint on the norm of the weight update [6]. A robust set-membership affine projection adaptive-filtering algorithm uses two error bounds, one of which is used to achieve faster convergence and the other to suppress impulsive measurement noise [7].

In gradient-based adaptive filtering, the next weight estimate is determined to reduce the cost function associated with the *a posteriori* output errors by using the gradient of the cost function with respect to the current weight vector, where the step size in $(0, 1]$ is usually designed to enhance the rate of convergence. Therefore, the cost function and the step size directly determine the performance of gradient-based adaptive algorithms. To handle impulsive measurement noise, researchers focus mainly on the cost function, rather than the step size. Recently, however, a novel method to improve the robustness against

* Corresponding author. Tel.: +82 54 279 2238; fax: +82 54 279 2903.
E-mail addresses: weedsis@postech.ac.kr (I. Song), ppg@postech.ac.kr (P. Park).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.09.008

impulsive measurement noise in gradient-based adaptive filtering was introduced [8]. In this method, the concept of the step-size scaler, which instantly scales down the step size whenever impulsive measurement noise appears, is applied, which eliminates the possibility of updating weight estimates being based on incorrect information attributed to impulsive measurement noise.

In this letter, we present a variable step-size (VSS) APA that uses a step-size scaler and equations for updating the step size, which are constructed by interpreting the behavior of the MSD of the conventional APA [9]. In [8], the step-size scaler is derived from a gradient-based adaptive algorithm based on a new cost function, which was developed by investigating and modifying the *tanh* cost function. Similarly, we derive the step-size scaler, which is suitable for application in the APA, from the cost function using the *a priori* error vector. To reduce the computational complexity of the step-size scaler, this letter also introduces a simplified version of the step-size scaler. Simulations show that the proposed algorithm leads to an excellent transient and steady-state behavior with colored inputs in impulsive-noise environments.

2. Step-size scaler in the APA

In a model for system identification, the desired signal d_i is represented as

$$d_i = \mathbf{u}_i^T \mathbf{w} + v_i. \tag{1}$$

Here, $\mathbf{w} \in \Re^{n\times 1}$ is a coefficient vector of the unknown system, $\mathbf{u}_i \in \Re^{n\times 1}$ denotes an input vector such as $[u_i, \ldots, u_{i-n+1}]^T$, a scalar variable $v_i = b_i + \eta_i$, where a scalar variable u_i is colored with variance σ_u^2 and b_i denotes the white Gaussian measurement noise with $N(0, \sigma_b^2)$, and η_i is impulsive measurement noise. Let $\widehat{\mathbf{w}}_i$ be an estimate of $\mathbf{w}$ at the ith iteration and e_i be an *a priori* measurement error vector, defined as

$$\mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i^T \widehat{\mathbf{w}}_i \in \Re^{M\times 1}, \tag{2}$$

where $\mathbf{d}_i = [d_i, \ldots, d_{i-M+1}]^T \in \Re^{M\times 1}$ and $\mathbf{U}_i = [\mathbf{u}_i, \ldots, \mathbf{u}_{i-M+1}] \in \Re^{n\times M}$. Here, M is the projection order and subscript $()^T$ denotes the matrix/vector transpose operation.

In [8], a new cost function that uses the square value of the normalized instant output error with respect to the input vector was introduced. Similarly, we apply the cost function using the *a priori* measurement error vector $\mathbf{e}_i$ as follows:

$$J_1(\widehat{\mathbf{w}}_i) = \frac{1}{\beta_1} \tanh\left(\frac{\beta_1 c_i}{2}\right) = \frac{1}{\beta_1} \frac{1-\exp(-\beta_1 c_i)}{1+\exp(-\beta_1 c_i)}, \tag{3}$$

$$c_i \triangleq \mathbf{e}_i^T (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i, \tag{4}$$

where $\beta_1 > 0$ determines the sharpness of the shape. When impulsive measurement noise does not appear, c_i is very small, and then, Taylor expansions provide the following approximation:

$$J_1(\widehat{\mathbf{w}}_i) \cong \mathbf{e}_i^T (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i, \tag{5}$$

using which this function works like the conventional APA. If impulsive measurement noise occurs, c_i becomes very large, which turns the cost function into $1/\beta_1$, *i.e.*, $J(\widehat{\mathbf{w}}_i) \cong 1/\beta_1$, and thus its derivative into zero, which plays an essential role in achieving the robustness against impulsive measurement noise.

Let us develop the adaptive gradient algorithm based on this cost function:

$$\widehat{\mathbf{w}}_{i+1} = \widehat{\mathbf{w}}_i - \mu_i \nabla_{\widehat{\mathbf{w}}_i} J_1(\widehat{\mathbf{w}}_i), \tag{6}$$

where μ_i is a step size at time i and $\nabla_{\widehat{\mathbf{w}}_i} J_1(\widehat{\mathbf{w}}_i)$ denotes the gradient of $J_1(\widehat{\mathbf{w}}_i)$ with respect to $\widehat{\mathbf{w}}_i$ as follows:

$$\nabla_{\widehat{\mathbf{w}}_i} J(\widehat{\mathbf{w}}_i) = -S_1(c_i) \mathbf{U}_i (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i, \tag{7}$$

$$S_1(c_i) \triangleq \frac{4\exp(-\beta_1 c_i)}{(1+\exp(-\beta_1 c_i))^2}. \tag{8}$$

The resulting algorithm can be summarized as

$$\widehat{\mathbf{w}}_{i+1} = \widehat{\mathbf{w}}_i + \mu_i S_1(c_i) \mathbf{U}_i (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i, \tag{9}$$

which exactly has the same form as the conventional APA with the step size μ_i, except for term $S_1(c_i)$. When c_i is small, *i.e.*, $S_1(c_i) \cong 1$, the adaptive algorithm performs like the conventional APA. A larger c_i leads to a smaller $S_1(c_i)$, which means the step-size scaler $S_1(c_i)$ performs the role of scaling the step size μ_i to eliminate possible malfunctions due to a large amount of impulsive measurement noise. Moreover, the step-size scaler can be easily applied to adaptive algorithms: it is just multiplied by each step size.

Remark 1. Since the step-size scaler $S_1(c_i)$ uses the exponential function, it need high computational complexity and it is difficult to apply in adaptive filtering applications. To reduce the computational complexity of $S_1(c_i)$, a simplified version of the step-size scaler was introduced in [8]. We modify the simplified version of the step-size scaler and propose a step-size scaler, which is suitable for application in the APA, *say* $S_2(c_i)$, defined as

$$S_2(c_i) \triangleq \frac{1}{1+\beta_2 c_i}, \tag{10}$$

which is derived from a cost function

$$J_2(\widehat{\mathbf{w}}_i) = \frac{1}{2\beta_2} \ln(1+\beta_2 c_i).$$

The resulting algorithm can be summarized as

$$\widehat{\mathbf{w}}_{i+1} = \widehat{\mathbf{w}}_i + \mu_i S_2(c_i) \mathbf{U}_i (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i.$$

Although this step-size scaler $S_2(c_i)$ is obtained experimental, it has the advantage of the computational complexity than $S_1(c_i)$ and improves the robustness against impulsive measurement noise by scaling the step size like $S_1(c_i)$.

3. The proposed APA

The conventional APA with a fixed step size μ is described as follows:

$$\widehat{\mathbf{w}}_{i+1} = \widehat{\mathbf{w}}_i + \mu \mathbf{U}_i (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{e}_i. \tag{11}$$

Then, the estimation error vector, $\tilde{\mathbf{w}}_i \triangleq \mathbf{w} - \widehat{\mathbf{w}}_i$, follows the recursion:

$$\tilde{\mathbf{w}}_{i+1} = \mathbf{F}_i \tilde{\mathbf{w}}_i - \mu \mathbf{U}_i (\mathbf{U}_i^T \mathbf{U}_i)^{-1} \mathbf{v}_i, \tag{12}$$