Signal Processing 152 (2018) 160–164

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

Bias-compensated normalized maximum correntropy criterion algorithm for system identification with noisy input

Wentao Ma[a,*], Dongqiao Zheng[a], Yuanhao Li[b], Zhiyu Zhang[a], Badong Chen[b]

[a] School of Automation and Information Engineering, Xi'an University of Technology, China
[b] School of Electronic and Information Engineering, Xi'an Jiaotong University, China

ARTICLE INFO

Article history:
Received 16 November 2017
Revised 7 May 2018
Accepted 28 May 2018
Available online 29 May 2018

2010 MSC:
00-01
99-00

Keywords:
Bias-compensated
Normalized maximum correntropy criterion (NMCC)
Noisy input
Impulsive output noise
System identification

ABSTRACT

This paper proposes a bias-compensated normalized maximum correntropy criterion (BCNMCC) algorithm charactered by its low steady-state misalignment for system identification with noisy input in an impulsive output noise environment. The normalized maximum correntropy criterion (NMCC) is derived from a correntropy based cost function, which is rather robust with respect to impulsive noises. To deal with the noisy input, we introduce a bias-compensated vector to the NMCC algorithm, and then an unbiasedness criterion and some reasonable assumptions are used to compute the bias-compensated vector. Taking advantage of the bias-compensated vector, the bias caused by the input noise can be effectively suppressed. System identification simulation results demonstrate that the proposed BCNMCC algorithm can outperform other related algorithms with noisy input especially in an impulsive output noise environment.

1. Introduction

Recently, bias-compensated adaptive filtering algorithms (BCAFAs) [1–5] based on the unbiasedness criterion (UC) are paid attention to in several signal processing applications, such as channel estimation, echo cancellation, and system identification in a noisy input case, that is, a bias-compensated vector is introduced to reduce the bias caused by the input noise. In particular, the bias-compensated normalized least mean square (BCNLMS) algorithm is popular due to its simplicity and effectiveness [1,2]. In [3], the bias-compensated affine projection algorithm (APA) was developed to reduce the performance degradation caused by highly correlated input. The bias-compensated normalized subband adaptive filter algorithm was proposed in [4], which has better performance and does not require knowledge of the input-output variance ratio. The bias-compensated normalized least mean fourth (NLMF) was presented in [5], which can offer a faster convergence rate and a lower steady-state misalignment. Furthermore, the BCNLMS with L_1-norm was proposed in [6] to address the noisy input problem in sparse system identification. At present in [7,8], the convergence analysis of the BCNLMS has been performed. All the above BCAFAs have been successfully utilized to solve the noisy input problem in different applications. However, they are sensitive to output noise with impulsive characters.

In order to improve the robustness with respect to output noise, some improved adaptive filtering algorithms (AFAs) have been proposed to eliminate the bad influence of the output noise in different literatures [9–13]. Particularly, different kinds of AFAs based on maximum correntropy criterion (MCC) were developed [14–16], such as sparse MCC [17], diffusion MCC [18], kernel MCC [19] and so on. Although the MCC based AFAs can improve the robustness in non-Gaussian signal processing, a noisy input case is however not considered in these solutions and thus they are sensitive to the scaling of input signals.

Considering the drawbacks of the existing BCAFAs and the MCC based AFAs, we take the advantages of the UC and robust property of the MCC to develop a novel bias-compensated normalized MCC (BCNMCC) algorithm in this study in order to eliminate the influence of the input noise and the impulsive output noise.

The rest of this paper is structured as follows. In Section 2, the NMCC algorithm is briefly reviewed. In Section 3, we develop the bias-compensated NMCC algorithm. In Section 4, simulation experiments are conducted to demonstrate the performance of the new method. Finally, the paper is concluded in Section 5.

* Corresponding author.
E-mail address: mwt7980@aliyun.com (W. Ma).

https://doi.org/10.1016/j.sigpro.2018.05.029

2. Review of the NMCC algorithm

For an adaptive filter under a common system identification (SI) framework, the desired signal is generally denoted by

$$d(i) = \mathbf{u}^T(i)\mathbf{w}^o + v(i) \tag{1}$$

where $\mathbf{w}^o = [w_1^o, w_2^o, \ldots w_L^o]^T$ denotes an unknown system parameter vector with L-tap to be estimated, and the perturbation signal $v(i)$ is the output noise at time index i. $\mathbf{u}(i) = [u_1(i), u_2(i), \ldots, u_L(i)]^T$ denotes the input vector. In [11], the update of the MCC based AFA is given by

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \exp(-\frac{e^2(i)}{2\sigma^2})e(i)\mathbf{u}(i) \tag{2}$$

where $\mathbf{w}(i) = [w_1(i), w_2(i), \ldots w_L(i)]^T$ denotes the tap-coefficients vector of an adaptive filter which is employed to find an estimate of $\mathbf{w}^o$ from the observed input-output data. μ is the step size and σ denotes the kernel bandwidth, which is a positive parameter that induces a trade-off between convergence speed and steady-state accuracy. $e(i) = d(i) - \mathbf{u}^T(i)\mathbf{w}(i)$ denotes the ith instantaneous error. Considering the MCC and the idea of the normalized least mean square, the normalized MCC (NMCC) updating equation can be represented as

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu \exp(-\frac{e^2(i)}{2\sigma^2})\frac{e(i)\mathbf{u}(i)}{\mathbf{u}^T(i)\mathbf{u}(i)+\varepsilon} \tag{3}$$

where $\varepsilon > 0$ is a small constant positive parameter, which is to prevent the denominator from being divided by zero, and it can provide a stable solution.

3. Bias-compensated NMCC

In this section, we focus on developing the bias-compensated NMCC algorithm based on UC and the NMCC algorithm in (3) for SI problem in a noisy input and output environment. Considering the input noise, we define the input vector as

$$\bar{\mathbf{u}}(i) = \mathbf{u}(i) + \boldsymbol{\eta}(i) \tag{4}$$

where $\boldsymbol{\eta}(i) = [\eta(i), \eta(i-1), \ldots \eta(i-L+1)]^T$ is the noise vector, and $\eta_l(i)(l \in [1, L])$ is with zero-mean Gaussian and variance δ^2. One can rewrite the filtered output error as

$$\begin{aligned}\bar{e}(i) &= d(i) - \bar{\mathbf{u}}^T(i)\mathbf{w}(i)\\ &= d(i) - (\mathbf{u}(i)+\boldsymbol{\eta}(i))^T\mathbf{w}(i)\\ &= \mathbf{u}(i)^T\tilde{\mathbf{w}}(i) + v(i) - \boldsymbol{\eta}(i)^T\mathbf{w}(i)\\ &= e_w(i) + v(i) - \boldsymbol{\eta}(i)^T\mathbf{w}(i)\end{aligned} \tag{5}$$

where $e_w(i) = \mathbf{u}^T(i)\tilde{\mathbf{w}}(i)$ is the priori error and the weight-error vector is denoted as $\tilde{\mathbf{w}}(i) = \mathbf{w}_o - \mathbf{w}(i)$. To compensate the bias caused by the input noise, we introduce a bias-compensation vector $\mathbf{B}(i)$ into (5), and replace $\mathbf{u}(i)$ and $e(i)$ with $\bar{\mathbf{u}}(i)$ and $\bar{e}(i)$ simultaneously. The Eq. (5) is improved as

$$\mathbf{w}(i+1) = \mathbf{w}(i) + \mu f(\bar{e}(i))\frac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon} + \mathbf{B}(i) \tag{6}$$

where $f(\bar{e}(i))$ is a non-linear function of the estimation error, which is defined as

$$f(\bar{e}(i)) = \exp\left(-\frac{\bar{e}^2(i)}{2\sigma^2}\right) = \exp\left(-\frac{(e_w(i)+v(i)-\boldsymbol{\eta}^T(i)\mathbf{w}(i))^2}{2\sigma^2}\right) \tag{7}$$

Then, combining (6) and the definition of the $\tilde{\mathbf{w}}(i)$, one can obtain

$$\tilde{\mathbf{w}}(i+1) = \tilde{\mathbf{w}}(i) - \mu f(\bar{e}(i))\frac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon} - \mathbf{B}(i) \tag{8}$$

Now, we employ the unbiasedness criterion [2] in (9) to get the bias-compensated vector.

$$E(\tilde{\mathbf{w}}(i+1)|\bar{\mathbf{u}}(i)) = \mathbf{0} \quad \text{whenever} \quad E(\tilde{\mathbf{w}}(i)|\bar{\mathbf{u}}(i)) = \mathbf{0} \tag{9}$$

Taking expectation on both sides of (8) with the given $\bar{\mathbf{u}}(i)$ and using criterion (9), one can obtain

$$\begin{aligned}&E[\tilde{\mathbf{w}}(i+1)|\bar{\mathbf{u}}(i)] =\\ &E[\tilde{\mathbf{w}}(i)|\bar{\mathbf{u}}(i)] - \mu E\Big[f(\bar{e}(i))\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big] - E[\mathbf{B}(i)|\bar{\mathbf{u}}(i)]\end{aligned} \tag{10}$$

According to (9) and (10), the following equation is obtained

$$E[\mathbf{B}(i)|\bar{\mathbf{u}}(i)] = -\mu E\left[f(\bar{e}(i))\frac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\right] \tag{11}$$

In order to calculate the gradient of the BCNMCC algorithm, the following commonly-used assumptions [20,21] are given:

Assumption 1. The signals $v(i)$, $\boldsymbol{\eta}(i)$, $\mathbf{u}(i)$ and $\tilde{\mathbf{w}}(i)$ are statistically independent of each other, and $\boldsymbol{\eta}(i)$ is of zero-mean.

Assumption 2. The non-linear function of the estimation error $f(v(i))$, $\boldsymbol{\eta}(i)$ and $\bar{e}(i)$ are statistically independent.

To simplify the following analysis, we take the Taylor expansion of $f(\bar{e}(i))$ with respect to $e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)$ around $v(i)$. Combining (5) one can obtain

$$\begin{aligned}&f(\bar{e}(i)) \approx f(v(i)) + f'(v(i))[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]\\ &+o\Big[[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]^2\Big]\end{aligned} \tag{12}$$

From (11), the following approximation can be obtained

$$\begin{aligned}&E\Big[f(\bar{e}(i))\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\\ &\approx E\Big[f(v(i))\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\\ &+E\Big[f'(v(i))[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\\ &+E\Big[o\Big[[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]^2\Big]\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\end{aligned} \tag{13}$$

In the steady-state, the priori error $e_w(i)$ converges to a small value which is ignorable with respect to the environmental noise when the step size is small [22]. Considering the Assumptions 1 and 2, the second term of Eq. (13) becomes

$$\begin{aligned}&E\Big[f'(v(i))[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\\ &\approx -E\Big[f'(v(i))\boldsymbol{\eta}^T(i)\mathbf{w}(i)\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]=0\end{aligned} \tag{14}$$

Similarly the third term of Eq. (13) is

$$E\left[o\Big[[e_w(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)]^2\Big]\frac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\right]=0 \tag{15}$$

Combining (13)–(15), and using Assumption 2, we have

$$\begin{aligned}&E\Big[f(\bar{e}(i))\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big] \approx E\Big[f(v(i))\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\\ &= E[f(v(i))|\bar{\mathbf{u}}(i)]E\Big[\tfrac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(i)\Big]\end{aligned} \tag{16}$$

Considering the fact that $\bar{e}(i) = e(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)$

$$E\left[\frac{\bar{e}(i)\bar{\mathbf{u}}(i)}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon}|\bar{\mathbf{u}}(n)\right] = \frac{E[\bar{e}(i)\bar{\mathbf{u}}(i)|\bar{\mathbf{u}}(i)]}{\bar{\mathbf{u}}^T(i)\bar{\mathbf{u}}(i)+\varepsilon} \tag{17}$$

$$\begin{aligned}&E[\bar{e}(i)\bar{\mathbf{u}}(i)|\bar{\mathbf{u}}(i)]\\ &= E\Big[[e(i) - \boldsymbol{\eta}^T(i)\mathbf{w}(i)][\mathbf{u}(i)+\boldsymbol{\eta}(i)]|\bar{\mathbf{u}}(i)\Big]\\ &= E[e(i)\bar{\mathbf{u}}(i)|\bar{\mathbf{u}}(i)] + E[e(i)\boldsymbol{\eta}(i)|\bar{\mathbf{u}}(i)]\\ &-E\Big[\boldsymbol{\eta}^T(i)\mathbf{w}(i)\mathbf{u}(i)|\bar{\mathbf{u}}(i)\Big] - E\Big[\boldsymbol{\eta}^T(i)\mathbf{w}(i)\boldsymbol{\eta}(i)|\bar{\mathbf{u}}(i)\Big]\end{aligned} \tag{18}$$