Signal Processing 143 (2018) 1–6

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

An adjusting-block based convex combination algorithm for identifying block-sparse system

Seung Hun Kim[a], Gyogwon Koo[a], Jae Jin Jeong[b], Sang Woo Kim[a,*]

[a] Department of Electrical Engineering, Pohang University of Science and Technology, Pohang, 37673, Republic of Korea
[b] Daegu Center, Defense Agency for Technology and Quality, Daegu, 42037, Republic of Korea

ARTICLE INFO

Article history:
Received 7 June 2017
Revised 17 August 2017
Accepted 19 August 2017

Keywords:
Adaptive filter
NLMS algorithm
Echo cancellation
Proportionate matrix
Zero-attracting penalty

ABSTRACT

A novel block wise convex combination algorithm with adjusting blocks is proposed for block-sparse system identification. The proposed algorithm unifies the complementary advantages of different block-induced algorithms, which are based on block proportionate matrix and block zero attracting penalty. A mixing parameter for block wise combination is designed as a block diagonal matrix. The mixing parameter is obtained using the conventional mixing parameter, which represents convergence state, and a block activeness indicator. The indicator for each block is derived from the l_ϵ^0-norm measure of the block. Moreover, a block adjustment algorithm is developed using the indicator to overcome the main disadvantage of block-induced algorithms, i.e., the dependency on cluster location. The simulations for system identification are performed on several block-sparse systems including systems with single cluster and double clusters. The simulation results show that the proposed algorithm not only combines the different block-induced algorithms effectively but also improves the performance via the block adjustment algorithm.

1. Introduction

Adaptive filtering algorithms have been researched for many applications such as acoustic and network echo cancellation (EC) and active noise control [1]. The normalized least-mean-square (NLMS) algorithm is the most popular adaptive filter owing to its wide applicability and robustness against background measurement noise. However, when a system has a long number of taps and sparse distribution, e.g., an echo channel, the NLMS algorithm shows an ineffective convergence. The sparse system indicates that its impulse response incorporates mainly zero or near-zero coefficients of minor significance, and it can be divided into two types: general-sparse and block-sparse systems.

In order to identify the sparse system, the adaptive filters based on a proportionate matrix and a zero-attracting penalty have been researched [2–10]. The proportionate NLMS (PNLMS) algorithm [2] and its several modifications [3–6] accelerate the convergence rate of the active tap by multiplying the diagonal matrix, which is proportionate to the tap magnitude, with the adaptation term. Owing to trade-off, the PNLMS algorithm exhibits a larger steady-state misalignment compared to the NLMS algorithm.

* Corresponding author.
E-mail addresses: pisics@postech.ac.kr (S.H. Kim), ggkoo99@postech.ac.kr (G. Koo), j_jeong@dtaq.re.kr (J.J. Jeong), swkim@postech.edu (S.W. Kim).

http://dx.doi.org/10.1016/j.sigpro.2017.08.014
0165-1684/

On the other hand, zero-attracting (ZA) algorithms focus on the inactive taps rather than the active taps [7–10]. The ZA algorithms attract the near-zero coefficients to the zero by applying penalty terms to their cost functions. Several sparsity measures including l^0-norm [7], l^1-norm [8], correntropy [9] and l^p-norm [10] have been suggested as penalty functions. With the relevant selection of the penalty step size, the ZA algorithms can have lower steady-state misalignment than the NLMS algorithm owing to the step-size diminishing effect on the inactive taps. Moreover, the cost functions of PNLMS and ZA algorithms have been combined [11]; however, such a combination cannot address the problem of the large steady-state misalignments of the active taps.

Recently, block-induced algorithms [12–14] have been introduced to further enhance the performance when the system is block sparse. In applications such as the typical acoustic EC and satellite-linked communication application, the systems have one or more clusters and near-zero coefficients elsewhere. The research in [12] has obtained a block-proportionate NLMS (BP-NLMS) algorithm by applying a block diagonal matrix, whose component blocks are proportional to block tap magnitudes, to the adaptation term. Other studies [13,14] have derived block-zero-attracting (BZA) algorithms by substituting previous penalties into block-based penalties. Similar to the PNLMS and ZA algorithm, the BP-NLMS algorithm accelerates convergence rate for cluster regions, and the BZA algorithm improves the steady-state misalignment for non-cluster regions. However, all these block-induced algorithms suffer from unavoidable dependency on the cluster location which is usually unknown in block-sparse systems [12].

In this paper, a novel block wise combination scheme is proposed to achieve following features: (1) combine the complementary advantages of the BP-NLMS and BZA-NLMS algorithms, and (2) overcome the dependency of block-induced algorithms on the cluster location. A block activeness indicator is designed to explicitly identify the active/inactive regions using the l_ϵ^0-norm measure [15]. From dual mixing parameters consist of the conventional mixing parameter and the indicator, the block wise combination of BP-NLMS and BZA-NLMS algorithms is achieved according to the block activeness. Further, a block adjustment algorithm is induced from the indicator, which can exactly fit blocks to the cluster location. Therefore, different from previous scheme [16] which updates the block-based mixing parameters individually, the proposed scheme updates only the single mixing parameter and exactly fits blocks to the cluster location. Simulation results show that the proposed algorithm effectively combines the advantages of the component filters in both cluster and non-cluster regions. Moreover, it is shown that the proposed algorithm can further improve the performance using the block adjustment algorithm.

2. Block sparsity induced adaptive filters

2.1. Block proportionate NLMS algorithm

The typical weight update equation for the PNLMS algorithm [2] is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \frac{\mathbf{G}_i \mathbf{u}_i e(i)}{\mathbf{u}_i^T \mathbf{G}_i \mathbf{u}_i + \delta_P} \tag{1}$$

where i denotes the number of iterations, μ is a step size, δ_P is a regularization parameter, $\hat{\mathbf{w}}_i = [\hat{w}_1(i) \cdots \hat{w}_M(i)]^T$ is a weight coefficient vector, $\mathbf{u}_i = [u(i)\, u(i-1) \cdots u(i-M+1)]^T$ is a vector consisting of input sequence, and $\mathbf{G}_i$ is a diagonal matrix, whose elements are proportional to the tap magnitude $|\hat{w}_j(i)|$. Further, M is the tap length, and $e(i)$ is an output error.

A block-induced $\mathbf{G}_i$ has been introduced in [12] to enhance the suitability of the PNLMS algorithm for block sparse system. The $\mathbf{G}_i$ is defined as $\mathrm{diag}(g_1(i)\mathbf{1}_P, g_2(i)\mathbf{I}_P, \cdots g_N(i)\mathbf{I}_P)$ where N is the total number of blocks, P is the length of each block, i.e., $M = N \times P$, and $\mathbf{I}_P$ is an identity matrix of size P. Subsequently, for $t = 1, 2, \cdots, N$, $g_t(i)$ is given as $g_t(i) = \frac{\gamma_t(i)}{\frac{1}{N}\sum_{k=1}^N \gamma_k(i)}$ with $\gamma_t(i) = \max\left\{\rho\zeta(i), \|\hat{\mathbf{w}}_{i,[t]}\|_2\right\}$ and $\zeta(i) = \max\left\{q, \|\hat{\mathbf{w}}_{i,[1]}\|_2, \cdots, \|\hat{\mathbf{w}}_{i,[N]}\|_2\right\}$. Here, $\hat{\mathbf{w}}_{i,[t]}$ is the tth coefficient block, i.e., $\hat{w}_j(i)$s with $j \in [(t-1)P+1, tP]$, q and ρ are parameters to circumvent possible stalling problems [2], and $\|\cdot\|_2$ means the Euclidean norm.

2.2. Block zero-attracting algorithm

The BZA algorithms are characterized by the selection of the norm order for the penalty function. Block induced $l^{2,0}$- [13] and $l^{2,1}$-norms [14] have been introduced as penalty functions to handle the block-sparse system. Among them, the BZA algorithm based on the $l^{2,1}$-norm penalty [14] is used here because its NLMS version was already studied. We tried to combine the NLMS-type algorithms which are known to be more robust against input sequence characteristics [1]. The block weight update equation for [14] is derived from the cost function $J(\hat{\mathbf{w}}_i) = \frac{1}{2}e^2(i) + \kappa_i f_i(\hat{\mathbf{w}}_i)$ where κ is a penalty step size and $f_i(\hat{\mathbf{w}}_i)_{[t]} = \sum_{t=1}^N \frac{\|\hat{\mathbf{w}}_{i,[t]}\|_2}{\|\hat{\mathbf{w}}_{i,[t]}\|_2 + \delta_W}$. From this, the equation is obtained as

$$\hat{\mathbf{w}}_{i+1,[t]} = \hat{\mathbf{w}}_{i,[t]} + \mu \frac{\mathbf{u}_{i,[t]} e(i)}{\|\mathbf{u}_i\|_2^2 + \delta_{ZA}} - \kappa^* \frac{\hat{\mathbf{w}}_{i,[t]}}{(\|\hat{\mathbf{w}}_{i,[t]}\|_2 + \delta_w)^2}, \tag{2}$$

where κ^* is decided as same as the ρ^* in [14], δ_{ZA} and δ_w are regularization parameters. Such κ^* can ensure the dominance of BZA-NLMS algorithm compared to the NLMS algorithm under a few assumptions. Note that the BZA-NLMS algorithm operates as the NLMS algorithm in the cluster regions through the denominator of the penalty function in (2).

3. Block wise convex combination

Both the BP-NLMS and BZA-NLMS algorithms are designed for the block-sparse system, but they target different regions of the system; the BP-NLMS aims at accelerating the convergence rate in the cluster regions, and the BZA-NLMS focuses on the shrinkage of the non-cluster regions. Therefore, a convex combination of these two algorithms can further improve the performance in block-sparse systems.

3.1. Mixing parameter

A block diagonal mixing parameter, which is $\mathbf{\Lambda}_i = \lambda_1(i)\mathbf{\Lambda}_{2i}$, is proposed to design the block wise convex combination. Accordingly, for $1 \le t \le N$, the tth blocks of component filters are combined as

$$\hat{\mathbf{w}}_{i,[t]} = \lambda_1(i)\mathbf{\Lambda}_{2i,[t]}\hat{\mathbf{w}}_{1i,[t]} + \left(\mathbf{I}_P - \lambda_1(i)\mathbf{\Lambda}_{2i,[t]}\right)\hat{\mathbf{w}}_{2i,[t]}, \tag{3}$$

where $\hat{\mathbf{w}}_i$ is the combined filter coefficient, $\hat{\mathbf{w}}_{1i}$ and $\hat{\mathbf{w}}_{2i}$ are the first and second component filter coefficients, respectively. In addition, $\lambda_1(i)$ is the conventional mixing parameter that indicates the convergence state, and $\mathbf{\Lambda}_{2i,[t]}$ is a diagonal matrix consisting of 1 and 0, which is related to the activeness of the tth block. Same as the previous convex combination schemes, $\lambda_1(i)$ is defined as $\lambda_1(i) = 1/(1+\exp(-a(i)))$ [16]. Subsequently, $a(i)$ of the sigmoid function is updated using the gradient of the squared error as follows:

$$\begin{aligned} a(i+1) &= a(i) - \frac{\mu_a}{2}\frac{1}{\hat{p}(i)}\frac{\partial e^2(i)}{\partial a(i)} \\ &= a(i) - \mu_a \frac{e(i)}{\hat{p}(i)}\lambda_1(i)(1-\lambda_1(i))\mathbf{u}_i^T\mathbf{\Lambda}_{2i}(\hat{\mathbf{w}}_{2i} - \hat{\mathbf{w}}_{1i}), \end{aligned} \tag{4}$$

where μ_a is the step size for $a(i)$ adaptation, $\mathbf{\Lambda}_{2i} = \mathrm{diag}(\mathbf{\Lambda}_{2i,[1]}, \cdots, \mathbf{\Lambda}_{2i,[N]})$, and $\hat{p}(i)$ is the normalization term [16] defined as

$$\hat{p}(i) = \alpha\hat{p}(i-1) + (1-\alpha)(\mathbf{u}_i^T\mathbf{\Lambda}_{2i}(\hat{\mathbf{w}}_{2i} - \hat{\mathbf{w}}_{1i}))^2. \tag{5}$$

Here, α, which is slightly less than 1, is a parameter for the moving-average.

3.2. Block adjustment algorithm

The block-induced algorithms [12–14] have dependency on the cluster location. For example, the block-induced algorithms with $P = 16$ cannot totally cover the regions of $\mathbf{w}_{[2]}^o$ and $\mathbf{w}_{[4]}^o$ if the impulse response of an unknown system has a dispersive region in the range of (25, 56) (Fig. 1). This is an unavoidable problem because the exact location of the cluster is usually unknown [12] and the algorithms are based on equi-partitioned blocks. Therefore, to relieve the dependency, a block adjustment (BA) algorithm is proposed in this subsection.

To derive the BA algorithm, a block activeness indicator $s_t(i)$ is introduced using the $l\epsilon^0$-norm measure [15] such as

$$s_t(i) = 1 - \frac{l^0_{\epsilon_t(i)}}{\text{length of } t^{\text{th}} \text{ block}} \tag{6}$$