A Variable Step-size Diffusion Affine Projection Algorithm Using a Novel Combiner Requiring No Additional Communication

Min Seok Seo^a, Seung Hun Kim^a, Gyogwon Koo^a, and Sang Woo Kim^{a,*}

^aDepartment of Electrical Engineering, POSTECH, 77, Cheongam-Ro, Nam-Gu, 37673, Pohang, Gyeongbuk, Republic of Korea

*Corresponding Author: swkim@postech.edu

Abstract

We propose a combination method for a diffusion affine projection algorithm. There are two steps to estimate unknown optimal weight in Adapt-then-combine strategy of diffusion adaptive filtering. In an adaptation step, variable step-size affine projection algorithm works to get an improvement of convergence rate and small steady-state error. In a diffusion step, the proposed combination method is applied to achieve small steady-state errors without additional communicational complexity. The cooperative strategy using information across the network achieves fast convergence rate, small steady-state error and robustness against outliers compared to a non-cooperative way. The proposed algorithm is motivated from the optimal combination rule and it has a physical meaning of weighting on how close the estimates reach in steady state.

Keywords: distributed network, adaptive network, diffusion adaptive filtering, combiner, adaptive filter, communicational complexity, optimal step size

1. Introduction

Distributed networks have been studied as they reduce power requirement and amount of communications between nodes of the network compared to centralized networks.⁽¹⁾ The network is interested in estimating same parameter, which is an optimal weight, through a connection of the nodes linked to each other and a cooperation by transmitted information across the network. Diffusion strategies have been suggested as one of the solution for the problem⁽²⁾. In a diffusion adaptive network, the nodes adaptively process the problem updating an estimate for an optimal weight where the update is based on adaptive filter and diffuses the updated

estimates to their linked neighborhood to cooperate by combining them.

Diffusion Least-mean-squares algorithm was introduced to estimate the optimal weight using input and output data⁽²⁾ since the algorithm is representatively method. When the given input is correlated, it causes a decrease of the convergence rate. For this reason, diffusion affine projection algorithm (APA) was suggested since the algorithm makes the given colored input whitened so that we can obtain desired convergence rate⁽³⁾. Moreover, variable step-size theories were studied^(4,5) as a solution of a general problem with a trade-off in step size. That is, the variable step-size theories provided the effective result in fast convergence rate and small steady-state error.

After updating the estimates, the nodes diffuse the estimates and combine them with a specific combiner. Several simple combiners such as uniform, metropolis⁽⁶⁾, and relative degree methods⁽⁷⁾ have been suggested. They consider the number of the linked nodes, but not other properties. As the convergence rate and steady-state error depend on the various noise variances of the individual nodes, some papers have suggested adaptive combination methods considering the different variances⁽⁸⁻¹⁰⁾, but they still carry out suboptimal performance with different step sizes or highly correlated input signals. Recently, combination methods based on mean-square deviation (MSD) of each node were introduced^(4,11-12) and it improved the convergence rate and steady-state error compared to other combiners. However, it requested additional communicational complexity which comes to the problem with channel impairments such as path loss, multi-path fading and shadowing⁽¹³⁾.

We propose a combiner to achieve fast convergence rate and small steady-state error without additional

communicational complexity compared to the combination method based on MSD. The proposed algorithm is motivated from the optimal combination rule. Furthermore, it considers update quantities of the individual nodes compared to the conventional algorithm considering the distance to the optimal weight as well as different step sizes of each node.

2. Diffusion Variable Step-size APA

A typical structure and process of diffusion adaptive filtering are shown in Fig. 1. Several nodes of the network are connected to each other where the line between the nodes means a connection transmitting data in Fig. 1(a). Node k has an access to a vector input data $\mathbf{u}_k(i) = [u_k(i), u_k(i-1), \dots, u_k(i-M+1)]$ and a scalar desired output data $d_k(i)$ at every time instant i . They are related via

$$d_k(i) = \mathbf{u}_{k,i}^T \mathbf{w}^o + v_k(i) \quad (1)$$

where $\mathbf{w}^o$ is an unknown column vector with M dimension we hope to estimate and $v_k(i)$ is a scalar measurement noise. $\psi_{n,i}$ is an intermediate estimate of node n . N_k is a

(a) Structure of the network

(b) Process of the network.

Fig. 1 Diffusion adaptive filtering.

neighborhood that a group of nodes linked to node k . In Fig. 1 (b), adapt-then-combine(ATC) diffusion strategy is applied to the network as it outperforms combine-then-adapt(CTA) diffusion strategy⁽¹⁴⁾. ATC strategy consists of two stages: adaptation stage and combination stage.

2.1 APA in the Adaptation stage

In ATC diffusion strategy, a combined estimate $\mathbf{w}_k$ is updated to an intermediate estimate ψ_k using the vector input data and the scalar desired output $\{\mathbf{u}_{k,i}, d_k(i)\}$ at each node before the combination stage. Then, the intermediate estimates are transmitted to their linked neighbor nodes. The conventional APA is applied to the distributed network updating equation^(3,4). Diffusion APA is followed as

$$\psi_{k,i} = \mathbf{w}_{k,i-1} + \mu_k \mathbf{U}_{k,i} (\mathbf{U}_{k,i}^T \mathbf{U}_{k,i})^{-1} \mathbf{e}_{k,i}, \quad (2)$$

where μ_k is step size of node k . $\mathbf{U}_{k,i}$ is a matrix of input data, and $\mathbf{e}_{k,i}$ is a *priori* output error of node k at iteration i defined as

$$\begin{aligned} \mathbf{U}_{k,i} &= [\mathbf{u}_{k,i} \ \mathbf{u}_{k,i-1} \ \dots \ \mathbf{u}_{k,i-P+1}], \\ \mathbf{e}_{k,i} &= \mathbf{d}_{k,i} - \mathbf{U}_{k,i}^T \mathbf{w}_{k,i-1}, \\ \mathbf{d}_{k,i} &= \mathbf{U}_{k,i}^T \mathbf{w}^o + \mathbf{v}_{k,i} \end{aligned} \quad (3)$$

where P is the number of the input vectors, and $\mathbf{v}_{k,i}$ is measurement noise vector defined as $\mathbf{v}_{k,i} = [v_k(i), v_k(i-1), \dots, v_k(i-K+1)]$.

2.2 Estimation of MSD

Individual nodes update the estimates with their different step sizes as the nodes have different estimation error.⁽⁴⁾ MSD of the node k should be derived to find their optimal step size to accurately estimate the optimal weight. Weight-error vectors of the intermediate estimate and the combined estimate at node k are defined as $\tilde{\psi}_k = \mathbf{w}^o - \psi_k$ and $\tilde{\mathbf{w}}_k = \mathbf{w}^o - \mathbf{w}_k$ respectively. The update equation can be rewritten as follows respect to (2)

$$\tilde{\psi}_{k,i} = \tilde{\mathbf{w}}_{k,i} - \mu_k \mathbf{U}_{k,i} (\mathbf{U}_{k,i}^T \mathbf{U}_{k,i})^{-1} \mathbf{e}_{k,i}. \quad (4)$$

At every iteration, the MSD of the node k is derived by squaring and taking expectation to the both sides in the rewritten update equation as

$$\begin{aligned} \bar{\mathbf{P}}_{k,i} &= \hat{\mathbf{Q}}_{k,i-1} - 2\mu_k(i) \mathbf{U}_{k,i} (\mathbf{U}_{k,i}^T \mathbf{U}_{k,i})^{-1} \mathbf{U}_{k,i}^T \hat{\mathbf{Q}}_{k,i-1} \\ &\quad + \mu_k^2(i) \mathbf{U}_{k,i} (\mathbf{U}_{k,i}^T \mathbf{U}_{k,i})^{-1} \mathbf{E}[\mathbf{e}_{k,i} \mathbf{e}_{k,i}^T] (\mathbf{U}_{k,i}^T \mathbf{U}_{k,i})^{-1} \mathbf{U}_{k,i}^T \end{aligned} \quad (5)$$

where $\bar{\mathbf{P}}_{k,i} = \mathbf{E}[\tilde{\psi}_{k,i} \tilde{\psi}_{k,i}^T | \mathbf{U}_{k,i}]$, and $\hat{\mathbf{Q}}_{k,i-1} = \mathbf{E}[\tilde{\mathbf{w}}_{k,i} \tilde{\mathbf{w}}_{k,i}^T | \mathbf{U}_{k,i}]$. Taking trace and expectation, the MSD update equation is expressed as

$$\bar{p}_k(i) = \hat{p}_k(i-1) - 2\mu_k(i)E\left\{\text{Tr}\left[\mathbf{U}_{k,i}(\mathbf{U}_{k,i}^T\mathbf{U}_{k,i})^{-1}\right.\right. \\ \left.\left.\times \mathbf{U}_{k,i}^T\hat{p}_k(i-1)\right]\right\} + \mu_k^2(i)\text{Tr}[\mathbf{R}_{e_{k,i}} \cdot \mathbf{R}_{u_{k,i}}^{-1}], \quad (6)$$

where $\bar{p}_k(i) \triangleq E[\text{Tr}(\bar{\mathbf{P}}_{k,i})]$, $\hat{p}_k(i) \triangleq E[\text{Tr}(\hat{\mathbf{Q}}_{k,i})]$, $\mathbf{R}_{e_{k,i}} \triangleq E[\mathbf{e}_{k,i}\mathbf{e}_{k,i}^T]$, and $\mathbf{R}_{u_{k,i}} \triangleq E[\mathbf{U}_{k,i}^T\mathbf{U}_{k,i}]$. It follows as

$$\bar{p}_k(i) = \hat{p}_k(i-1) - 2\mu_k(i)\frac{K}{M}\hat{p}_k(i-1) + \\ \mu_k^2(i)\text{Tr}[\mathbf{R}_{e_{k,i}}\mathbf{R}_{u_{k,i}}^{-1}] \triangleq \hat{p}_k(i-1) + f_2(\mu_k(i)). \quad (7)$$

We can approximate $\text{Tr}[\mathbf{U}_{k,i}(\mathbf{U}_{k,i}^T\mathbf{U}_{k,i})^{-1}\mathbf{U}_{k,i}^T\hat{\mathbf{p}}_k(i-1)]$ to

$$\frac{P}{M}\hat{p}_k(i-1) \text{ as it is demonstrated}^{(15)}.$$

2.3 Optimal Step Size

Optimal step size at every iteration should be determined to minimize the MSD. Reduction of $f_2(\mu_k(i))$ plays the role decreasing the expected MSD as iteration increases. We differentiate eq. (7) as

$$\frac{\partial f_2(\mu_k(i))}{\partial \mu_k(i)} = \frac{-2K}{M}\hat{p}_k(i-1) + 2\mu_k(i)\text{Tr}(\mathbf{R}_{e_{k,i}} \cdot \mathbf{R}_{u_{k,i}}^{-1}). \quad (8)$$

where $\mathbf{R}_{e_{k,i}}$ and $\mathbf{R}_{u_{k,i}}$ are the estimates of output errors and input data respectively as

$$\mathbf{R}_{e_{k,i}} = \alpha\mathbf{R}_{e_{k,i-1}} + (1-\alpha)\mathbf{e}_{k,i}\mathbf{e}_{k,i}^T, \\ \mathbf{R}_{u_{k,i}} = \alpha\mathbf{R}_{u_{k,i-1}} + (1-\alpha)\mathbf{U}_{k,i}^T\mathbf{U}_{k,i}. \quad (10)$$

where α is forgetting factor. Reduction of MSD is maximized when it sets the equation equal to 0. Therefore, the optimal step size is obtained as

$$\mu_k^*(i) = \frac{P\cdot\hat{p}_k(i-1)}{M\cdot\text{Tr}(\mathbf{R}_{e_{k,i}}\mathbf{R}_{u_{k,i}}^{-1})}, \quad (9)$$

For the stability of the algorithm, we restrict the value of step size

$$\mu_k^*(i) = \min\left(1, \frac{P\cdot\hat{p}_k(i-1)}{M\cdot\text{Tr}(\mathbf{R}_{e_{k,i}}\mathbf{R}_{u_{k,i}}^{-1})}\right). \quad (11)$$

3. Combination stage

The intermediate estimates updated in the adaptation step are diffused to the neighbor nodes from their own nodes. Then, the diffused estimates are combined with a specific combination method through a convex combination as

$$\mathbf{w}_{k,i} = \sum_{l \in N_k} a_{lk}\boldsymbol{\psi}_{l,i}, \quad (12)$$

where a_{lk} means a combination weight of node l at node k . In diffusion adaptive filter, such cooperation process improves the convergence rate, the robustness against outliers and decreases the steady-state errors while it requires the communication among the nodes of the network.

Recently, combination methods based on MSD were

introduced. It uses estimated MSD transmitted in combination stage. It improves the convergence rate and reduces steady-state error. However, reduction of communicational complexity is required in practice because the transmission power is limited and data can be distorted due to the channel impairments.

3.1 Optimal combination rule

Optimal combination rule was suggested to fuse the information from the neighbor⁽¹⁰⁾. The rule considers the variation in noise profile across the nodes, while other structurally defined combination methods such as uniform rule, metropolis rule ignore the variance so that can lead to the performance degradation. Considering ATC strategy, the corresponding upper bound of the network MSD of N nodes is expressed⁽¹⁶⁾ as

$$\text{Tr}(\mathcal{A}^T\mathcal{M}\mathcal{S}\mathcal{M}\mathcal{A}), \quad (13)$$

$$\mathcal{A} = A \otimes I_M,$$

$$\mathcal{M} = \text{diag}\{\mu_1 I_M, \mu_2 I_M, \dots, \mu_N I_M\}, \quad (14)$$

$$\mathcal{S} = \text{diag}\{\sigma_{v_1}^2 \mathbf{R}_{u_1}, \sigma_{v_2}^2 \mathbf{R}_{u_2}, \dots, \sigma_{v_N}^2 \mathbf{R}_{u_N}\},$$

where A consists of $\{a_{lk}, l = 1, 2, \dots, N\}$, I_M is a M -dimensional identity matrix, and $\otimes$ denotes Kronecker product operation. $\sigma_{v_k}^2$ is a noise variance of node k , and $\mathbf{R}_{u_k}$ is a input covariance of node k . To cope with the combination weights, we have to consider the problem minimizing the network MSD. The upper bound of network MSD is rewritten in terms of combination weight as

$$\min_{\{a_{lk}\}_{l=1}^N} \sum_{l=1}^N \mu_l^2 a_{lk}^2 \sigma_{v_l}^2 \text{Tr}(\mathbf{R}_{u_l}), \quad k = 1, \dots, N \quad (15)$$

Subjects to $a_{lk} \geq 0, \sum_{l=1}^N a_{lk} = 1, a_{lk} = 0$ if $l \notin N_k$. As the network MSD should be minimized, we refer to this combination rule as the relative variance combination rule⁽⁸⁾ and the solution is given by:

$$a_{lk} = \begin{cases} \frac{\gamma_{l,k}^{-2}}{\sum_{j \in N_k} \gamma_{j,k}^{-2}}, & \text{if } l \in N_k \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where $\gamma_{l,k}^2$ is defined as $\mu_l^2 \sigma_{v_l}^2 \text{Tr}(\mathbf{R}_{u_l})$, which is a variance product. It is, however, not available to evaluate the combination weights defined in terms of the step sizes μ_l , the noise variances $\sigma_{v_l}^2$ and the input covariance matrices $\mathbf{R}_{u_l}$ across the network as the nodes do not have an access to the information from the nodes.

The nodes across the network are required adaptively to implement the combination of the intermediate estimates. In other words, the variance product should be evaluated using the instantaneous data $\{\mathbf{u}_l, d(i)\}$ which each node is able to access, and transmitted intermediate estimates $\boldsymbol{\psi}_l$

from the neighbor nodes. Under the condition that the nodes have small step size or reach steady state, the update equation of the intermediate estimate is approximated by (1) as

$$\boldsymbol{\psi}_{k,i} \approx \mathbf{w}^o + \mu_k \mathbf{u}_{k,i}^T v_k(i). \quad (17)$$

We obtain a variance equation from the rewritten (17) as follows

$$E\|\boldsymbol{\psi}_{k,i} - \mathbf{w}^o\|^2 \approx \mu^2 \sigma_{vk}^2 \text{Tr}(\mathbf{R}_u). \quad (18)$$

It means that we can regard this estimated equation as the variance product $\gamma_{l,k}^2$. Using the instantaneous approximation that $\mathbf{w}^o \approx \mathbf{w}_k(i-1)$, (18) can be expressed as

$$E\|\boldsymbol{\psi}_{k,i} - \mathbf{w}^o\|^2 \approx \|\boldsymbol{\psi}_{k,i} - \mathbf{w}_{k,i-1}\|^2. \quad (19)$$

Then, $\gamma_{l,k}^2$ is estimated at every iteration in a way of moving average as

$$\gamma_{l,k}^2(i) = \alpha \cdot \gamma_{l,k}^2(i-1) + (1-\alpha) \cdot \|\boldsymbol{\psi}_{l,i} - \mathbf{w}_{k,i-1}\|^2 \quad (20)$$

In this way, the distance between the intermediate estimate $\boldsymbol{\psi}_l$ and optimal weight $\mathbf{w}^o$ is used to estimate the variance product.

However, in the variable step size method, all nodes do not use same step size. Even though it approaches steady-state, the different step sizes of individual nodes are not sufficiently small as they continuously change to reduce the estimated MSD. There is, therefore, the performance degradation in the convergence rate and steady-state error.

3.2 Proposed algorithm

The proposed algorithm is motivated from the optimal combination rule and it is appropriate to variable step-size algorithms. This accurately explicit the noise variance of each node. In the general minimization problem, the estimate is updated in a way of minimizing $\|\mathbf{w}_i - \mathbf{w}_{i-1}\|^2$ as iteration increases. The proposed rule uses the update quantity that is difference between the intermediate estimates at iteration i and $i-1$ as

$$E\|\boldsymbol{\psi}_{k,i} - \mathbf{w}^o\|^2 \approx \|\boldsymbol{\psi}_{k,i} - \boldsymbol{\psi}_{k,i-1}\|^2 \quad (21)$$

Also, a variance product of the proposed rule is estimated as

$$\beta_{l,k}^2(i) = \alpha \cdot \beta_{l,k}^2(i-1) + (1-\alpha) \cdot \|\boldsymbol{\psi}_{l,i} - \boldsymbol{\psi}_{l,i-1}\|^2 \quad (22)$$

In this way, the combination weight of each node is evaluated in terms of only the values of the corresponding nodes. It means that there is no requirement for additional communication. The proposed combination weight is following

$$a_{l,k} = \begin{cases} \frac{\beta_{l,k}^{-2}}{\sum_{j \in N_k} \beta_{j,k}^{-2}}, & \text{if } l \in N_k \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

4. Simulation

In this section, we simulated the mean square performance in the scenarios of the diffusion adaptive filtering to compare the proposed algorithm with other combiners. The normalized network MSD curves of 12 nodes were simulated in the given both topologies in Fig. 2^(4,11). In the simulations, an unknown impulse response was generated randomly with tap length $M=16$, and the inputs, at all nodes except the nodes $\{3, 6, 12\}$, were zero-mean white Gaussian random sequence filtered through a first-order autoregressive model with the pole located at 0.9. The number of input vector P was set to 4. The SNRs of the noise parameters at each node were generated as shown in Fig. 3 where $\text{SNR} \triangleq 10 \log_{10}(E[y(i)^2]/E[v(i)^2])$ with $y(i) \triangleq \mathbf{u}^T(i)\mathbf{w}^o$. In addition, the forgetting factor α was

(a) Structure of the network⁽⁴⁾

(b) Structure of the network⁽¹¹⁾

Fig. 2 Network topology for diffusion.

Table 1. Computational and communicational complexity

combiner	Computational complexity	Communicational Complexity
Uniform	1	0
Metropolis	1	n_k
Relative-degree	1	n_k
Optimal Rule	$M + n_k + 4$	0
Method based on MSD	n_k	n_k
proposed	$M + n_k + 4$	0

set to $1 - \frac{P}{N \cdot M}^{(4)}$ with the total number of nodes $N = 12$.

The optimal combination rule⁽¹⁰⁾, uniform rule, and combination method based MSD⁽⁴⁾ were compared with the proposed algorithm in the same simulation environment. The uniform rule is as follows

$$a_{l,k} = \begin{cases} \frac{1}{n_k}, & \text{if } l \in N_k \\ 0, & \text{otherwise} \end{cases} \quad (24)$$

where n_k denotes the number of N_k . Also, the combination method based on MSD⁽⁴⁾ is as follows

$$a_{l,k} = \begin{cases} \frac{c_{lk} \cdot \bar{P}_l^{-1}(i)}{\sum_{j \in N_k} c_{jk} \cdot \bar{P}_j^{-1}(i)}, & \text{if } l \in N_k \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

where c_{lk} is the metropolis rule which is one of the representative combiner as follows

$$c_{lk} = \begin{cases} \frac{1}{\max\{n_k, n_l\}}, & \text{if } l \neq N_k \\ 1 - \sum_{j \in N_k, j \neq k} c_{jk}, & \text{if } l = N_k \end{cases} \quad (26)$$

In the adaptation stage, as the intermediate estimate is transmitted, $M n_k$ communicational complexity is generally required. In the combination stage, computational and the communicational complexity of combination methods are shown in Table 1. The normalized MSD was given by $\frac{\|\tilde{\mathbf{w}}_l\|^2}{\|\mathbf{w}^o\|^2}$ and the ensemble-average value of 20 independent trials were obtained in Fig. 4. The results showed that the proposed algorithm achieved similar performance of convergence rate to the method based on MSD although the proposed algorithm does not require additional communication in Fig. 4.

5. Conclusion

In this paper, a combination weight based on the update quantity is applied to a diffusion adaptive network with

Fig. 3. SNR at each node.

(a) Network MSD over Fig. 2. (a)-topology network

(b) Network MSD over Fig. 2. (b)-topology network

Fig. 4. Simulation results

variable step-size affine projection algorithm to work without additional data in communication. As a variable step-size strategy is applied to the adaptation step, it provides fast convergence rate and small steady-state error. As a result, individual nodes have different optimal step sizes. To reflect the properties of each node, the proposed method appropriately works since it considers the update quantity. Compared to other combiners, it showed faster convergence rate and smaller steady-state error as well as it does not require additional communication.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2015R1D1A1A09059377).

References

- (1) D. Estrin, L. Girod, G. Pottie, and M. Srivastava : "Instrumenting the world with wireless sensor networks", *Acoustics, Speech, and Signal Processing*, Vol. 4, No. 1, pp. 2033-2036, 2001
- (2) Cassio G. Lope and A.H. Sayed : "Diffusion Least-Mean Squares Over Adaptive Networks: Formulation and Performance Analysis", *IEEE Transactions on Signal Processing*, Vol. 56, No. 7, pp. 3122-3136, 2008
- (3) Mohammad Shams Esfand Abadi, Zahra Saffari : "Distributed estimation over an adaptive diffusion network based on the family of affine projection algorithms", *Telecommunications (IST)*, pp. 607-611, 2012
- (4) J. W. Yoo, I. S. Song, J. W. Shin, and P. G. Park : "A variable step-size diffusion affine projection algorithm", *International Journal of Communication Systems*, Vol. 29, No. 5, pp. 1012-1025, 2016
- (5) Jingen Ni, and Jian Yang : "Variable step-size diffusion least mean fourth algorithm for distributed estimation", *Signal Processing*, Vol. 122, pp. 93-97, 2016
- (6) L. Xiao, and S. Boyd : "Fast linear iterations for distributed averaging", *Systems & Control Letters*, Vol. 53, No. 1, pp. 65-78, 2004
- (7) F.S. Cattivelli, C.G. Lopes, and A.H. Sayed : "Diffusion recursive least-squares for distributed estimation over adaptive networks", *IEEE Transactions on Signal Processing*, Vol. 56, No. 5, pp. 1865-1877, 2008
- (8) F.S. Cattivelli, and A.H. Sayed : "Diffusion LMS strategies for distributed estimation", *IEEE Transactions on Signal Processing*, Vol. 58, No. 3, pp. 1035-1048, 2009
- (9) X. Zhao, A.H. Sayed : "Performance limits for distributed estimation over LMS adaptive networks", *IEEE Transactions on Signal Processing*, Vol. 60, No. 10, 2012
- (10) Sheng-Yuan Tu, and A.H. Sayed "Optimal Combination Rules for Adaptation and Learning over Networks", *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2011 4th IEEE International Workshop on, pp. 317-320, 2011
- (11) J. H. Seo, S. M. Jung, and P. G. Park : "A diffusion subband adaptive filtering algorithm for distributed estimation using variable step size and new combination method based on the MSD", *Digital Signal Processing*, Vol. 48, pp. 361-369, 2016
- (12) S. M. Jung, J. H. Seo, and P. G. Park : "A Variable Step-Size Diffusion Normalized Least-Mean-Square Algorithm with a Combination Method Based on Mean-Square Deviation", *Circuits, Systems, and Signal Processing*, Vol. 34, No. 10, pp. 3291-3304, 2015
- (13) Reza Abdolee, and Benoit Champagne : "Diffusion LMS Algorithms for Sensor Networks over Non-ideal Inter-sensor Wireless Channels", 2011 International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS), pp. 1-6, 2011
- (14) A. H. Sayed : "Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior", *IEEE Signal Processing Magazine*, pp. 155-171, 2013
- (15) P. G. Park, C.H. Lee, and J. W. Ko : "Mean-Square Deviation Analysis of Affine Projection Algorithm", *IEEE Transactions on Signal Processing*, Vol. 59, No. 12, 2011
- (16) A.H. Sayed : "Diffusion adaptation over networks", Academic Press Library in Signal Processing, 2013
- (17) A.H. Sayed : "Fundamentals of adaptive filtering", New York: Wiley, 2003