ISSN 1598-0170 (Print)
ISSN 2287-1136 (Online)
http://www.jksii.or.kr

상호 정보 에너지와 델타함수를 이용한 알고리즘의 충격성 잡음에 대한 강인성

Robustness to Impulsive Noise of Algorithms based on Cross-Information Potential and Delta Functions

김 남 용[1*]
Namyong Kim

요 약

이 논문에서는 델타함수와 상호 정보 에너지(cross information-potential with the delta functions, CIPD)에 기반한 블라인드 등화 알고리즘의 최적 가중치를 유도하고 충격성 잡음에 대해 가지는 강인성에 대해 분석하였다. CIPD 알고리즘의 입력에 대한 크기조절 기능이 정상상태 가중치를 충격성 잡음으로부터 안정되게 유지하는 주된 역할을 하는 것으로 분석되었으며 시뮬레이션 결과를 통하여, CIPD 알고리즘의 정상상태 가중치는 MSE 성능기준의 최적해를 가지면서도, 충격성 잡음에서 MSE에 기반한 LMS 알고리즘과 달리, 안정된 값을 유지함을 보였다.

☞ 주제어 : 상호정보에너지, 델타함수, CIPD, 충격성 잡음, 강인성

ABSTRACT

In this paper, the optimum weight of the algorithm based on the cross information-potential with the delta functions (CIPD) is derived and its robustness against impulsive noise is studied. From the analysis of the behavior of optimum weight, it is revealed that the magnitude controlling operation for input plays the main role of keeping optimum weight of CIPD stable from the impulsive noise. The simulation results show that the steady state weight of CIPD is equivalent to that of MSE criterion. Also in the simulation environment of impulsive noise, unlike the LMS algorithm based on MSE, the steady state weight of CIPD is shown to be kept stable.

☞ keyword : Cross-information potential, delta function, CIPD, impulsive noise, robustness

1. INTRODUCTION

Wireless communication systems are often affected by impulsive noise from a variety of sources [1][2]. In the environment with impulsive noise that causes large instantaneous errors and system instability, algorithms based on the criterion of mean squared error (MSE) may fail to compensate for intersymbol interference (ISI) [3].

As a cost function based on information theoretic learning, cross information potential with delta functions (CIPD) has been developed into a blind algorithm that yields superior ISI cancelation performance under impulsive noise environment[4].

[1] Division of Electronics, Information & Communication Engineering, Kangwon National Unversity, Samcheok, Gangwon-Do, 245-711, Republic of Korea
* Corresponding author (namyong@kangwon.ac.kr)
[Received 29 September 2015, Reviewed 19 October 2015, Accepted 16 February 2016]

One of the problems of the CIPD algorithm was that its computational complexity was heavy caused by gradient estimation. This computational burden has been dealt with in the work [5] and significantly reduced by estimating the gradient based on a recursive approach so that the CIPD algorithm has been better suited to practical situations. However, analytic research on its optimum solutions and their behavior has not been carried out yet.

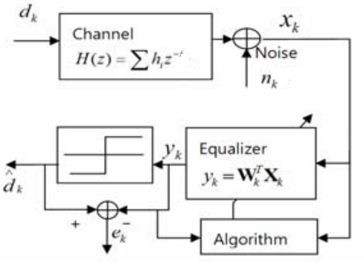

Fig. 1. Base-band communication system model

http://dx.doi.org/10.7472/jksii.2016.17.2.11

In this paper, the optimum weight of the CIPD algorithm is derived and analyzed in the aspect of the robustness against impulsive noise. And through simulation under impulsive noise, the stability of the steady state weight of the CIPD algorithm is compared with MSE-based least mean squared (LMS) algorithm [6].

2. MSE CRITERION AND RELATED ALGORITHMS

A symbol d_k at time k is transmitted through the channel $H(z)$ distorted by multipath and additive noise n_k as shown in Fig. 1 [6]. The equalizer input x_k becomes

$$x_k = \sum h_i d_{k-i} + n_k \tag{1}$$

With the input $\mathbf{X}_k = [x_k, x_{k-1}, ..., \ x_{k-L+1}]^T$ and weight $\mathbf{W}_k = [w_{0,k}, w_{1,k}, \ .., w_{L-1,k}]^T$ of a tapped delay line (TDL) filter structure, the equalizer output y_k and the error e_k become

$$y_k = \mathbf{W}_k^T \mathbf{X}_k \tag{2}$$

$$e_k = d_k - y_k = d_k - \mathbf{W}_k^T \mathbf{X}_k \tag{3}$$

Taking statistical average $E[\cdot]$ to the error power $e_k^{\ 2}$, the MSE criterion $E[e_k^{\ 2}]$ is defined and the optimum weight $\mathbf{W}_{MSE}^{opt}$ for MSE criterion is

$$\mathbf{W}_{MSE}^{opt} = \frac{E[d_k \mathbf{X}_k]}{E[\mathbf{X}_k \mathbf{X}_k^T]} \tag{4}$$

With this optimum weight $\mathbf{W}_{MSE}^{opt}$, $E[e_k \mathbf{X}_k]$ in the optimum state becomes

$$E[e_k \mathbf{X}_k] = E[d_k \mathbf{X}_k] - \mathbf{W}_{MSE}^{opt} \cdot E[\mathbf{X}_k \mathbf{X}_k^T] = 0 \tag{5}$$

The statistical average $E[\cdot]$ is commonly replaced with sample average or time average operation in practice. The LMS algorithm based on the sample average version of MSE criterion is to use one error power e_k^2 instead of $E[e_k^2]$ for practical reasons [6]. The gradient $\frac{\partial e_k^2}{\partial \mathbf{W}}$ for minimization of e_k^2 becomes

$$\frac{\partial e_k^2}{\partial \mathbf{W}} = 2e_k \frac{\partial (d_k - y_k)}{\partial \mathbf{W}} = -2e_k \mathbf{X}_k \tag{6}$$

Using the steepest descent method leads to the LMS algorithm as

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu \cdot \frac{\partial e_k^2}{\partial \mathbf{W}} = \mathbf{W}_k + 2\mu \cdot e_k \mathbf{X}_k \tag{7}$$

By letting the gradient $\frac{\partial e_k^2}{\partial \mathbf{W}}$ be zero, we have the optimum weight of LMS algorithm $\mathbf{W}_{LMS}^{opt}$ as

$$\mathbf{W}_{LMS}^{opt} = \frac{d_k \mathbf{X}_k}{\mathbf{X}_k \mathbf{X}_k^T} \tag{8}$$

We see that $\mathbf{W}_{LMS}^{opt}$ may fluctuate since noise exists within the input $\mathbf{X}_k$, but the statistically averaged $E[\mathbf{W}_{LMS}^{opt}]$ becomes equal to $\mathbf{W}_{MSE}^{opt}$ in Gaussian noise environments. Non-Gaussian noise such as impulsive noise, however, may defeat the averaging operation since even an impulse can dominate the averaging operation. Therefore we may speculate that $\mathbf{W}_{LMS}^{opt}$ can become unstable under impulsive noise environment.

In the following section, we will discuss the optimum weight behavior of CIPD algorithm known for its robustness against impulsive noise.

3. CIPD ALGORITHM AND OPTIMUM WEIGHT

With N samples (sample size is N) of output $\{y_k, y_{k-1}, ..., y_{k-N+1}\}$ the probability density

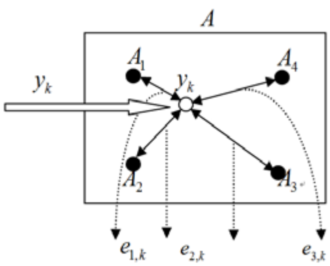

Fig. 2. Symbol space A and error samples for $M = 4$.

function of output $f_Y(y)$ can be constructed as in (9) based on Kernel density estimation [7].

$$f_Y(y) = \frac{1}{N}\sum_{i=k-N+}^{k} G_\sigma(y - y_i)$$
$$= \frac{1}{N}\sum_{i=k-N+1}^{k}\frac{1}{\sigma\sqrt{2\pi}}\exp[\frac{-(y-y_i)^2}{2\sigma^2}] \tag{9}$$

The desired symbols $\{A_1, A_2, ..., A_M\}$ in the symbol space A of CIPD algorithm are assumed to be i.i.d and the transmitted symbol d_k is one element of the symbol space A. Therefore

$$f_A(a) = \frac{1}{M}[\delta(a - A_1) + ... + \delta(a - A_M)] \tag{10}$$

Then the cross information-potential (CIP) with the delta functions $CIPD_{AY}$ can be expressed using the kernel density estimation method as [5].

$$CIPD_{AY} = \frac{1}{M}\frac{1}{N}\sum_{m=1}^{M}\sum_{i=k-N+1}^{k} G_\sigma(A_m - y_i) \tag{11}$$

The gradient of (11) for the maximization of $CIPD_{AY}$ becomes

$$\frac{\partial CIPD_{AY}}{\partial \mathbf{W}} = \frac{2}{MN\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{m=1}^{M}(A_m - y_i)$$
$$\cdot G_\sigma(A_m - y_i)\cdot \mathbf{X}_i \tag{12}$$

If we consider the sample-averaged operation $\frac{1}{N}\sum_{i=k-N+1}^{k}(\cdot)$ in (12) is treated as the same as the statistical average $E[\cdot]$ for practical purposes, we can rewrite (12) as

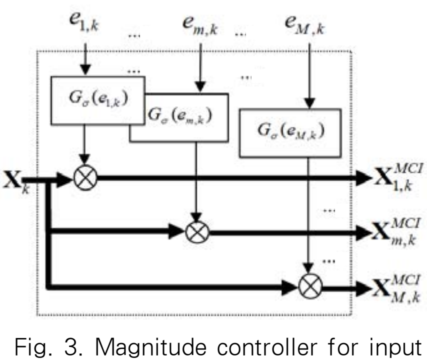

Fig. 3. Magnitude controller for input

$$\frac{\partial CIPD_{AY}}{\partial \mathbf{W}} = \frac{2}{MN\sigma^2}E[\sum_{m=1}^{M}(A_m - y_k)$$
$$\cdot G_\sigma(A_m - y_k)\cdot \mathbf{X}_k] \tag{13}$$

At the optimum state, the gradient becomes zero.

$$E[\sum_{m=1}^{M}(A_m - y_k)\cdot G_\sigma(A_m - y_k)\cdot \mathbf{X}_k] = 0 \tag{14}$$

Since the term $(A_m - y_k)$ implies how far the current output y_k is from each symbol A_m, we may define the difference $(A_m - y_k)$ as an symbol error $e_{m,k}$ for each symbol A_m. N symbol errors are generated from the symbol space at each iteration time as in Fig. 2 for a simple case of $M = 4$.

Then, (14) can be written as

$$E[\sum_{m=1}^{M} e_{m,k}\cdot G_\sigma(e_{m,k})\mathbf{X}_k] = 0 \tag{15}$$

Comparing the optimum condition of LMS algorithm $E[e_k\mathbf{X}_k] = 0$ in (5) to (15) , we may regard $G_\sigma(e_{m,k})\mathbf{X}_k$ as a kind of modified input. And we see that the term $G_\sigma(e_{m,k})\mathbf{X}_k$ in (15) implies that $\mathbf{X}_k$ is magnitude-controlled by $G_\sigma(e_{m,k})$ according to error values. For example, when symbol error $e_{m,k}$ has a very large value induced from some strong noise like impulses, the Gaussian function output $G_\sigma(e_{m,k})$ becomes very small (its exponential is a decay function of error power), so that the value of input $\mathbf{X}_k$ is cut down by the multiplication of $G_\sigma(e_{m,k})$. With the definition of $\mathbf{X}_{m,k}^{MCI}$ as a magnitude controlled input (MCI) in (16), this process is described in Fig. 3.

$$\mathbf{X}_{m,k}^{MCI} = G_\sigma(e_{m,k})\mathbf{X}_k \tag{16}$$

With $\mathbf{X}_{m,k}^{MCI}$ and (12), the CIPD algorithm can be rewritten as

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \frac{2\mu}{MN\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{m=1}^{M} e_{m,i}\cdot \mathbf{X}_{m,i}^{MCI} \tag{17}$$

Compared with the LMS algorithm in (7), the two algorithms are very similar in the aspect of error and input terms. On the other hand, it can be noticed that the MCI $\mathbf{X}_{m,k}^{MCI}$ can keep the algorithm stable even at large error occurrences such as when the input is contaminated by impulse noise, while the LMS has no such protection measures. Another different aspect is the summation process over symbol points and $e_{m,i}\mathbf{X}_{m,i}^{MCI}$. But this process does not seem to contribute much to deterring the influence of large errors since even an impulse can dominate the averaging operation.

Now, the optimum condition $\frac{\partial CIPD_{AY}}{\partial \mathbf{W}}=0$ with $y_k=\mathbf{X}_k^T\mathbf{W}^{opt}$ becomes

$$\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}(A_m-\mathbf{X}_i^T\mathbf{W}^{opt})\cdot\mathbf{X}_{m,i}^{MCI}$$
$$=\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}A_m\mathbf{X}_{m,i}^{MCI}-\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}\mathbf{X}_i^T\mathbf{W}^{opt}\mathbf{X}_{m,i}^{MCI}$$
$$=0 \tag{18}$$

That is

$$\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}A_m\mathbf{X}_{m,i}^{MCI}=\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}\mathbf{X}_i^T\mathbf{W}^{opt}\mathbf{X}_{m,i}^{MCI} \tag{19}$$

This relationship yields the optimum weight of CIPD algorithm as

$$\mathbf{W}^{opt}=\frac{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}A_m\mathbf{X}_{m,i}^{MCI}}{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}\mathbf{X}_i^T\mathbf{X}_{m,i}^{MCI}} \tag{20}$$

Since we can assume that most error samples are concentrated in the steady state at around zero, the kernel $G_\sigma(e_{m,k})$ can be treated as a constant $1/\sigma\sqrt{2\pi}$. That is,

$$\lim_{k\to\infty}G_\sigma(e_{m,k})=\frac{1}{\sigma\sqrt{2\pi}} \tag{21}$$

And

$$\lim_{k\to\infty}\mathbf{X}_{m,k}^{MCI}=\mathbf{X}_k \tag{22}$$

Therefore the expectation of the optimum weight $E[\mathbf{W}^{opt}]$ for (20) can be rewritten as

$$E[\mathbf{W}^{opt}]=\frac{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}E[A_m\mathbf{X}_{m,i}^{MCI}]}{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}E[\mathbf{X}_i^T\mathbf{X}_{m,i}^{MCI}]}$$
$$=\frac{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}E[A_m\cdot\mathbf{X}_i]}{\sum_{i=k-N+1}^{k}\sum_{m=1}^{M}E[\mathbf{X}_i^T\cdot\mathbf{X}_i]} \tag{23}$$

Since the transmitted symbol d_k is an element of the symbol space A, $E[\mathbf{W}^{opt}]$ becomes

$$E[\mathbf{W}^{opt}]=\frac{E[d_k\cdot\mathbf{X}_k]}{E[\mathbf{X}_k^T\cdot\mathbf{X}_k]} \tag{24}$$

This indicates

$$E[\mathbf{W}^{opt}]=E[\mathbf{W}_{LMS}^{opt}]=\mathbf{W}_{MSE}^{opt} \tag{25}$$

When the input becomes contaminated with impulsive noise even in the steady state, $\mathbf{X}_k$ can have an excessive value, this in turn makes the steady state (optimum) weight $\mathbf{W}_{LMS}^{opt}$ in (8) wildly fluctuate. However, thanks to the magnitude-controlling $G_\sigma(e_{m,k})$ cutting outliers, the MCI $\mathbf{X}_{m,k}^{MCI}$ in both the nominator and denominator of (22) can be in an acceptable range, so that we can be sure that $\mathbf{W}^{opt}$ remains stable in the steady state without getting shaky under such an impulsive noise situation. Assuming that most error samples are located at around zero in the steady state, we see that the behavior of steady state weight can be regarded as that of optimum weight. So in the following section, on behalf of $\mathbf{W}_{LMS}^{opt}$ and $\mathbf{W}^{opt}$, their weight traces will be investigated through simulation in order to verify the property $E[\mathbf{W}^{opt}]=E[\mathbf{W}_{LMS}^{o}]$ in (25) and the stability of $\mathbf{W}^{opt}$ in impulsive noise environments.

4. SIMULATION RESULTS

As one of 4 symbol points $\{d_1=-3, d_2=-1, d_3=1, d_4=3\}$, d_k is sent at time k from the transmitter ($M=4$) through $H(z)=0.304+0.903z^{-1}+0.304z^{-2}$ [8]. For the observation of the trace of the steady state weight, the impulse noise is applied in the steady state (after convergence, 8000) to the channel output while the additive white Gaussian noise (AWGN) n_k with its variance of 0.001 is added to it throughout the whole time as depicted in Fig. 4. The variance of the impulse is 50 and its incident rate is 0.01 based on the generation of impulsive noise described in [9].

The number of tap weights of the TDL equalizer is L=11. The sample size N, the kernel size σ and convergence parameter μ for CIPD algorithm are 20, 0.5 and 0.007, respectively. The convergence parameter for LMS is 0.001. These parameter values are chosen to have the lowest steady state MSE for this simulation.

Figure 5 shows the MSE learning curves comparing the two algorithms of LMS and CIPD. At around 6000 samples, both algorithms converge completely and then they undergo the impulsive noise described in Fig. 4. The error power of LMS when hit by impulses shows sharp spikes and then stays in very high MSE up to 10 dB or decreases very slowly from that MSE value. On the other hand, the curve of CIPD shows no staying in high MSE state after the sharp spikes induced from the impulse noise. This indicates that CIPD has no turbulence or instability of steady state weight. For more detailed inspection, the results of weight trace are compared in Fig. 6. For the trace observation, only $w_{3,k}$, $w_{4,k}$ and $w_{5,k}$ are shown in Fig. 6 considering page-limits. The thick gray line is the trace of the LMS algorithm and the thin black line is the one of CIPD.

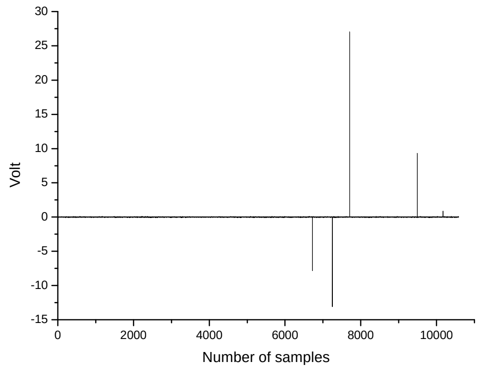

Fig. 4. The impulse noise and background AWGN.

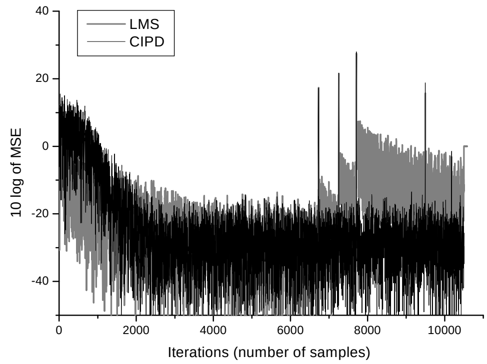

Fig. 5. MSE learning curves.

Assuming the steady state weight to be in the optimum state, it is reasonable for us to investigate if the steady state weight keeps the optimum value under impulsive noise. In Fig. 6, it is observed firstly that LMS and CIPD both have the same steady state weight values as described in (25). As the second point we found, the each weight trace of CIPD in the steady state shows no fluctuations at all remaining undisturbed under the strong impulses. This is obviously in contrast to the case of LMS algorithm where traces of all $w_{3,k}$, $w_{4,k}$ and $w_{5,k}$ have sharp perturbations at each impulse occurrence and remain perturbed although gradually dying.

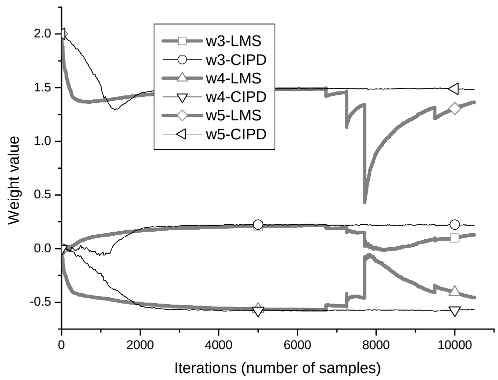

Fig. 6. The behavior of weight values of $w_{3,k}$, $w_{4,k}$ and $w_{5,k}$ with impulsive noise being added in the steady state.

By comparing the weight update equations (7) and (17), as main differences between them, CIPD is found to have sample average operations and MCI. As for the average or summation operation, we can notice that the dominant role of robustness against impulsive noise is not the average operation but the MCI since impulsive noise can defeat the average operation as explained in Section 2.

For a clearer quantitative comparison of the robustness against impulsive noise between $\mathbf{X}_k$ and $\mathbf{X}_{m,k}^{MCI}$, their traces are shown under the impulsive noise given in Fig. 4 are depicted in Fig. 7. Considering page-limits, MCI inputs $x_{0,k}^{MCI}$ and $x_{1,k}^{MCI}$ are compared to the original input x_k. As it can be expected, the input element x_k of $\mathbf{X}_k$ shows strong spikes up to 27 volts at the exact time instants affected directly from the impulsive noise given in Fig. 4. The magnitude controlled inputs, however, are staying unaffected by the impulsive noise. This result explains that the MCI $\mathbf{X}_{m,k}^{MCI}$ magnitude-controlled by $G_\sigma(e_{m,k})$ cutting outliers keeps both the nominator and denominator of (22) being in an acceptable range, so that

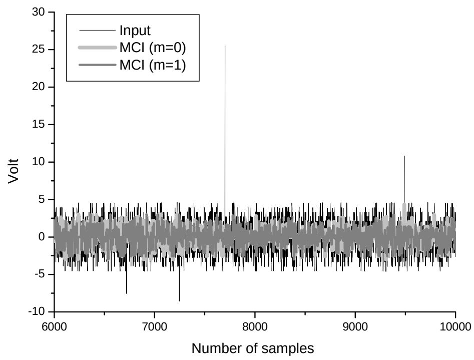

Fig. 7. The trace of the original input x_k (black line) and MCI ($x_{0,k}^{MCI}$, gray line and $x_{1,k}^{MCI}$, dark gray line) under the impulsive noise given in Fig. 4

$\mathbf{W}^{opt}$ remains stable under such impulsive noise situations.

5. CONCLUSION

The CIPD algorithm is known to outperform MSE-based algorithms in most blind signal processing applications in impulsive noise environment. But the optimum solutions and properties of the CIPD in regard to the robustness against impulsive noise have not been studied sufficiently. Through the derivation of the optimum weight of the CIPD algorithm and the analysis of the behavior of its optimum weight under impulsive noise, it can be concluded that the optimum weight of CIPD is equivalent to that of MSE criterion and the optimum weight optimum weight of CIPD is kept stable from impulsive noise. The MCI in CIPD is expected to be studied in more detail for further enhancement of performance in future research.

REFERENCES

[1] J. Xu, and D. Pham, “Robust Impulse-noise Filtering for Biomedical Images using Numerical Interpolation”, Image Analysis and Recognition,Vol. 7325 of LNCS, pp. 146-155. 2012. http://link.springer.com/chapter/10.1007%2F978-3-642-31298-4_18#page-1

[2] K. Blackard, T. Rappaport, and C. Bostian, “Measurements and Models of Radio Frequency Impulsive Noise for Indoor Wireless Communications,” IEEE Journal on selected areas in communications, Vol. 11, pp. 991 - 1001, Sep. 1993. http://dx.doi.org/10.1109/49.233212

[3] S. Unawong, S. Miyamoto, and N. Morinaga, “A Novel Receiver Design for DS-CDMA Systems under Impulsive Radio Noise Environments,” IEICE Trans. Comm., Vol. E82-B, pp. 936 -943, June 1999. http://search.ieice.org/bin/summary.php?id=e82-b_6_936&category=B&year=1999&lang=E&abst

[4] N. Kim, H. Byun, Y. You and K. Kwon, “Blind Signal Processing for Impulsive Noise Channels,” JCN, Vol. 14, pp. 27-33, Feb. 2012. http://dx.doi.org/10.1109/JCN.2012.6184548

[5] N. Kim, “Complexity Reduction of Blind Algorithms based on Cross-Information Potential and Delta Functions”, Journal of Internet Computing and Services (JICS), pp. 2187-2197, Vol. 15, pp. 9-18, June 2014. http://www.dbpia.co.kr/Article/NODE06187616

[6] S. Haykin, Adaptive Filter Theory, Prentice Hall, Upper Saddle River, 4th edition, 2001.

http://tocs.ulb.tu-armstadt.de/110863747.pdf
[7] E. Parzen, “On the Estimation of a Probability Density Function and the Mode,” Ann. Math. Stat. Vol. 33, p.1065, 1962.
http://bayes.wustl.edu/Manual/parzen62.pdf
[8] J. Proakis, Digital Communications, McGraw-Hill, 4th ed, 2012.
http://www.slideshare.net/hoangphuong2808/digital-communications-by-john-proakis-4th-edition
[9] I. Santamaria, P. Pokharel, and J. Principe, “Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization,” IEEE Trans. Signal Processing, Vol. 54, pp. 2187-2197, June 2006.
http://dx.doi.org/10.1109/TSP.2006.872524

◐ 저 자 소 개 ◑

김 남 용 (Namyong Kim)
1986년 연세대학교 전자공학과 졸업(학사)
1988년 연세대학교 대학원 전자공학과 졸업(석사)
1991년 연세대학교 대학원 전자공학과 졸업(박사)
1992-1998년 관동대학교 전자통신공학과 부교수
1998～현재 강원대학교 공학대학 전자정보통신공학부 교수
관심분야 : Adaptive Equalization, RBFN Odour Sensing Systems
E-mail : namyong@kangwon.ac.kr