논문 15-40-12-02 The Journal of Korean Institute of Communications and Information Sciences '15-12 Vol.40 No.12
http://dx.doi.org/10.7840/kics.2015.40.12.2324

충격성 잡음에 강인한 코렌트로피 기반 블라인드 알고리듬의 성능분석

김 남 용•

Performance Analysis of Correntropy-Based Blind Algorithms Robust to Impulsive Noise

Namyong Kim•

요 약

충격성 잡음하의 블라인드 신호처리 분야에서 최대 상호 코렌트로피 알고리듬 (MCC)이 MSE 기반의 알고리듬에 비해 우수한 성능을 보인다. 그러나 MCC 알고리듬에 대한 최적 가중치 조건들이나 충격성 잡음에 대한 내성과 관련된 특성들은 아직 충분히 연구되지 못한 상태이다. 이 논문에서는 MSE기반의 LMS 알고리듬과 비교를 통해 MCC의 최적 가중치의 성질을 분석하여 MCC 알고리듬의 최적 가중치가 MSE기반의 LMS 알고리듬과 같다는 보인다. 또한 MCC의 최적 가중치가 충격성 잡음 하에서도 동요 없이 안정을 유지하는 요인이 입력 크기 조정에 있다는 것을 시뮬레이션을 통해 입증하였다.

Key Words : Cross-correntropy, Random symbols, Impulsive noise, Optimum weight, Magnitude controlled, Equalizer

ABSTRACT

In blind signal processing in impulsive noise environment the maximum cross-correntropy (MCC) algorithm shows superior performance compared to MSE-based algorithms. But optimum weight conditions of MCC algorithm and its properties related with robustness to impulsive noise have not been studied sufficiently. In this paper, through the analysis of the behavior of its optimum weight and the relationship with the MSE-based LMS algorithm, it is shown that the optimum weight of MCC and MSE-based LMS have an equal solution. Also the factor that keeps optimum weight of MCC undisturbed and stable under impulsive noise is proven to be the magnitude controlled input through simulation.

Ⅰ. Introduction

Multipath propagation in wireless channel and impulsive noise from a variety of sources affects the communication systems adversely[1,2]. In the environment with impulsive noise, many signal processing methods based on MSE fail to function properly because of large instantaneous errors and instability[3].

As an alternative to the MSE criterion, the correntropy criterion similar to auto-correlation has been introduced[3]. The cross-correlation (CC) concept can be expressed with inner products of two different distribution functions constructed by kernel

• First Author : Division of Electronic, Information and Commun. Eng., Kangwon National Univ., 종신회원
논문번호 : KICS2015-09-310, Received September 15, 2015; Revised November 19, 2015; Accepted December 2, 2015

density estimation methods with Gaussian kernel [3,4]. Through maximization of CC (MCC) with steepest descent method and a set of symbol samples generated randomly at the receiver, the MCC algorithm has been developed for blind channel equalization in the environment of impulsive noise and multipath distortions[5].

One of the drawbacks of MCC algorithm is heavy computational complexity related with summation operations at each iteration time since gradients are calculated based on block processing method. This problem of computational complexity has been significantly reduced in the work [6] by utilizing recursive estimation of the gradient. Though the MCC algorithm has been developed to be better suited to practical situations and problems, analytic research on its optimum solutions and their behavior has not been carried out yet deterring further enhancement of the algorithm.

Ⅱ. MSE Criterion and MCC Algorithm

As depicted in Fig. 1, the baseband model of communication system, the transmitted symbol point d_k at time k is distorted by the channel's multipath fading and added noise n_k. For the multipath channel model $H(z)=\sum h_i z^{-i}$ in z-transform, the equalizer input x_k becomes (1)[7].

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

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

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

Then the MSE criterion P_{MSE} is defined as a statistical average $E[\cdot]$ of squared error $e_k^{\ 2}$.

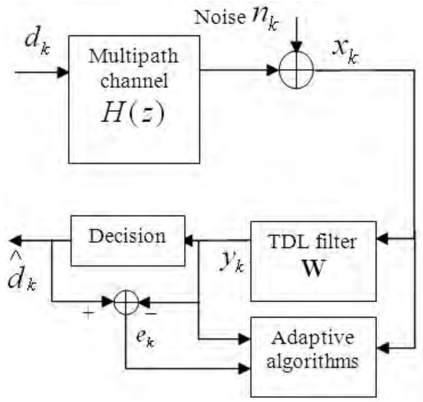

Fig. 1. Base-band communication system and adaptive equalizer structure

$$\begin{aligned} P_{MSE} &= E[e_k^{\ 2}] \\ &= E[d_k^2] + \mathbf{W}_k^T E[\mathbf{X}_k^T \mathbf{X}_k]\mathbf{W}_k - 2\mathbf{W}_k^T E[d_k \mathbf{X}_k] \end{aligned} \tag{3}$$

Letting the gradient $\frac{\partial P_{MSE}}{\partial \mathbf{W}}$ be zero, the optimum weight vector $\mathbf{W}_{MSE}^o$ for MSE criterion can be obtained[8].

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

Inserting this optimum weight $\mathbf{W}_{MSE}^o$ in the correlation $E[e_k \mathbf{X}_k]$ leads to

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

As a typical algorithm based on the MSE criterion, LMS (least mean square) is to use the instant error power $e_k^{\ 2}$ instead of $E[e_k^2]$ for practical reasons[7]. Then the gradient of LMS becomes

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

And

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

The optimum weight vector of LMS algorithm $\mathbf{W}_{LMS}^o$ can be obtained as below by letting the gradient $\frac{\partial e_k^2}{\partial \mathbf{W}}$ in (6) be zero.

$$\mathbf{W}_{LMS}^o = \frac{d_k \mathbf{X}_k}{\mathbf{X}_k^T \mathbf{X}_k} \quad (8)$$

And

$$E[\mathbf{W}_{LMS}^o] = \frac{E[d_k \mathbf{X}_k]}{E[\mathbf{X}_k^T \mathbf{X}_k]} = \mathbf{W}_{MSE}^o \quad (9)$$

While taking the averaging operation $E[\cdot]$ to (8) can mitigate the influence of the Gaussian noise on the steady state weights, non-Gaussian noise such as impulsive noise may defeat the averaging operation because even an impulse can dominate the mean operation. Therefore algorithms based on the MSE criterion can become unstable under impulsive noise environment.

Among blind algorithms known for its robustness against impulsive noise, a blind algorithm based on MCC criterion and a random symbol set has been developed for impulsive noise environment[7]. We assume that M symbol points are equally likely to be transmitted a priori with a probability $1/M$, and the transmitted symbol points A_m are $A_m = 2m-1-M$, $m = 1,2,...,M$. Since modulation schemes are normally known to receivers, the receiver generates a set of random symbol samples $\mathbf{D}_N = [d_1, d_2, ..., d_N]^T$ in order for the MCC method to have the same distribution as the transmitted symbol points $\{A_m\}$[5]. For that purpose, the number of random symbol samples corresponding to each symbol point A_m is N/M. Then the probability density can be constructed based on Kernel density estimation[9].

$$f_D(d) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp[\frac{-(d-d_i)^2}{2\sigma^2}] \quad (10)$$

The cross-correntropy concept can be expressed with inner products of two different probability density functions constructed by Gaussian-kernel density estimation methods[4]. Then the cross-correntropy criterion with the output distribution $f_Y(y)$ and $f_D(d)$ in (10) is $\int f_D(\alpha) f_Y(\alpha) d\alpha$

$$= \frac{1}{N^2} \sum_{i=k-N+1}^{k} \sum_{j=1}^{N} \frac{1}{\sigma\sqrt{2}\sqrt{2\pi}} \exp[\frac{-(d_j - y_i)^2}{4\sigma^2}] \quad (11)$$

For maximization of the cross-correntropy (MCC), the gradient can be utilized.

$$\frac{\partial \int f_d(\alpha) f_y(\alpha) d\alpha}{\partial \mathbf{W}}$$
$$= \frac{1}{2N^2\sigma^3\sqrt{\pi}} \sum_{i=k-N+1}^{k} \sum_{j=1}^{N} (d_j - y_i) \cdot \exp[\frac{-(d_j - y_i)^2}{4\sigma^2}] \cdot \mathbf{X}_i \quad (12)$$

With the gradient (12), the MCC algorithm is expressed as

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{\partial \int f_d(\alpha) f_y(\alpha) d\alpha}{\partial \mathbf{W}} \quad (13)$$

Ⅲ. Weight Behavior of MCC Algorithm

The term $\frac{1}{N} \sum_{i=k-N+1}^{k} (\cdot)$ in (12) can be considered as a sample-averaged version of $E[\cdot]$ so that the optimum condition $\frac{\partial \int f_d(\alpha) f_y(\alpha) d\alpha}{\partial \mathbf{W}} = 0$ leads to

$$E[\sum_{j=1}^{N}(d_j - y_k)\cdot\exp(\frac{-(d_j-y_k)^2}{4\sigma^2})\mathbf{X}_k]=0 \quad (14)$$

In (14), the term $d_j - y_k$ means how far the current output y_k is from each symbol sample d_j. Since N/M samples of d_j are located in each transmitted symbol point A_m as depicted in Fig. 2, the term $d_j - y_k$ corresponds to the distance between the current y_k and symbol points ($A_1,...,A_M$). From this perspective, we may define the difference $d_j - y_k$ as an error $e_{j,k}$ for each symbol sample d_j. For convenience sake, this error $e_{j,k}$ will be referred to as symbol sample

(SS) error. N SS errors are produced from the symbol space at each iteration time as in Fig. 2 for a simple case of $M=4$ and $N=16$. Then, (14) can be written as

$$E[\sum_{j=1}^{N}e_{j,k}\cdot\exp(\frac{-e_{j,k}^2}{4\sigma^2})\mathbf{X}_k]=0 \quad (15)$$

The term $\exp(\frac{-e_{j,k}^2}{4\sigma^2})\mathbf{X}_k$, that is, $\mathbf{X}_k$ multiplied by $\exp(\frac{-e_{j,k}^2}{4\sigma^2})$ can be considered as a magnitude controlled value of $\mathbf{X}_k$ according to error values. For example, when SS error $e_{j,k}$ has a very large value due to some strong noise like impulses, the exponential $\exp(\frac{-e_{j,k}^2}{4\sigma^2})$ becomes very small (the exponential function is a decay function of SS error power) and the current input $\mathbf{X}_k$ is mitigated by the multiplication of $\exp(\frac{-e_{j,k}^2}{4\sigma^2})$. This process of input magnitude control is depicted in Fig. 3, defining $\mathbf{X}_{j,k}^{MC}$ as a magnitude controlled input.

$$\mathbf{X}_{j,k}^{MC}=\exp(\frac{-e_{j,k}^2}{4\sigma^2})\mathbf{X}_k \quad (16)$$

With the definition $\mathbf{X}_{j,k}^{MC}$ and (12), the MCC algorithm can be rewritten as

$$\mathbf{W}_{k+1}=\mathbf{W}_k+\frac{\mu}{2N^2\sigma^3\sqrt{\pi}}\sum_{i=k-N+1,}^{k}\sum_{j=1}^{N}e_{j,i}\cdot\mathbf{X}_{j,i}^{MC} \quad (17)$$

It is noticed in (17) that the magnitude controlled $\mathbf{X}_{j,k}^{MC}$ plays the role in stabilizing the algorithm in situations of large error occurrence when compared with the LMS algorithm in (7) though the two algorithms have a very similar form that comprises error and input.

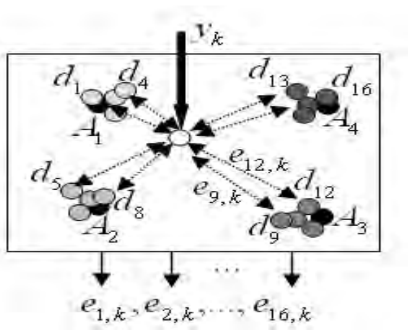

Fig. 2. Symbol space and error samples for $N=16$ and $M=4$.

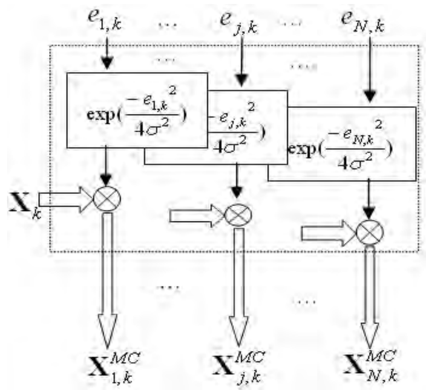

Fig. 3. Input magnitude controller

The steady state condition (15) becomes

$$E[\sum_{j=1}^{N} e_{j,k} \cdot \mathbf{X}_{j,k}^{MC}] = 0 \tag{18}$$

It is also observable when we compare (5) with (18) the optimum condition of MCC is very similar to MSE criterion except the summation process over symbol samples and the magnitude controlled input $\mathbf{X}_{j,k}^{MC}$.

From (12), the optimum condition becomes

$$\sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} (d_j - \mathbf{X}_i^T \mathbf{W}^o) \cdot \mathbf{X}_{j,i}^{MC} = 0 \tag{19}$$

$$\sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} d_j \cdot \mathbf{X}_{j,i}^{MC} = \sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} \mathbf{X}_i^T \mathbf{W}^o \cdot \mathbf{X}_{j,i}^{MC} \tag{20}$$

The optimum weight for MCC algorithm is

$$\mathbf{W}^o = \frac{\sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} d_j \cdot \mathbf{X}_{j,i}^{MC}}{\sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} \mathbf{X}_i^T \cdot \mathbf{X}_{j,i}^{MC}} \tag{21}$$

In the steady state, we may assume that most of the error samples are located at around zero as depicted in Fig. 4.

This assumption leads us to treat $\exp(\frac{-e_{j,k}^2}{4\sigma^2})$ as a constant. That is, in the steady state,

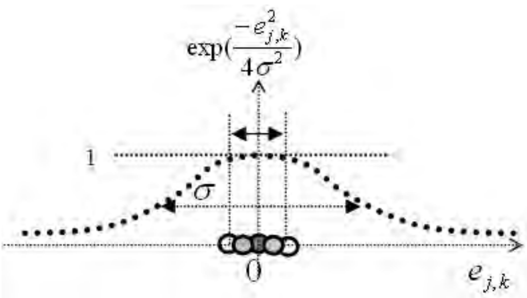

Fig. 4. Exponential function and error samples gathered around zero.

$$\lim_{k\to\infty} \exp(\frac{-e_{j,k}^2}{4\sigma^2}) = 1 \tag{22}$$

$$\lim_{k\to\infty} \mathbf{X}_{j,k}^{MC} = \mathbf{X}_k \tag{23}$$

Using (22) and (23), we may rewrite the expected value of optimum weight $E[\mathbf{W}^o]$ for (21) as

$$E[\mathbf{W}^o] = \frac{\sum_{i=k-N+1,}^{k} E[\sum_{j=1}^{N} d_j \cdot \mathbf{X}_{j,i}^{MC}]}{\sum_{i=k-N+1,}^{k} E[\sum_{j=1}^{N} \mathbf{X}_i^T \cdot \mathbf{X}_{j,i}^{MC}]}$$

$$= \frac{\sum_{j=1}^{N} E[d_j \cdot \mathbf{X}]}{\sum_{j=1}^{N} E[\mathbf{X}^T \cdot \mathbf{X}]} = \frac{E[d \cdot \mathbf{X}]}{E[\mathbf{X}^T \cdot \mathbf{X}]} \tag{24}$$

The equation (24) indicates

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

From the perspective of large error situation such as due to impulsive noise, the equation (21) gives another important property that the magnitude controlled $\mathbf{X}_{j,k}^{MC}$ both in the nominator and denominator cuts outliers that are abnormally large input contaminated with large noise. This in turn makes $\mathbf{W}^o$ remain stable without wild fluctuation in the steady state. Compared to the optimum weight of MCC, the one of LMS algorithm (8) has no protection measures against damage from large errors or impulsive noise. Assuming optimum condition that most error samples are located at around zero in the steady state, this property will be discussed through observations of the behavior of steady state weight (21) and (8) under impulsive noise situations in the following section.

Ⅳ. Results and Discussion

In this section, For the observation of the behavior of the steady state weights (21) and (8) under impulsive noise situations, the channel environment of [7] with the impulse noise being applied in the steady state is used in this paper as depicted in Fig. 5. The symbol point set in the transmitter is $\{d_1=-3, d_2=-1, d_3=1, d_4=3\}$ and a symbol point d_k is transmitted at time k through the channel models as

$$H_1(z)=0.26+0.93z^{-1}+0.26z^{-2} \quad (26)$$

$$H_2(z)=0.304+0.903z^{-1}+0.304z^{-2} \quad (27)$$

The additive Gaussian white noise (AWGN) is added to the channel output throughout the whole time as a background noise. The impulse noise is added after convergence (8000) as in Fig. 5. The impulsive noise n_k is generated according to the following PDF of Gaussian mixture model[3].

$$f_{NOISE}(n_k)=\frac{1-\varepsilon}{\sigma_1\sqrt{2\pi}}\exp[\frac{-n_k^2}{2\sigma_1^2}]+\frac{\varepsilon}{\sigma_2\sqrt{2\pi}}\exp[\frac{-n_k^2}{2\sigma_2^2}] \quad (28)$$

where ε<1, $\sigma_1=\sigma_{GN}$, $\sigma_2=\sqrt{\sigma_{GN}^2+\sigma_{IN}^2}$

The variance σ_{IN}^2 and incident rate ε of the impulse in this section are given by 50 and 0.01, respectively. The TDL equalizer has 11 weights. The number of random symbol samples N is 32, the kernel size σ is 0.5.

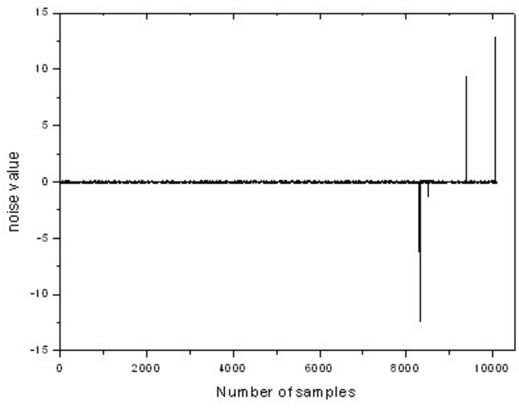

Fig. 5. Background and impulsive noise.

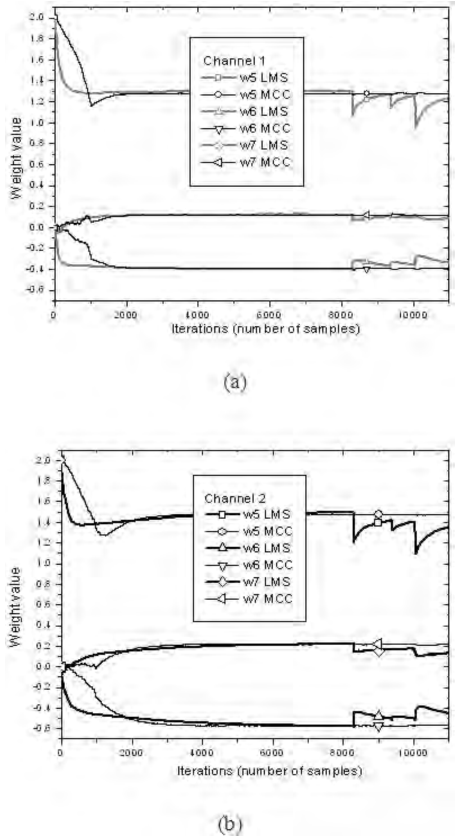

Fig. 6. Trace of weight values in the steady state with impulsive noise for $w_{5,k}$, $w_{6,k}$ and $w_{7,k}$ (the other weights are not included just for the page-limit); (a) is for channel $H_1(z)$ and (b) is for $H_2(z)$.

The convergence step-sizes are $\mu_{MCC}=0.007$ for MCC1 and $\mu_{LMS}=0.001$ for LMS algorithm. All the parameters are selected to have the lowest steady state MSE in this simulation.

The trace of $w_{5,k}$, $w_{6,k}$ and $w_{7,k}$ (the other tap weights are not included just for the page-limit) in Fig. 6. The dotted line is the trace of the LMS algorithm and the solid line is the one of MCC1.

Since the steady state weight vectors can be considered to be reached the optimum state, it is reasonable to investigate whether the steady state weights satisfy the relation in (25) and the steady state weight can keep the optimum values under impulsive noise situations. The first thing we can observe in Fig. 6 is that both algorithms have the same steady state weight values as explained in (25). Secondly, the weight traces of MCC1, each weight presents no fluctuations staying undisturbed under the strong impulses while the ones of LMS algorithm for all $w_{5,k}$, $w_{6,k}$ and $w_{7,k}$ show sharp perturbations at each impulse occurrence.

Comparison of (7) and (17) shows that the mainly different terms between the two weight update equations are sample averaging and magnitude controlling processes. Since impulsive noise may defeat the averaging operation as mentioned in Section 2, we can explain that the dominant role in the robustness against impulsive noise is the magnitude controlled input $\mathbf{X}_{j,k}^{MC} = \exp(\frac{-e_{j,k}^{2}}{4\sigma^{2}})\mathbf{X}_{k}$ and therefore the steady state weights of LMS algorithm without the input controlling function cannot avoid wild fluctuations in impulsive noise environment.

Ⅴ. Conclusion

In most blind signal processing applications in impulsive noise environment the MCC algorithm outperforms MSE-based algorithms. However, the optimum solutions and properties related with MCC algorithms have not been sufficiently studied. In this paper, through analysis of the relationship with the behavior of optimum weight of MSE-based LMS algorithm, it has been proven that the optimum weight of MCC is equal to the one of MSE criterion. Furthermore, it is the magnitude controlled input that keeps optimum weight of MCC undisturbed and stable by mitigating the influence of impulsive noise. Studies on application of the magnitude controlled input to enhanced methods are desirable in future research.

References

[1] K. Blackard, T. Rappaport, and C. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Sel. Area. Commun.*, vol. 11, pp. 991-1001, Sept. 1993.

[2] S. Unawong, S. Miyamoto, and N. Morinaga, "A novel receiver design for DS-CDMA systems under impulsive radio noise environments," *IEICE Trans. Commun.*, vol. E82-B, pp. 936 -943, Jun. 1999.

[3] I. Santamaria, P. Pokharel, and J. Principe, "Generalized correlation function: Definition, properties, and application to blind equalization," *IEEE Trans. Signal Process.*, vol. 54, pp. 2187-2197, Jun. 2006.

[4] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, pp. 5286-5298, Nov. 2007.

[5] N. Kim, S. Kang, and D. Hong, "Blind equalization based on maximum cross-correntropy criterion using a set of randomly generated symbols," *J. KICS,* vol. 35, pp. 33-39, Jan. 2010.

[6] N. Kim, "Practical approach to blind algorithms using random-order symbols and cross-correntropy," *J. KICS,* vol. 39A, pp. 149-154, Mar. 2014.

[7] J. Proakis, *Digital Communications*, McGraw-Hill, 2nd Ed., 1989.

[8] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, 4th Ed., 2001.

[9] E. Parzen, "On the estimation of a probability density function and the mode," *Ann. Math. Stat.*, vol. 33, p. 1065, 1962.

김 남 용 (Namyong Kim)

1986년 2월 : 연세대학교 전자공학과 졸업
1988년 2월 : 연세대학교 전자공학과 석사
1991년 8월 : 연세대학교 전자공학과 박사
1992년 8월~1998년 2월 : 관동대학교 전자통신공학과 부교수
1998년 3월~현재 : 강원대학교 공학대학 전자정보통신공학부 교수
<관심분야> Adaptive equalization.