

# 랜덤 심볼에 기반한 정보이론적 학습법의 스텝 사이즈 정규화

## Step-size Normalization of Information Theoretic Learning Methods based on Random Symbols

김 남 용\*  
Namyong Kim

### 요 약

랜덤 심볼열을 기반으로 한 정보이론적 학습법 (ITL)은 특정 확률분포를 갖도록 랜덤하게 발생시킨 심볼열을 타겟 데이터로 활용하고, 입력 데이터 사이의 확률분포 거리 최소화를 비유함수로 하여 설계된다. 이 방식의 단점으로, 고정상수를 알고리즘 갱신의 스텝사이즈로 사용하므로 입력 전력의 통계적 추이를 활용할 수 없다. 정보포텐셜 출력(Information potential output, IPO)와 연관된 기울기에서는 정보포텐셜 입력(Information potential input, IPI)이, 정보포텐셜 오차(Information potential error, IPE)와 관련된 기울기에서는 입력자체가 입력으로 작용함을 이 연구에서 밝혀내고, 입력의 전력 추이를 따로 계산하여 스텝사이즈 (step size)를 정규화하도록 제안하였다. 제안된 알고리즘은 충격성잡음과 다중경로 페이딩 환경의 통신시스템 실험에서 기존 방식보다 약 4dB 정도 더 낮은 정상 상태 오차 전력, 약 2배 이상 빠른 수렴속도를 나타냈다.

☞ 주제어 : 분포거리, 랜덤심볼, 입력 전력 정규화, 스텝사이즈, 정보포텐셜, 충격성 잡음.

### ABSTRACT

Information theoretic learning (ITL) methods based on random symbols (RS) use a set of random symbols generated according to a target distribution and are designed nonparametrically to minimize the cost function of the Euclidian distance between the target distribution and the input distribution. One drawback of the learning method is that it can not utilize the input power statistics by employing a constant stepsize for updating the algorithm. In this paper, it is revealed that firstly, information potential input (IPI) plays a role of input in the cost function-derivative related with information potential output (IPO) and secondly, input itself does in the derivative related with information potential error (IPE). Based on these observations, it is proposed to normalize the step-size with the statistically varying power of the two different inputs, IPI and input itself. The proposed algorithm in an communication environment of impulsive noise and multipath fading shows that the performance of mean squared error (MSE) is lower by 4dB, and convergence speed is 2 times faster than the conventional methods without step-size normalization.

☞ keyword : Distribution distance, random symbol, step-size, information potential, Impulsive noise.

## 1. 서 론

통신 채널상에 발생하는 왜곡과 잡음을 극복할 적응신호처리 기술은 설정된 성능기준을 바탕으로 최적 가중치를 찾아가도록 설계된다[1]. 가장 보편적인 성능기준인 MSE (mean squared error)가 충격성 잡음에 대해 취약한 반면 샘플간 거리를 변수로 가우시안 커널에 적용하여 충격성 잡음의 영향을 무력화시키는 ITL (information theoretic learning) 기술이 연구되고 있다[2][3][4].

ITL 학습법은 르나이 (Reny)의 정보 엔트로피 (entropy)에 파잔 (Parzan)의 커널분포 추정법 (kernel density estimation)을 도입하여, 정보포텐셜 (information potential, IP) 개념으로 설계되었다 [5][6][7].

송신된 훈련 심볼열 (training sequence)는 알 수 없으나, 수신단이 송신심볼의 확률분포 정보를 보유한 상태라면, 동일 확률분포의 심볼 샘플을 무작위로 발생시켜서 정보 복원 방식에 활용 할 수 있다. 이에 착안하여, 랜덤 심볼열과 수신단 신호처리의 출력 샘플들 사이에 분포 유사성을 극대화하도록 설계된 랜덤발생 심볼 기반의 분포거리 (distribution-distance) 최소화 알고리즘이 제안되었다 [8]. 랜덤 샘플을 발생해 활용하는 것은, 수식화가 어려운 그 어떤 송신 심볼의 확률분포에 대해서도 랜덤샘플로 발생시켜 활용 가능하다는 장점이 있으며, 많

1 Electronics, information and Communications Eng, Kangwon National University, Gangwon-do, 24341, Republic of Korea

\* Corresponding author: (namyong@kangwon.ac.kr)

[Received 19 July 2019, Reviewed 27 September 2019(R2 14 November 2019, R3 20 January 2020), Accepted 12 February 2020]

은 공학응용분야에서 확률분포 수식 모델링 대신 그 확률분포에 부합하는 랜덤 샘플들을 발생 시켜 활용하고 있다 [9][10].

논문 [8]의 분포거리 최소화 알고리즘은 두가지 확률분포, 즉, 출력샘플들이 가지는 확률분포  $f_Y(y)$ 와 랜덤 생성 심볼열이 확률분포  $f_D(d)$ 사이의 유클리드 거리 (Euclidian distance, ED)  $ED = \int [f_D(x) - f_Y(x)]^2 dx$ 를 최소화하며 이 논문에서는 편의상 MED-RS (minimum ED with random symbols) 알고리즘이라 부르기로 한다. 한편, MED-RS와 동일한 성능을 지니면서 계산량을 크게 줄인 recursive MED-RS (RMED-RS)가 제안되었다 [11].

그러나 타겟에 대한 확률분포 수식화가 불가능 한 경우에 사용가능한 이 방식은 수식화에 기반한 알고리즘보다 수렴성능이 떨어진다. 이 논문은 논문 [8]과 [11]의 MED-RS 에 대해 스텝 사이즈를 정규화하여 향상된 수렴 성능을 입증하고자 한다. 최근, 확률분포를 델타함수 집합(a set of delta functions)으로 수식화하고 비유함수 ED를 적용한 논문 [12]에서, 출력샘플간 엔트로피 함수에 엔트로피로 통제된 입력 (Entropy-governing input)의 전력 추이를 스텝 사이즈에 정규화하는 방식을 제안하였다. 이 접근방식을 활용하여, 본 논문에서는, 송신단 심볼들이 수학적으로 규명될 수 없는 확률분포를 가지는 조건에서도 적응신호처리가 가능한 MED-RS 알고리즘에 스텝사이즈 정규화를 적용한다.

## 2. MSE 기준 및 관련 알고리즘

적응신호처리가 TDL (tapped delay line) 이며 다중 경로채널을  $H(z) = \sum h_i z^{-i}$ 으로 표현할 때, 기저대역 통신시스템은 그림 1과 같이 표현될 수 있다. 시간 k에서 송신단의 M개의 심볼점 ( $S_1, S_2, \dots, S_M$ ) 중 하나인 송신 심볼  $a_k$ 가 전송된다고 가정한다. 입력신호  $\mathbf{X}_k = [x_k, x_{k-1}, \dots, x_{k-L+1}]^T$ 는 필터 가중치  $\mathbf{W}_k = [w_{0,k}, w_{1,k}, \dots, w_{L-1,k}]^T$ 를 통과하여 출력신호  $y_k = \mathbf{W}_k^T \mathbf{X}_k$ 를 만든다. 훈련열을 사용하는 경우, 오차신호  $e_k = a_k - y_k$ 를 활용하여 가중치를 갱신하며  $e_k^2$ 의 통계적 평균을 취한 MSE를 대부분 성능기준으로 한다.

현실적 구현을 위해 오차 순시전력의 기울기와 스텝 사이즈  $\mu_{LMS}$ 를 도입한 LMS (least mean square) 알고리즘은 다음과 같다 [1].

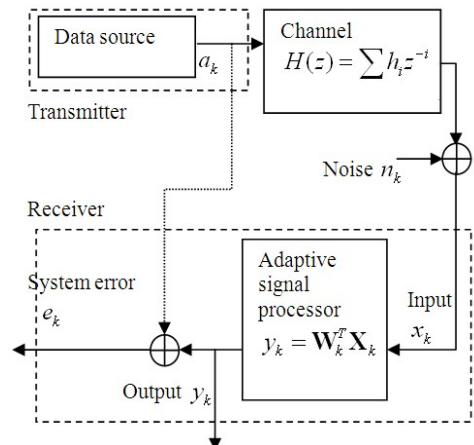
$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu_{LMS} \frac{\partial e_k^2}{\partial \mathbf{W}} \quad (1)$$

여기서  $\frac{\partial e_k^2}{\partial \mathbf{W}} = -2e_k \mathbf{X}_k$  (2)

한편, 입력전력  $\|\mathbf{X}_k\|^2 = \mathbf{X}_k^T \mathbf{X}_k = \sum_{m=0}^{L-1} x_{k-m}^2$ 에 스텝사이즈  $\mu_{LMS}$ 가 반비례하도록 (3)과 같이 정의하면 가중치 흔들림 (weight perturbation,  $\|\mathbf{W}_{k+1} - \mathbf{W}_k\|^2$ )을 최소화 할 수 있어서 실시간 적응시스템에 보다 효과적임이 밝혀졌다 [13][14].

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \frac{\mu_{NLMS}}{\|\mathbf{X}_k\|^2} e_k \mathbf{X}_k \quad (3)$$

한편, 훈련열 없이 송신단 심볼의 통계적 정보와 출력만 활용하는 blind learning 알고리즘인 CMA(constant modulus algorithm)는 송신심볼 통계적 정보  $R_2 = E[|a_k|^4]$ ,  $E[|a_k|^2]$ 와 출력전력의 차이인  $e_{CMA,k} = |y_k|^2 - R_2$ 를 오차로 하고 MSE 성능기준  $E[|e_{CMA,k}|^2]$ 에 적용하여 설계되었다 [15].



(그림 1) 기저대역 통신시스템 모델 (Figure 1) Baseband communication system model.

## 3. 랜덤심볼을 사용하는 분포거리 알고리즘

커널밀도 추정법 (kernel density estimation)은  $N$  개의 샘플  $\{y_1, y_2, \dots, y_i, \dots, y_N\}$ 이 있다고 할 때, 커널 사이즈  $\sigma$

의 가우시안 커널  $G_\sigma(y-y_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(y-y_i)^2}{2\sigma^2}]$ 을 각 샘플값에 위치시키고 평균하여 확률밀도를 추정한다 [7]. 현재시간  $k$ 에서 버퍼에 저장된  $N$ 개의 출력샘플들  $\{y_k, y_{k-1}, \dots, y_{k-N+1}\}$ 로 구한 확률분포함수는 다음과 같다.

$$f_Y(y) = \frac{1}{N} \sum_{i=0}^{N-1} G_\sigma(y-y_{k-i}) = \frac{1}{N} \sum_{i=k-N+1}^k G_\sigma(y-y_i) \quad (4)$$

블라인드 알고리즘인 MED-RS는 송신 심볼 확률분포 정보를 활용하여, 이 분포에 맞도록 수신단에서 랜덤 생성 심볼들을 사용한다. 균등한 발생확률로 무작위 발생하는 송신단  $M$ 개의 심볼점  $S=(S_1, S_2, \dots, S_m, \dots, S_M)$ 에 대해 수신단에서는 샘플집합  $D=\{d_1, d_2, d_3, \dots, d_N\}$ 이  $S$ 에 균등한 발생확률을 갖도록 발생시킨다. 즉,

$$f_D(d) = \frac{1}{N/M} \sum_{i=1}^{N/M} G_\sigma(d-S_1) + \dots + \frac{1}{N/M} \sum_{i=N-M+1}^N G_\sigma(d-S_M) \quad (5)$$

이제, 두 확률분포에 대한 유사성을 유클리드 분포거리로 정의하면

$$ED = \int [f_D(x) - f_Y(x)]^2 dx \quad (6)$$

식(4)와 (5)에 의해 식 (6)은 다음과 같아진다.

$$ED = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma\sqrt{2}}(d_j - d_i) + \frac{1}{N^2} \sum_{i=k-N+1}^k \sum_{j=k-N+1}^k G_{\sigma\sqrt{2}}(y_j - y_i) - 2 \frac{1}{N^2} \sum_{i=k-N+1}^k \sum_{j=1}^N G_{\sigma\sqrt{2}}(d_j - y_i) \quad (7)$$

한편, (5)에 의해 식 (7)의 첫 항은 다음과 같다.

$$\begin{aligned} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N G_{\sigma\sqrt{2}}(d_j - d_i) &= \frac{1}{N} \sum_{j=1}^N \frac{1}{N} \sum_{i=1}^N G_{\sigma\sqrt{2}}(d_j - d_i) \\ &= \frac{1}{N^2/M} \sum_{j=1}^N \sum_{i=1}^{N/M} G_{\sigma\sqrt{2}}(d_j - S_1) + \dots \\ &+ \frac{1}{N^2/M} \sum_{j=1}^N \sum_{i=N-M+1}^N G_{\sigma\sqrt{2}}(d_j - S_M) \end{aligned} \quad (8)$$

정보포텐셜 개념에서는 가우시안 커널이 마치 두 입자

사이의 상호작용(interaction)을 일으키는 포텐셜 장(potential field)를 만들고 있는 것으로 해석하므로,  $\sum_{j=1}^N G_{\sigma\sqrt{2}}(d_j - S_m)$ 는  $m$  번째 입자  $S_m$ 에 가해지는 힘들의 합에 해당되고, 식(8)은  $d_j$ 와  $S_m$  샘플들을 짝 지을 때, 모든 짝에 가해지는 상호 힘의 작용을 모은 전체 포텐셜 에너지로서 정보포텐셜(information potential)이 된다 [5].

식 (7)의 둘째, 세 짝 항은 수신 시스템에서 통제할 수 있는 변수이므로 채택하여 통제가능 유클리드 분포거리로 하고 최소화 과정을 시행한다. 또한 반복적(recursive) 계산법을 도입하여 계산량을 획기적으로 줄여 분포거리를 유도하는 방법이 개발되었다 [11].

식 (7)의 둘째 항과 셋째 항을 시간  $k$ 에 대해 다음과 같이 정의하면,

$$A_k = \frac{1}{N^2} \sum_{i=k-N+1}^k \sum_{j=k-N+1}^k G_{\sigma\sqrt{2}}(y_j - y_i) \quad (9)$$

$$B_k = -2 \frac{1}{N^2} \sum_{i=k-N+1}^k \sum_{j=1}^N G_{\sigma\sqrt{2}}(d_j - y_i) \quad (10)$$

$A_k$ 와  $B_k$ 는 다음과 같이 반복적 계산에 의해 구해질 수 있다 (초기값은 0으로 한다).

$$\begin{aligned} A_{k+1} &= A_k + \frac{2}{N^2} \sum_{j=k-N+1}^k G_{\sigma\sqrt{2}}(y_i - y_{k+1}) \\ &- \frac{2}{N^2} \sum_{j=k-N+1}^k G_{\sigma\sqrt{2}}(y_i - y_{k-N+1}) \\ &- \frac{2}{N^2} G_{\sigma\sqrt{2}}(y_{k+1} - y_{k-N+1}) + \frac{2}{N^2} G_{\sigma\sqrt{2}}(0) \end{aligned} \quad (11)$$

$$B_{k+1} = B_k - \frac{2}{N^2} \sum_{j=1}^N [G_{\sigma\sqrt{2}}(d_j - y_{k+1}) - G_{\sigma\sqrt{2}}(d_j - y_{k-N+1})] \quad (12)$$

#### 4. 정보포텐셜 입력의 전력 추이를 이용한 스텝사이즈 정규화

반복적 기울기 계산에 의한 RMED-RS 알고리즘의 가중치 계산과정을 정리하면 다음과 같다[11].

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu \left[ \frac{\partial A_k}{\partial \mathbf{W}} + \frac{\partial B_k}{\partial \mathbf{W}} \right] \quad (13)$$

여기서

$$\begin{aligned} \frac{\partial A_{k+1}}{\partial \mathbf{W}} &= \frac{\partial A_k}{\partial \mathbf{W}} + \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k (y_{k+1} - y_j) \\ &\quad \cdot G_{\sigma\sqrt{2}}(y_{k+1} - y_j)(\mathbf{X}_i - \mathbf{X}_{k+1}) \\ &\quad - \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k (y_{k-N+1} - y_j) \\ &\quad \cdot G_{\sigma\sqrt{2}}(y_{k-N+1} - y_j)(\mathbf{X}_i - \mathbf{X}_{k-N+1}) \\ &\quad - \frac{1}{N^2 \sigma^2} (y_{k-N+1} - y_{k+1}) \\ &\quad \cdot G_{\sigma\sqrt{2}}(y_{k-N+1} - y_{k+1})(\mathbf{X}_{k+1} - \mathbf{X}_{k-N+1}) \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial B_{k+1}}{\partial \mathbf{W}} &= \frac{\partial B_k}{\partial \mathbf{W}} - \frac{1}{N^2 \sigma^2} \sum_{j=1}^N [(y_{k+1} - d_j) \cdot G_{\sigma\sqrt{2}}(d_j - y_{k+1}) \mathbf{X}_{k+1} \\ &\quad - (y_{k-N+1} - d_j) \cdot G_{\sigma\sqrt{2}}(d_j - y_{k-N+1}) \mathbf{X}_{k-N+1}] \end{aligned} \quad (15)$$

입력의 상호 거리  $(\mathbf{X}_i - \mathbf{X}_j)$ 를  $\mathbf{X}_{i,j}$ 로, 출력 샘플간 상호거리  $(y_j - y_i)$ 를  $y_{j,i}$ 로, 랜덤심볼과 출력샘플간 상호거리, 또는 랜덤심볼 오차 거리  $d_j - y_i$ 를  $e_{j,i}$ 라고 정의하면, 새로 정의된 변수  $\mathbf{X}_{i,j}$ ,  $y_{j,i}$ 와  $e_{j,i}$ 는 각각 정보 포텐셜에 작용하는 입력변수, 정보 포텐셜에 작용하는 출력변수, 그리고 정보 포텐셜을 관장하는 오차변수로 볼 수 있다.

한편, 식(5)가 랜덤 발생된 심볼열에 기반하여 표현된 확률분포인 반면, 고정된  $M$  개의 송신 심볼 집합  $\{D_1, D_2, \dots, D_M\}$ 의 확률분포를 델타함수 집합으로 수식화한 논문 [12]에서는  $f_d(d)$ 가  $f_d(\alpha) = \frac{1}{M} [\delta(\alpha - D_1) + \delta(\alpha - D_2) + \dots + \delta(\alpha - D_M)]$ 로 표현된다. 여기서 고정 심볼값  $D_m$ 과 출력  $y_i$ 의 거리를 오차거리  $(D_m - y_i)$ 로 정의하고 있어서, 이 논문에서 정의하고 있는 랜덤심볼 오차 거리  $e_{j,i} = d_j - y_i$ 와는 완전히 다른 값이 된다. (14)와 (15)을 다시 쓰면,

$$\begin{aligned} \frac{\partial A_{k+1}}{\partial \mathbf{W}} &= \frac{\partial A_k}{\partial \mathbf{W}} + \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k y_{k+1,i} \\ &\quad \cdot G_{\sigma\sqrt{2}}(y_{k+1,i}) \mathbf{X}_{i,k+1} \\ &\quad - \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k y_{k-N+1,i} \cdot G_{\sigma\sqrt{2}}(y_{k-N+1,i}) \mathbf{X}_{i,k-N+1} \\ &\quad - \frac{1}{N^2 \sigma^2} y_{k-N+1,k+1} \cdot G_{\sigma\sqrt{2}}(y_{k-N+1,k+1}) \mathbf{X}_{k+1,k-N+1} \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial B_{k+1}}{\partial \mathbf{W}} &= \frac{\partial B_k}{\partial \mathbf{W}} - \frac{1}{N^2 \sigma^2} \sum_{j=1}^N [-e_{j,k+1} \cdot G_{\sigma\sqrt{2}}(e_{j,k+1}) \mathbf{X}_{k+1} \\ &\quad + e_{j,k-N+1} \cdot G_{\sigma\sqrt{2}}(e_{j,k-N+1}) \mathbf{X}_{k-N+1}] \end{aligned} \quad (17)$$

식(15)에서  $y_{k,i} \cdot G_{\sigma\sqrt{2}}(y_{k,i})$ 는 정보 포텐셜에 작용하는 출력  $y_{k,i}$ 의 함수이므로 정보포텐셜 출력 (IPO)는  $y_{j,i}^{IPO}$  정의하기로 한다.

$$y_{j,i}^{IPO} = y_{k,i} \cdot G_{\sigma\sqrt{2}}(y_{k,i}) \quad (18)$$

마찬가지로 정보포텐셜 오차(IPE)  $e_{j,i}^{IPE}$ 는

$$e_{j,i}^{IPE} = e_{j,i} \cdot G_{\sigma\sqrt{2}}(e_{j,i}) \quad (19)$$

식(18)과 (19)를 사용하여 반복적 기울기 추정식 (16)와 (17)에 대입하여 정리하면 (초기 기울기  $\frac{\partial A_0}{\partial \mathbf{W}} = 0, \frac{\partial B_0}{\partial \mathbf{W}} = 0$ ),

$$\begin{aligned} \frac{\partial A_{k+1}}{\partial \mathbf{W}} &= \frac{\partial A_k}{\partial \mathbf{W}} + \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k y_{k+1,i}^{IPO} \mathbf{X}_{i,k+1} \\ &\quad - \frac{1}{N^2 \sigma^2} \sum_{j=k-N+1}^k y_{k-N+1,i}^{IPO} \mathbf{X}_{i,k-N+1} \\ &\quad - \frac{1}{N^2 \sigma^2} y_{k-N+1,k+1}^{IPO} \mathbf{X}_{k+1,k-N+1} \end{aligned} \quad (20)$$

$$\frac{\partial B_{k+1}}{\partial \mathbf{W}} = \frac{\partial B_k}{\partial \mathbf{W}} - \frac{1}{N^2 \sigma^2} \sum_{j=1}^N [-e_{j,k+1}^{IPE} \mathbf{X}_{k+1} + e_{j,k-N+1}^{IPE} \mathbf{X}_{k-N+1}] \quad (21)$$

기울기 (20)에서는 입력거리에 의해 정의된 정보포텐셜 입력 (IPI)  $\mathbf{X}_{i,j}$ 가 정보포텐셜 출력  $y_{j,i}^{IPO}$ 와 곱해지고, 기울기 (21)에서는 입력 자체  $\mathbf{X}_i$ 가 정보포텐셜 오차  $e_{j,i}^{IPE}$ 와 곱해지는 형태로 구성되어있음을 알 수 있다. LMS 알고리즘의 식(2)와 비교할 때, 밀접한 유사성을 발견할 수 있다. 정보포텐셜 출력  $y_{j,i}^{IPO}$ 가 소속된 기울기  $\frac{\partial A_k}{\partial \mathbf{W}}$ 에서는 정보포텐셜 입력  $\mathbf{X}_{i,j}$ 이, 정보포텐셜 오차  $e_{j,i}^{IPE}$ 가 소속된 기울기  $\frac{\partial B_k}{\partial \mathbf{W}}$ 에서는  $\mathbf{X}_i$ 가 입력으로서 관여하므로 이  $\mathbf{X}_{i,j}$ ,  $\mathbf{X}_i$ 의 통계적 추이를 스텝사이즈 정규화에 활용할 수 있다.

여기서 식(20)과 (21)을, 식(2)의 기울기  $\frac{\partial e_k^2}{\partial \mathbf{W}} = -2e_k \mathbf{X}_k$ 와 비교하면, LMS 알고리즘의 기울기에서 오차가 입력

에 의해 곱해진 상태여서 오차와 입력이 직교 관계일 때 기울기가 0 되는 최적 가중치로 접근했다고 볼 수 있다.

한편, 이 절에서는 입력 전력의 추이를 고려하여 가중치 흔들림 (weight perturbation)을 최소화하도록 설계된 식(3)의 NLMS (normalized LMS)처럼 입력전력의 추이를 스텝사이즈의 정규화에 반영한 Normalized RMED-RS (NRMED-RS) 알고리즘을 제안한다. 식(20)의 기울기는 정보포텐셜 출력과 정보포텐셜 입력이 곱해진 상태로, (21)도 정보포텐셜 오차와 입력이 곱해진 상태이므로 기울기 성분중 (21)에는 정보포텐셜 입력의 전력 추이를, (21)에는 입력신호의 전력 추이를 반영할 필요가 있다.

즉, 식(14)에서  $\mu_A$ 에는 정보포텐셜 입력전력  $\frac{1}{N}\|\mathbf{X}_{i,j}\|^2 = \frac{1}{N}\sum_{i=k-N+1}^k \sum_{j=k-N+1}^k x_{i,j}^2$ 으로,  $\mu_B$ 에는  $\frac{1}{N}\|\mathbf{X}_k\|^2 = \frac{1}{N}\sum_{i=k-N+1}^k x_i^2$ 으로 나눈 새로운 시변 스텝사이즈를 채용한다.

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \mu_A \frac{\partial A_k}{\partial \mathbf{W}} - \mu_B \frac{\partial B_k}{\partial \mathbf{W}} \quad (22)$$

$$\mu_A = \mu \left/ \frac{1}{N} \sum_{i=k-N+1}^k \sum_{j=k-N+1}^k x_{i,j}^2 \right. \quad (23)$$

$$\mu_B = \mu \left/ \frac{1}{N} \sum_{i=k-N+1}^k x_i^2 \right. \quad (24)$$

한편, 입력 전력이 스텝 사이즈에 분모로 들어가므로 불안정성을 일으킬 수 있다. 전력을 평균하는  $\frac{1}{N}\sum_{i=k-N+1}^k$  만으로 충격성 잡음의 영향을 막기 어렵다고 판단되어 현 입력전력과 이전 산출전력에 배분인자  $\alpha$  ( $0 < \alpha < 1$ )로 적절한 배분으로 다음 산출 전력을 얻도록 한다 (초기  $Power_A(0) = Power_B(0) = 1$ ).

$$Power_A(k) = \alpha \cdot Power_A(k-1) + (1-\alpha) \sum_{j=k-N+1}^k x_{k,j}^2 \quad (25)$$

$$Power_B(k) = \alpha \cdot Power_B(k-1) + (1-\alpha)x_k^2 \quad (26)$$

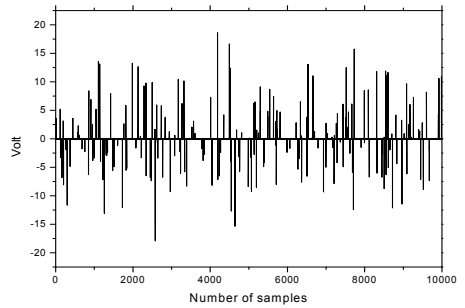
$$\mu_A = \mu / Power_A(k) \quad (27)$$

$$\mu_B = \mu / Power_B(k) \quad (28)$$

이 논문에서 제안한 NRMED-RS 알고리즘을 정리하면 식(20), (21), (22)과 (27), (28)으로 요약될 수 있다.

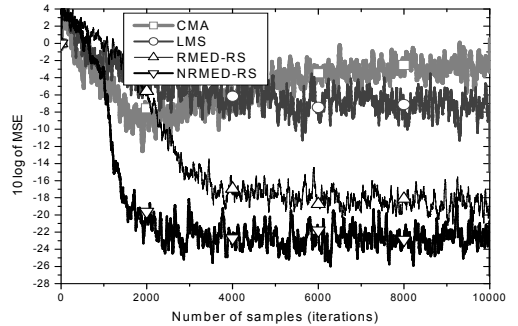
## 5. 결과 및 토론

그림 1의 기저대역 통신 시스템에서 송신 심볼은 독립적이며 동일 발생 확률을 가지는 4개 ( $M=4$ )의 값 (-3, -1, 1, 3)이며, 무작위 선택되어 전송된다. 전송되는 심볼을 모르는 상태로 수신단은 랜덤 심볼들( $N=32$ )을 발생시켜 사용한다. 다중경로 채널은  $H_1(z) = 0.26 + 0.93z^{-1} + 0.26z^{-2}$ 을 사용하였다 [16]. 잡음  $n_k$ 는 식(30)과 같은 확률밀도분포 함수  $f(n_k)$ 를 가지며, 포아송 발생률  $c=0.03$ , 분산  $\sigma_{IN}^2=50$ 을 가지는 충격 잡음(impulses)과 분산  $\sigma_{GN}^2=0.001$ 의 영평균 배경백색잡음이 합해진 형태로 만들어진다 [2]. 수신단 신호처리의 가중치는  $L=11$ 로, CMA와 LMS의  $\mu_{CMA}=0.000001$ ,  $\mu_{LMS}=0.0002$ 를 사용하였다. 기존의 RMED-RS [11]과 제안한 NRMED-RS에  $\sigma=0.5$ ,  $\mu=0.007$ ,  $\alpha=0.9$ 를 사용하였다. 성능 비교는 (그림 3)에서 MSE의 수렴성능을, (그림 4)에 오차확률분포를 나타냈다.

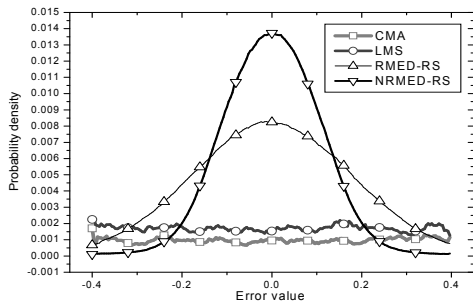


(그림 2) 충격성 잡음의 예.

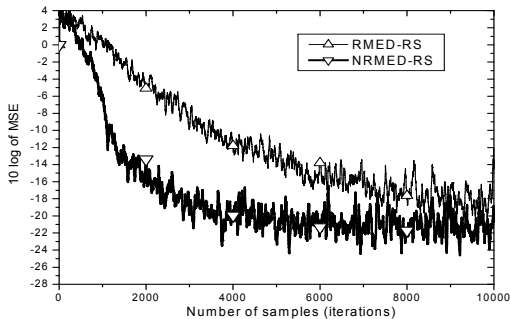
(Figure 2) An example of impulsive noise.



(그림 3) 채널모델  $H_1(z)$ 에 대한 MSE 수렴성능  
(Figure 3) MSE learning curves for  $H_1(z)$



(그림 4) 채널모델  $H_1(z)$  의 오차 확률분포  
(Figure 4) Error distribution for  $H_1(z)$



(그림 5) 채널모델  $H_2(z)$  에 대한 MSE 수렴성능  
(Figure 5) MSE learning curves for  $H_2(z)$

대표적인 블라인드 알고리즘인 CMA는 충격성 잡음 하에서 채널 왜곡을 보상하지 못하고 있다. 대표적인 supervised 알고리즘인 LMS조차도 -7dB 이하로 수렴하지 못하고 있다. 그러나 RMED-RS 와 이 논문에서 제안한 NRMED-RS는 현격한 성능 향상을 보이고 있다. RMED-RS는 -20 dB까지 내려갔으며 NRMED-RS는 이보다 약 4dB 정도 더 내려간 정상상태 오차 전력을 보이고 있다. 뿐만 아니라 수렴속도에서도 약 2배 이상 빠르다. (그림 4)에서 CMA의 오차 샘플들은 전혀 0에 집결하지 못하며 수렴을 보인 LMS조차 관찰 영역에서 0에 집결한다고 볼 수 없는 상태이다. 그러나 RMED-RS 와 NRMED-RS는 0을 기준으로 종모양의 분포형태를 나타내 오차 샘플들이 -0.2 와 0.2 사이에 대체로 집결하고 있다. 특히, NRMED-RS 알고리즘은 대부분 오차 샘플들이 -0.1 과 0.1 사이에 집결하고 있다.

한편, 그림 3은 정상상태 MSE가 약 -24dB이며 논문 [12]의 결과에서는 -26dB이다. 본 논문은 수식화가 불가능한 상황을 랜덤 발생 심볼열로 극복한 알고리즘에 기반하였지만, 논문[12]는 수학적 모델로 설계된 알고리즘에

스텝 사이즈를 정규화하였다는 것에 성능 차이가 있다.

제안한 알고리즘이 가지는 오차전력 수렴속도가 채널 특징에 따라 어떤 영향을 받는지 분석하기 위해 더 열린  $H_2(z) = 0.304 + 0.903z^{-1} + 0.304z^{-2}$ 에 대해 동일한 실험을 시행하였고 그 수렴 결과를 (그림 5)에 나타냈다. 수렴속도는 다소 느린 특징을 보였으나 제안한 방식인 NRMED-RS는 성능면에서 여전히 두 배 이상 빠른 수렴과 더 낮은 정상상태 MSE를 나타냈다.

## 6. 결 론

이 논문에서는 두 가지 서로 다른 기율기로 표현되는 RMED-RS 알고리즘에서 정보포텐셜 출력과 정보포텐셜 오차의 기율기에 각각 다른 입력이 관여함을 밝히고 전력 추이를 따로 계산하여 스텝사이즈를 정규화하였다. 이렇게 제안된 스텝사이즈 정규화에 의해 학습능력이 현저하게 향상된 것으로 실험결과에서 나타남에 따라, 정보포텐셜 입력의 전력 추이를 고려하여 스텝사이즈 정규화를 설계하고 이를 정보이론적 학습법에 기반한 가중치 갱신 알고리즘에 적용할 경우, 충격성 잡음하의 통신 환경이나, 특정 확률분포에 랜덤 샘플들을 발생 시켜 활용하는 신호처리 시스템에서 탁월한 성능향상을 기대할 수 있다고 판단된다.

## 참 고 문 헌

- [1] J. Principe, D. Xu and J. Fisher, *Information Theoretic Learning in: S. Haykin, Unsupervised Adaptive Filtering*, Wiley, (New York, USA), pp. 265-319, 2000.
- [2] 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.
- [3] W. Liu, P. P. Pokharel, and J. C. Principe, "Corentropy: Properties and applications in non-gaussian signal processing", *IEEE Trans. Signal Process.*, vol. 55, pp. 5286-5298, Nov. 2007.
- [4] N. Kim "Performance analysis of corentropy-based blind algorithms robust to impulsive noise", *The Journal of Korean Institute of Communications and Information Sciences* ' vol. 40, pp. 2324-2330, Dec. 2015.
- [5] N. Kim, "Information potential with shifted symbol

- points and related blind equalizer algorithms”, *Journal of The Institute of Electronics Engineers of Korea*, vol. 50, pp. 293-300, Feb. 2013.
- [6] H. Ting, F. Jun, W. Qiang, Z. Ding-Xuan, “Learning theory approach to minimum error entropy criterion”, *Journal of Machine Learning Research*, pp. 377-397, 2013.
- [7] E. Parzen, “On the estimation of a probability density function and the mode,” *Ann. Math. Stat.*, vol. 33, p. 1065, 1962.
- [8] N. Kim, and H. Byun, “Blind equalization based on Euclidian distance of information theoretic learning for impulsive noise environments”, *Proceedings of 2010 International conference on computer ccommunications and networks*, pp. 53-56, July 2010.
- [9] B. David, L. Wayne, “Efficient hardware generation of random variates with arbitrary distributions”, *14th Annual IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM’06)* April, 2006.
- [10] S. Richard, *Computer Generation of Statistical Distributions*, Storming Media, 2000.
- [11] N. Kim, “Recursive estimation of Euclidean distance between probabilities based on a set of random symbols”, *Journal of Internet Computing and Services*, vol. 15, pp. 119-124, Aug. 2014.
- [12] N. Kim, “Function analysis of the Euclidean distance between probability distributions”, *Entropy*, vol. 20, no. 48; doi:10.3390/e20010048
- [13] L. Bharani, P. Radhika, “FPGA implementation of optimal step size NLMS algorithm and its performance analysis”, *IJRET*, vol. 2, pp. 885-890, July 2013.
- [14] R. Chinaboina, “Adaptive algorithms for acoustic echo cancellation in speech processing, *IJRRAS*, vol. 7, pp. 38-42, April 2011.
- [15] R. Treichler, R. B. Agee, “A new approach to multipath correction of constant modulus signals”, *IEEE Trans. ASSP.*, vol. ASSP-31, pp. 349-372, Nov. 1983.
- [16] J. Proakis, *Digital Communications*, McGraw-Hill, 2th ed. USA, 1989.

## ● 저 자 소 개 ●



### 김 남 용

1986년 연세대학교 전자공학과 졸업(학사)  
 1988년 연세대학교 대학원 전자공학과 졸업(석사)  
 1991년 연세대학교 대학원 전자공학과 졸업(박사)  
 1992년~1998년 카톨릭관동대학교 전자통신공학과 부교수  
 1998년~현재 강원대학교 공학대학 전자정보통신공학부 교수  
 관심분야 : IITL learning