논문 15-40-02-01 The Journal of Korean Institute of Communications and Information Sciences '15-02 Vol.40 No.02
http://dx.doi.org/10.7840/kics.2015.40.2.247

영확률 성능기준에 근거한 결정궤환 알고리듬의 효율적인 계산

김 남 용◆

Efficient Calculation for Decision Feedback Algorithms Based on Zero-Error Probability Criterion

Namyong Kim◆

요 약

영확률을 성능기준으로 하는 적응 알고리듬은 충격성 잡음에 강인함을 나타내며 그 결정 궤환 알고리듬은 심각한 다경로 채널 왜곡을 효과적으로 보상하는 것으로 알려져 있다. 그러나 이러한 결정 궤환 영확률 알고리듬은 각 필터 구역에 대해 매 샘플시간마다 여러 합산 동작을 계산해야하는데 이것이 실제 구현에 장애가 되고 있다. 이 논문에서는 반복적 기울기 추정 방식을 가진 결정 궤환 영확률 알고리듬을 제안하며 이 알고리듬은 기존 계산량 $O(N)$을 샘플 사이즈 N에 무관한 상수량으로 줄일 수 있음을 보인다. 또한 초기상태와 안정상태의 가중치 갱신이 연속적인 과정으로 이루어져 결정 궤환에서 어떤 기울기 추정 오류 전파도 일으키지 않음을 보인다.

Key Words : decision feedback, ZEP, computational complexity, recursive gradient, continuity

ABSTRACT

Adaptive algorithms based on the criterion of zero-error probability (ZEP) have robustness to impulsive noise and their decision feedback (DF) versions are known to compensate effectively for severe multipath channel distortions. However the ZEP-DF algorithm computes several summation operations at each iteration time for each filter section and this plays an obstacle role in practical implementation. In this paper, the ZEP-DF with recursive gradient estimation (RGE) method is proposed and shown to reduce the computational burden of $O(N)$ to a constant which is independent of the sample size N. Also the weight update of the initial state and the steady state is a continuous process without bringing about any propagation of gradient estimation error in DF structure.

Ⅰ. Introduction

Wireless communication systems may suffer performance degradation by multipath propagation and impulsive noise in many communication systems such as satellite-mobile and underwater communications[1,2].

Conventional adaptive equalizer algorithms to cope with multipath problems are generally based on mean squared error (MSE) criterion. While the MSE-based algorithms are highly sensitive to impulsive noise, information theoretic learning methods utilizing kernel density estimation for

◆ First Author : Division of Electronic, Information & Comm. Engineering, Kangwon National Univ, namyong@kangwon.ac.kr, 종신회원
논문번호 : KICS2014-09-378, Received September 30, 2014; Revised December 19, 2014; Accepted January 12, 2015

nonparametric probability density function are robust to impulsive noise[3]. Maximization of zero-error probability (ZEP) has shown superior performance compared to MSE-based methods in supervised channel equalization applications[4].

In the work[5], the performance of ZEP with decision feedback (DF) algorithm under impulsive noise environments has been introduced. Compared to the non-DF ZEP algorithm, the algorithm with DF has shown significantly improved convergence in the situation of severely distorted channel and impulsive noise. However, one of the problems to be solved in that algorithm is the computational burden taking place in estimating the gradient for weight update equation at each iteration time. This problem must be coped with for practical implementation so that a new method of computing each section gradient efficiently is proposed in this paper. For that purpose, the efficient calculation method introduced in the linear algorithm[6] is modified and developed for the DF structure and it is investigated whether the method for DF might cause gradient-error propagation.

Ⅱ. Maximization of ZEP criterion for DF algorithms

The criterion of zero-error probability $f_E(0)$ constructed by kernel-based Parzen window density estimation is defined as in (1)[4].

$$f_E(0) = \frac{1}{N}\sum_{i=k-N+1}^{k} G_\sigma(-e_i) \tag{1}$$

where $G_\sigma(\cdot)$ is Gaussian kernel with kernel size σ, and$N$ is the sample size for the probability estimation. While large sample sizes ensure reliable density estimates, a computational cost scales directly with the sample size. Here lies the main practical difficulty with employing kernel-based Parzen window density estimators[7].

When error samples of adaptive systems are needed to be concentrated at zero for some applications, the criterion ZEP is maximized with respect to the system weights as $\max_W f_E(0)$.

When the adaptive system employs DF, it consists of a feed-forward filter section and a feedback filter section for producing corresponding decisions $\hat{d}_k$ from input x_k as depicted in Fig 1. With the feed-forward weight vector $\mathbf{W}_k^F = [w_{k,0}^F, w_{k,1}^F, w_{k,2}^F, ..., w_{k,P-1}^F]^T$, the feedback weight vector $\mathbf{W}_k^B = [w_{k,0}^B, w_{k,1}^B, w_{k,2}^B, ..., w_{k,P-1}^B]^T$ and the previously detected symbol vector $\hat{\mathbf{D}}_{k-1} = [\hat{d}_{k-1}, \hat{d}_{k-2}, ..., \hat{d}_{k-Q-2}]^T$, the output y_k to the input vector $\mathbf{X}_k = [x_k, x_{k-1} ..., x_{k-P+1}]^T$ is expressed as $y_k = \mathbf{X}_k^T \mathbf{W}_k^F + \hat{\mathbf{D}}_{k-l}^T \mathbf{W}_k^B$. Weights are updated toward maximizing the zero-error probability $f_E(0)$ (MZEP) using the gradient $\frac{\partial f_E(0)}{\partial \mathbf{W}^F} = \nabla_k^F$, $\frac{\partial f_E(0)}{\partial \mathbf{W}^B} = \nabla_k^B$ and the step size μ as in (2) and (3) that will be referred to as MZEP-DF algorithm as proposed in [5].

$$\mathbf{W}_{k+1}^F = \mathbf{W}_k^F + \mu \cdot \nabla_k^F \tag{2}$$

$$\mathbf{W}_{k+1}^B = \mathbf{W}_k^B + \mu \cdot \nabla_k^B \tag{3}$$

where the gradient vectors are

$$\nabla_k^F = \frac{1}{\sigma^2 N}\sum_{i=k-N+1}^{k} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i \tag{4}$$

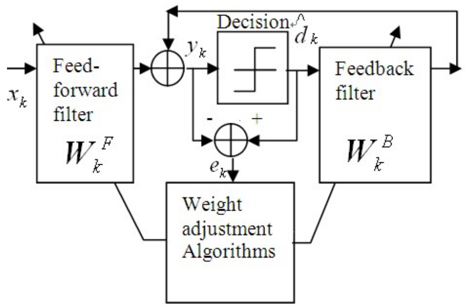

Fig. 1. Decision feedback equalizer structure.

$$\nabla_k^B = \frac{1}{\sigma^2 N} \sum_{i=k-N+1}^{k} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} \tag{5}$$

It is observed that the gradient vector ∇_k^F and ∇_k^B at each iteration are estimated through the computation of $O(N)$ (due to the summation operations) for each filter section. This can be a computational burden for practical implementation. For reduced computational complexity, a recursive gradient estimation (RGE) approach is proposed in the following section.

Ⅲ. RGE for MZEP-DF algorithms

Each gradient ∇_k^F and ∇_k^B can be calculated in the initial and steady state, separately. In the initial state $1 \le k \le N$ ($e_0 = 0$, $X_0 = 0$, and $\hat{\mathbf{D}}_{-1} = 0$), the number of available error samples are only k so that each gradient can be estimated as

$$\nabla_k^F = \frac{1}{\sigma^2 k} \sum_{i=1}^{k} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i \tag{6}$$

$$\nabla_k^B = \frac{1}{\sigma^2 k} \sum_{i=1}^{k} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} \tag{7}$$

We can separate the data related with the current time k from each summation as

$$\begin{aligned}\nabla_k^F &= \frac{1}{\sigma^2 k}[e_k \cdot G_\sigma(e_k) \cdot \mathbf{X}_k + \sum_{i=1}^{k-1} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i] \\ &= \frac{1}{\sigma^2 k} e_k \cdot G_\sigma(e_k) \cdot \mathbf{X}_k \\ &+ \frac{(k-1)}{\sigma^2 k(k-1)} \sum_{i=1}^{k-1} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i \\ &= \frac{1}{\sigma^2 k} e_k \cdot G_\sigma(e_k) \cdot \mathbf{X}_k + \frac{(k-1)}{k} \nabla_k^F\end{aligned} \tag{8}$$

Similarly, the backward gradient in the initial state becomes

$$\begin{aligned}\nabla_k^B &= \frac{1}{\sigma^2 k}[e_k \cdot G_\sigma(e_k) \cdot \hat{\mathbf{D}}_{k-1} + \frac{(k-1)}{\sigma^2 k(k-1)} \sum_{i=1}^{k-1} e_i \\ &\cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1}] \\ &= \frac{1}{\sigma^2 k} e_k \cdot G_\sigma(e_k) \cdot \hat{\mathbf{D}}_{k-1} + \frac{(k-1)}{k} \nabla_k^B\end{aligned} \tag{9}$$

In the steady state for $k \ge N+1$, the number of available error samples are the same as the sample size N. The gradient for the feedforward section at time k can be divided into the terms related with the current sample time k and the terms related with the previous sample times as

$$\begin{aligned}\nabla_k^F &= \frac{1}{\sigma^2 N} \sum_{i=k-N+1}^{k} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i \\ &= \frac{1}{\sigma^2 N}[\sum_{i=k-N}^{k-1} e_i \cdot G_\sigma(e_i) \cdot \mathbf{X}_i + e_k \\ &\cdot G_\sigma(e_k) \cdot \mathbf{X}_k - e_{k-N+1} \cdot G_\sigma(e_{k-N+1}) \cdot \mathbf{X}_{k-N+1}]\end{aligned} \tag{10}$$

Utilizing the definition ∇_k^F in(4) leads to

$$\begin{aligned}\nabla_k^F &= \nabla_{k-1}^F + \frac{1}{\sigma^2 N}[e_k \cdot G_\sigma(e_k) \cdot \mathbf{X}_k - e_{k-N+1} \\ &\cdot G_\sigma(e_{k-N+1}) \cdot \mathbf{X}_{k-N+1}]\end{aligned} \tag{11}$$

Similarly,

$$\begin{aligned}\nabla_k^B &= \frac{1}{\sigma^2 N} \sum_{i=k-N+1}^{k} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} \\ &= \frac{1}{\sigma^2 N}[\sum_{i=k-N}^{k-1} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} + e_k \cdot G_\sigma(e_k) \\ &\cdot \hat{\mathbf{D}}_{k-1} - e_{k-N+1} \cdot G_\sigma(e_{k-N+1}) \cdot \hat{\mathbf{D}}_{k-N}] \\ &= \nabla_{k-1}^B + \frac{1}{\sigma^2 N}[e_k \cdot G_\sigma(e_k) \\ &\cdot \hat{\mathbf{D}}_{k-1} - e_{k-N+1} \cdot G_\sigma(e_{k-N+1}) \cdot \hat{\mathbf{D}}_{k-N}]\end{aligned} \tag{12}$$

Here, the MZEP-DF algorithm, (2) and (3) can become a computationally efficient one by use of RGE methods (11) and (12) for the steady state, and (8) and (9) for the initial state. This algorithm will be referred to as MZEP-DF with RGE

(MZEP-DF-RGE) in this paper.

Ⅳ. Simulation Results

In this section, it is investigated how much the proposed MZEP-DF-RGE of (11) and (12) reduces computational complexity in multiplication when we compare it with the original MZEP-DF of (4) and (5). For convenience, the Gaussian kernel $G_\sigma(e_i)$commonly included in both methods is treated as a function value for e_i and $\frac{1}{\sigma^2 N}$ is treated as a constant. The equations (4) and (5) demand 4N+2 multiplications at each iteration time. However, the proposed (11) and (12) only require 10 multiplications regardless of the sample size N. As mentioned in section 2, the sample size required to be large for reliable density estimates. The property of being independent of the sample size N indicates that the proposed method reduces the computational cost considerably. Clearer comparisons are shown in Fig. 2 which reveals that the proposed MZEP-DF-RGE is more appropriate to practical implementations.

As another issue, gradient-error propagation problems are investigated. The feedback filter section cancels the ISI that remains after the feed forward section under the assumption of past decisions being correct. This indicates that incorrect decisions can induce error propagation. Similarly, it is observed that the recursive method might cause

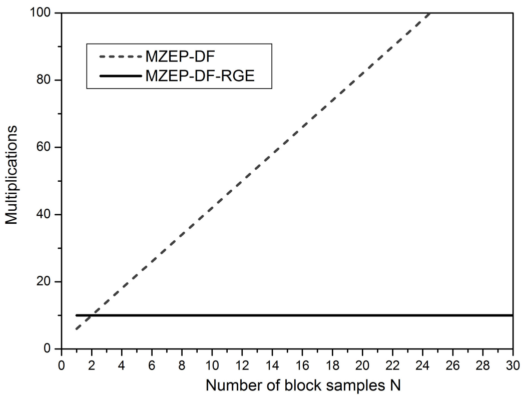

Fig. 2. Number of multiplications with respect to sample size N.

gradient-error propagation if the gradient value of theinitial state is discontinuously transferred to the steady state. This scenario shows that the proposed RGE method for DF can cause another serious problem to DF equalizers which are fragile to error propagation. So, it is required to investigate whether the estimation of each gradient has a continuous mode change from the initial state into the steady state. This can be done by seeing if the two modes yield the same gradient values at $k = N$. The initial state gradients (6) and (7) at $k = N$ are

$$\nabla_N^F = \frac{1}{\sigma^2 N}\sum_{i=1}^{N} e_i \cdot G_\sigma(e_i)\cdot \mathbf{X}_i \tag{13}$$

$$\nabla_N^B = \frac{1}{\sigma^2 N}\sum_{i=1}^{N} e_i \cdot G_\sigma(e_i)\cdot \hat{\mathbf{D}}_{i-1} \tag{14}$$

And the steady state gradients (10) and (12) at $k = N$ are

$$\nabla_N^F = \frac{1}{\sigma^2 N}[\sum_{i=0}^{N-1} e_i \cdot G_\sigma(e_i)\cdot \mathbf{X}_i + e_N \cdot G_\sigma(e_N)\cdot \mathbf{X}_N - e_1 \cdot G_\sigma(e_1)\cdot \mathbf{X}_1] \tag{15}$$

Inserting $e_0 = 0$ and $X_0 = 0$ into the summation in (15) and rearranging it leads to

$$\nabla_N^F = \frac{1}{\sigma^2 N}\sum_{i=1}^{N} e_i \cdot G_\sigma(e_i)\cdot \mathbf{X}_i \tag{16}$$

Also,

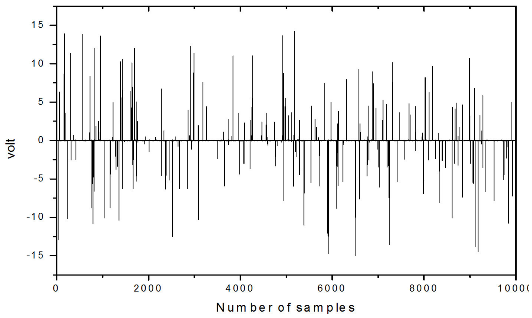

Fig. 3. The impulsive noise.

$$\nabla_N^B = \frac{1}{\sigma^2 N}[\sum_{i=0}^{N-1} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} + e_N \cdot G_\sigma(e_N) \cdot \hat{\mathbf{D}}_{N-1} - e_1 \cdot G_\sigma(e_1) \cdot \hat{\mathbf{D}}_0] \quad (17)$$

Since $e_0 = 0$, $X_0 = 0$, and $\hat{\mathbf{D}}_{-1} = 0$, ∇_N^B of (17) becomes

$$\nabla_N^B = \frac{1}{\sigma^2 N}\sum_{i=1}^{N} e_i \cdot G_\sigma(e_i) \cdot \hat{\mathbf{D}}_{i-1} \quad (18)$$

Observing that (13) and (14) are exactly the same as (16) and (18), respectively, we can judge that the gradient values of the initial state are continuously transferred to the steady state without bringing about any gradient error propagation.

To verify that the proposed MZEP-DF-RGE algorithm yield the same immunity against impulsive noise as MZEP-DF, MSE learning curves are compared in Fig. 2 under the same channel environment and equalizer factors used in [5]. That is, the transfer function the channel model is

$$H_1(z) = 0.304 + 0.903z^{-1} + 0.304z^{-2} \quad (19)$$

The channel output is added with impulsive noise as in Fig. 3. The variance 50, occurrence rate 0.03. The background noise AWGN of variance 0.001 is added so that the SNR for AWGN and binary symbol points is 30 dB. The step size μ for the proposed algorithm is varied from 0.0008 to 0.006 to investigate the behavior of residual MSE. The numbers of feed-forward and feedback weights are 7 and 4, respectively.

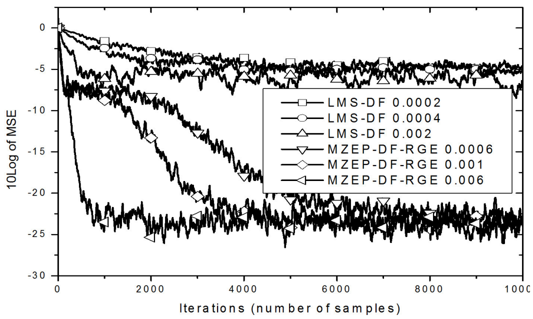

Fig. 4. MSE learning curves under impulsive noise environment

As shown in Fig. 4, the proposed MZEP-DF-RGE rapidly converges under impulsive noise and multipath channel distortion while the LMS - DF stays at about -6 dB with the gap of 19 dB compared to MZEP-DF-RGE. The residual MSE of the proposed algorithm varies from -23dB to -24dB when its step size is increased from 0.0006 to 0.006. The convergence speed shows wide difference as predicted but the difference of the residual MSE of the proposed algorithm is just 1 dB. Similar phenomenon is observed in the residual MSE of LMS-DF showing -4.6 and -6.6 dB for its step size is 0.0002 and 0.002, respectively.

V. Conclusions

The MZEP-DF algorithm has been known to have robustness to impulsive noise and severe multipath channel distortions. However the weight update process of MZEP-DF algorithm computes some summation operations at each iteration for each filter section. This computational burden being an obstacle for practical implementation can be avoided by the proposed MZEP-DF-GRE method which reduces the computational burden of $O(N)$ to a constant which is independent of the sample size N. Also the weight update of the initial state and the steady state is a continuous process without bringing about any propagation of gradient estimation error. These results lead us to conclude that the proposed MZEP-DF-GRE method is an appropriate candidate for practical implementations.

References

[1] M. Chitre, S. Shahabudeen, L. Freitag, and M. Stojanovic, "Recent advances in underwater acoustic communications & networking," in *Proc. MTS/IEEE OCEANS*, pp. 1-10, QC, Canada, Sept. 2008.

[2] M. Richharia, "Satellite communication systems: design principles," *Technology &*

Engineering, 1999.

[3] J. Principe, D. Xu, and J. Fisher, *Information Theoretic Learning, in: S. Haykin, Unsupervised Adaptive Filtering*, NY: Wiley, vol. I, pp. 265-319, 2000.

[4] N. Kim, K. H. Jeong, and L. Yang, "Maximization of zero error probability for adaptive channel equalization," *J. Commun. Networks(JCN)*, vol. 12, pp. 459-465, Oct. 2010.

[5] N. Kim, "Decision feedback equalizer based on maximization of zero-error probability," *J. KICS*, vol. 36, pp. 516-521, Aug. 2011.

[6] N. Kim, "Efficient adaptive algorithms based on zero-error probability," *J. KICS*, vol. 39A, pp. 237-243, May 2014.

[7] M. Girolami and C. He, "Probability density estimation from optimally condensed data samples," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 25, pp. 1253-1264, Oct. 2003.

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, RBFN algorithms, ITL algorithms, Odor sensing systems.