새로운 정보이론적 최적기준에 의한 블라인드 등화

A New Criterion of Information Theoretic Optimization and Application to Blind Channel Equalization

김 남 용* 리우킹 양**

Namyong Kim Liuqing Yang

요 약

인터넷에 관련된 많은 연구들이 멀티포인트 통신에 집중되어 있으며 이 멀티 포인트 통신에 블라인드 등화기술이 주로 사용되어 왔다. 이 논문에서는, 적응 블라인드 등화기를 위해, 두 확률밀도함수의 유클리드 거리를 최소화하는 기준을 소개한다. 파젼 확률밀도로 표현되는 유클리드 거리가 최소화되기 위해, 전송 심볼의 확률밀도함수와 같은 형태를 갖도록 하는 랜덤 심볼을 수신단에서 생성하여 사용한다. 시뮬레이션 결과로부터, 일반적으로 쓰이는 CMA 알고리듬이 열악한 성능을 보이는 채널모델에 대해 제안한 방식은 크게 향상된 성능을 나타냈다. CMA와 비교했을 때, 채널의 고유치분포비에 대해서도 비교적 민감하지 않은 일관된 성능을 보였으며, ITL의 한 영역으로서, 파젼 확률밀도를 사용한 유클리드 거리 최소화 기준이 블라인드 등화에 성공적으로 응용될 수 있음을 보였다.

Abstract

Blind equalization techniques have been used in multipoint communication on which the research on the internet has focused. In this paper, a criterion of minimizing Euclidian Distance between two PDFs for adaptive blind equalizers has been presented. In order for ED expressed with Parzen PDFs to be minimized, we propose to use a set of randomly generated desired symbols at the receiver so that the PDF of the generated symbols matches that of the transmitted symbols. From the simulation results, the proposed method has shown superior error performance even in severe channel environments in which CMA has shown severe performance degradation. This indicates that the proposed algorithm can be considered relatively insensitive to ESR variations compared to CMA. As a field of ITL, ED minimization using Parzen PDFs has shown possibilities of being successfully applied to blind equalization.

☞ keyword : Blind Equalization, PDF, Euclidian Distance.

1. INTRODUCTION

Multipoint communication has been an increasingly focused topic in computer communication networks, including the Internet, the ATM, and the wireless/mobile networks [1]

In applications such as broadcast and multipoint networks, blind equalizers to counteract multipath effects are very useful since they do not require a training sequence to start up or to restart after a communications breakdown [2][3].

Problems involving the training of adaptive equalizers have been developed through the use of information theoretic optimization criteria. As a way for solving these problems, information theoretic learning (ITL) has been introduced by Princepe [4]. This approach is to choose the parameters W of the mapping g(.) such that a figure of merit based on information theory is

* 정 회 원 : Prof. of Kanwon National Univ. Information & Communication Eng. namyong@kangwon.ac.kr

** 정 회 원 : Assistant Prof., Dept. of ECE in UF, U.S.A.

[2008/05/29 투고 - 2008/06/02 심사 - 2008/09/05 심사완료]

optimized at the output space of the mapper. ITL algorithms are based on a combination of a nonparametric probability density function (PDF) estimator and a procedure to compute entropy or information potential (IP). The difficulty in approximating Shannon's entropy is overcome by utilizing Renyi's generalized entropy. Estimating the data PDF nonparametrically is based on the Parzen window method using a Gaussian kernel. The combination of Renyi's quadratic entropy with the Parzen window leads to an estimation of entropy or information potential by computing interactions among pairs of output samples which is a practical cost function for ITL. This approach has been applied to Blind deconvolution of a linear channel by maximizing or minimizing the output entropy of an adaptive equalizer.

Maximum entropy deconvolution is based on the idea that when the nonlinear device output PDF is forced to uniform by maximzing its output entropy, the pdf of equalizer output matches the pdf that is the derivative of the nonlinearity, where the nonlinearity is selected to be the cdf of the source signal. On the other hand, minimum entropy deconvolution method is also available. Erdogmus and his coworkers show that the minimization of Renyi's entropy at the output of the equalizer can achieve Blind deconvolution [5]. But the global minimum of a minimum entropy Blind deconvolution criterion occurs for zero equalizer weights, corresponding to a zero equalizer output. This unwanted situation should be avoided by some measures such as modifying the cost function with the equalizer output variance.

In this paper we investigate the interactions among not only output samples but also ramdomly generated desired samples at the receiver by utilizing Euclidian distance (ED). It will be shown that this ITL algorithm can be applied to blind equalization on the condition that the randomly generated desired data have the same shape of PDF as that of the transmitted data.

2. ED BETWEEN TWO PDFS

Recently, Erdogmus introduced an information theoretic framework based on Kullback Leibler (KL) divergence [6] minimization for training adaptive systems in supervised learning settings using both labeled and unlabeled data [7].

The KL divergence is a way to estimate mutual information which is capable of quantifying the entropy between pairs of random variables. The KL divergence between two PDFs, f_x and f_y is:

$$KL[f_x, f_y] = \int f_x(\xi) \log[f_x(\xi)/f_y(\xi)]d\xi \,. \quad (1)$$

Since it is not quadratic in the PDFs, it can not be easily integrated with the information potential [4]. Based on the Euclidian distance, a new divergence measure between two PDFs [4] has been introduced which contains only quadratic terms to utilize the tools of information potential as

$$ED[f_x, f_y] = \int f_x^2(\xi)d\xi + \int f_y^2(\xi)d\xi - 2\int f_x(\xi) f_y(\xi)d\xi \,. \quad (2)$$

For equalization application, the Euclidian distance between f_d the transmitted symbols PDF and f_y the equalizer outputs PDF, can be minimized with respect to equalizer weight W as

$$\underset{W}{Min}\left(ED[f_d, f_y]\right) = \underset{W}{Min}\left(\int f_d^2(\xi)d\xi + \int f_y^2(\xi)d\xi\right.$$

$$-2\int f_d(\xi)f_y(\xi)d\xi\Big). \quad (3)$$

In other words, we create desired symbols for the input signal during training by utilizing the equalizer outputs PDF and the previously known PDF information of the transmitted symbols.

In this paper, we propose a method of minimizing the Euclidean distance based on Parzen PDFs which are computed directly from data samples. Computing ED directly from data samples requires also a continuous and differentiable estimator for the two probability density functions f_d and f_y. Parzen windowing is a suitable method, which is in general biased but the bias can be asymptotically reduced to zero by selecting an unimodal symmetric kernel function such as the Guassian and reducing the kernel size monotonically with increasing the number of samples. Selection of an optimal kernel size is one of the important steps in the Parzen windowing method but in this paper it will be left for future work.

For blind channel equalization, we assume here that the a priori probability f_d of transmitted symbols is known to the receiver but the exact training symbols are not available to the receiver. This assumption can be considered reasonable in most cases since the transmitter has a particular modulation scheme and the symbols are generally independent and identically distributed (i.i.d). as that of the transmitted data.

3. ED MINIMIZATION ALGORITHM FOR BLIND EQULIZATION

Given the randomly generated N independent and identically distributed (iid) symbols $\{d_1, d_2, d_3, ..., d_N\}$, the pdf can be approximated by

$$f_d(\xi) \cong \frac{1}{N}\sum_{i=1}^{N} G_\sigma(\xi - d_i) \quad (4)$$

where $G_\sigma(\cdot)$ is typically a zero mean Guassian kernel with standard deviation σ. If the symbols are generated randomly so as to match with the PDF of the transmitted symbols, the $f_d(\xi)$ in (4) can be considered the same as the PDF of the desired symbols. The point noticeable here is that for Parzen PDF calculation, instead of the exact training symbols, the randomly generated symbols are used at the receiver. For example, in case of bipolar transmission with equal probability, random numbers +1, 1 are generated equiprobably for Parzen PDF calculation but no exact desired symbols are used for it. The number of generated random numbers is the same as that of the output symbols to be used in cost function calculation. This makes blind equalization possible because exact desired symbols are not used.

One of the advantages of Parzen window with Gaussian kernel is that we can avoid the integral computation directly, since the integral of the two Gaussian kernels generates another Gaussian kernel with different double standard deviation.

$$\int f_d^2(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(d_j - d_i) \quad (5)$$

$$\int f_y^2(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(y_j - y_i) \quad (6)$$

$$\int f_d(\xi)f_y(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(d_j - y_i) \quad (7)$$

Equation (5) is the sum of interactions of all pairs of randomly generated symbols. This can be called IP of the set of randomly generated symbols, or $IP(d,d)$ in this paper. By summing the interactions among pairs of output samples we can obtain the $IP(y,y)$ of output samples as in (6). Equation (7), $IP(d,y)$, indicates the interactions between the two different variables. Equations (6) and (7) which contain system output are a function of weight but (5) is not a function of weight since.

d_j is the desired sample randomly generated

Now we derive a gradient descent method for the minimization of the cost function (3) with respect to equalizer weight W.

$$W_{new} = W_{old} - \mu \frac{\partial P}{\partial W}, \tag{8}$$

where $P = IP(y,y) - 2 \cdot IP(d,y)$.

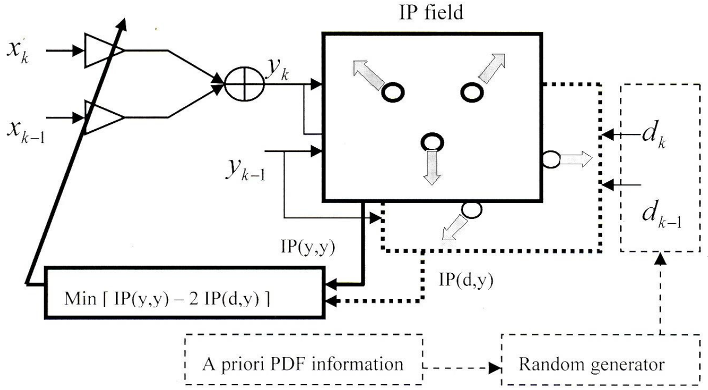
In case of on line linear equalization, a tapped delay line (TDL) can be used for input $X_k = [x_k, x_{k-1}, x_{k-2}, ..., x_{k-N+1}]^T$ and output $y_k = W_k^T X_k$ at time k (Fig. 1). The randomly generated desired symbols $D_N = \{d_1, d_2, ..., d_j, ..., d_N\}$ are used in the equalization process regardless of time k. Then the gradient is evaluated from

$$\frac{\partial P}{\partial W_k} = \frac{1}{2N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=k-N+1}^{k} (y_j - y_i) \cdot G_{\sigma\sqrt{2}}(y_j - y_i) \cdot (X_i - X_j) - \frac{1}{N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} (d_j - y_i) \cdot G_{\sigma\sqrt{2}}(d_j - y_i) \cdot X_i \tag{9}$$

This proposed method is referred to as minimum ED (MED) algorithm in this paper. The proposed blind equalizer structure for N=2 is shown in Fig. 1.

4. SIMULATION RESULTS

In this section we present and discuss simulation results that illustrate the comparative performance of the proposed MED algorithm versus constant modulus algorithm (CMA) in blind equalization for two linear channels. The 4 level



(Figure 1) Proposed structure for N=2.

random signal $\{\pm 3, \pm 1\}$ is transmitted to the channel and the impulse response, h_i of the channel model in [8] is

$$h_i = \frac{1}{2}\{1 + \cos[2\pi(i-2)/BW]\},\quad i = 1,2,3. \quad (10)$$

The parameter BW determines the channel bandwidth and controls the eigenvalue spread ratio (ESR) of the correlation matrix of the inputs in the equalizer [8].

Channel 1: BW =3.1, ESR=11.12,
Channel 2: BW =3.3, ESR=21.71.

The number of weights in the linear TDL equalizer structure is set to 11. The channel noise is zero mean white Gaussian for a SNR=30 dB. As a measure of equalizer performance, we use probability densities for errors of CMA and MED. The convergence parameters for CMA which have shown the lowest steady state MSE are 0.00001 and 0.0000005 for CH1 and CH2, respectively. For MED we used a data block size $N = 20$, a fixed kernel size $\sigma = 0.5$ and the convergence parameter $\mu = 0.007$.

To investigate the effect of orders on the performance, the randomly generated desired symbols $D_N = \{d_1, d_2, ..., d_j, ..., d_N\}$ are studied in 3 cases of different orders as described below;

Case1 :

$$d_j = \begin{cases} +3: j = 1,2,3,......................., N/4. \\ +1: j = N/4+1, N/4+2, ..., N/2. \\ -1: j = N/2+1, N/2+2, ..., 3N/4. \\ -3: j = 3N/4+1, 3N/4+2, ..., N. \end{cases}$$

Case 2 :

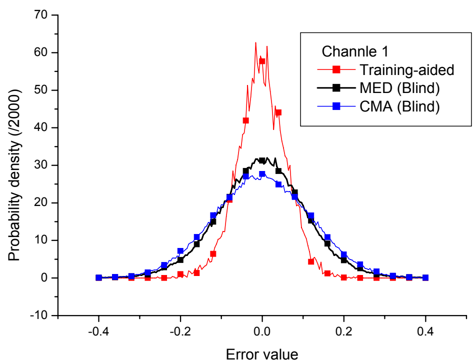
$$d_j = \begin{cases} +3: j = 1,5,9,..., N-3. \\ +1: j = 2,6,10,..., N-2. \\ -1: j = 3,7,11,..., N-1. \\ -3: j = 4,8,12,........, N. \end{cases}$$

Case 3 :

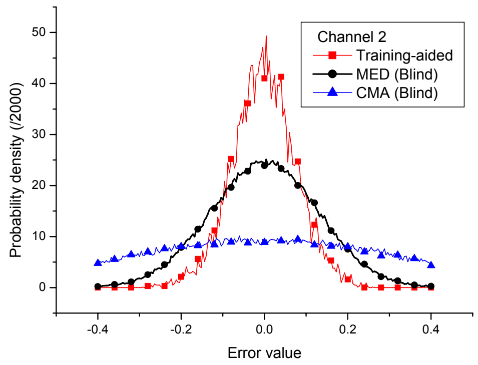
d_j = Randomly ordered $\{\pm 3, \pm 1\}$.

Where the symbols in all 3 cases are equiprobable.

We have studied the PDF of steady state errors of MED and CMA as a figure of merit. All three cases of different orders have been observed to show exactly the same performance in the result. This is readily understandable since all symbols of D_N are paired with all output samples and involved in the interaction calculations in (7) or (9). We see in Figs. 2 and 3 for error performance that increasing the ESR has the effect of increasing the steady state error of CMA. For comparison of the blind algorithms with the training aided method, the least mean square (LMS) algorithm with its convergence parameter $\mu = 0.005$ is used.



(Figure 2) Probability density for errors in channel 1.



(Figure 3) Probability density for errors in channel 2.

In case of channel 1 with ESR=11.12, the error performance of the MED has shown a slightly increased performance in comparison with CMA. In the severer channel model, channel 2, whose ESR is 21.71, CMA has shown severe performance degradation. On the other hand, the steady state error performance of MED has shown similar performance to that in channel 1, so the proposed MED can be considered relatively insensitive to ESR variations compared to CMA.

3. CONCLUSION

In this paper, an information theoretic unsupervised learning, namely, a criterion of minimizing ED between two PDFs criterion for adaptive blind equalizers has been proposed. In order for ED expressed with Parzen PDFs to be a cost function for blind equalization, we proposed to use a set of randomly generated desired symbols so that the PDF of the generated symbols matches that of the transmitted symbols. We have shown that using randomly generated symbols and minimizing ED which is a combination of sample interactions, the weight vector of the linear TDL equalizer can be successfully adjusted without training symbols. In both channel models, the error performance of MED has shown enhanced performance in comparison with CMA. In severe channel environment, MED has shown superior performance but CMA shows severe performance degradation. The error performances of MED have been observed to be similar in all simulation channel environments. This implies that the proposed MED can be considered relatively insensitive to ESR variations compared to CMA. Many challenging steps for better ideas lie ahead in this area of research, and we conclude that as a field of ITL, ED minimization using Parzen PDFs has shown possibilities of being successfully applied to blind equalization. In future work, it is considered that this method can be readily applicable to nonlinear equalizers such as weight adjustment in neural networks, and for QAM transmission, complex versions of this method should be studied. A research for reduced computational complexity is also needed for efficient implementation of blind equalization applications.

REFERENCES

[1] W. M. Moh, Y. Chen, "Multicasting flow control for hybrid wired/wireless ATM networks," Performance Evaluation, Vol. 40, Mar. 2000, pp. 161-194.

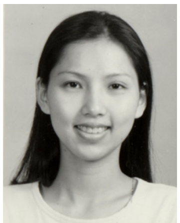
[2] L. M. Garth, "A dynamic convergence analysis of blind equalization algorithms," IEEE Trans. on Comm., Vol. 49, April. 2001, pp. 624-634

[3] F. Mazzenga, "Channel estimation and equalization for M QAM transmission with a hidden pilot sequence, "IEEE Trans. on

Broadcasting, Vol. 46, June. 2000, pp. 170-176
[4] J.C. Principe, D. Xu and J. Fisher, Information Theoretic Learning, in: S. Haykin, Unsupervised Adaptive Filtering, Wiley, New York, Vol. I, 2000, pp 265-319.
[5] D. Erdogmus, K. E. Hild, M. Lazaro, I. Santamaria, J. C. Principe, "Adaptive Blind Deconvolution of Linear Channels Using Renyi's Entropy with Parzen Estimation," Trans. on Signal Processing (Accepted).
[6] S. Kullback, Information Theory and Statistics, Dover Publications, New York, 1968
[7] D. Erdogmus, Y. Rao and J. C. Principe, "Supervised Training of Adaptive Systems with Partially Labeled Data" Proceedings of the International Conference on ASSP, Apr. 2005. pp. v321-v324.
[8] S. Haykin, Adaptive Filter Theory, Prentice Hall, Upper Saddle River, 4th edition, 2001.

◐ 저 자 소 개 ◑

김 남 용
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



LIUQING YANG (S'02-M'04-SM'06)
received her B.S. degree in Electrical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 1994, and her M.Sc. and Ph.D. degrees in Electrical and Computer Engineering from University of Minnesota, Minneapolis, in 2002 and 2004, respectively. Since August 2004, she has been an assistant professor with the Department of Electrical and Computer Engineering at University of Florida, Gainesville. Her current research interests include signal processing, communications theory and networking. Dr. Yang received the Best Dissertation Award in the Physical Sciences & Engineering from University of Minnesota in 2005, and the Office of Naval Research Young Investigator Award in 2007. She is the co-recipient of the Best Paper Award at IEEE international Conference on UWB (ICUWB) in 2006. She serves as an Associate Editor of the IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, and IEEE Transactions on Intelligent Transportation Systems. She is also a member of the Editorial Board of Signal, Image and Video Processing Journal, and PHYCOM: Physical Communication. Dr. Yang is a senior IEEE member, and has been the co-chair of the Mobile Communication Networks technical committee of the IEEE ITSS since 2006. She was the secretary of the IEEE Gainesville section from 2006 to 2007, and now serves as the vice chair.