랜덤 심볼을 사용한 최대 코렌트로피 기준의 블라인드 등화

종신회원 김 남 용*, 강 성 진**, 홍 대 기***

Blind Equalization based on Maximum Cross-Correntropy Criterion using a Set of Randomly Generated Symbol

Namyong Kim*, Sung-jin Kang**, Dae-ki Hong*** *Lifelong Members*

요 약

코렌트로피는 일반화된 상관함수로서 확률밀도함수의 고차 모멘트를 가지는데 이는 기존의 모멘트 확장 방식들보다 더 높은 고차 모멘트이다. 두 다른 랜덤 변수의 상호 코렌트로피를 최대화하는 이 기준 방식은 최소자승오차 기준 방식과 비교할 때, 비선형, 비 가우시안 신호 처리 환경에서 특히 탁월한 성능을 나타낸다. 이 논문에서는, 상호 코렌트로피 기준에 근거한 새로운 블라인드 등화 기법을 제안한다. 이 기법은 등화기 출력의 확률밀도함수와, 송신 심볼의 분포에 맞추어 발생시킨 랜덤심볼의 파전 확률밀도 추정치라는 두 확률변수에 상호 코렌트로피를 적용한다. 상호 코렌트로피에 근거한 제안 방식의 블라인드 등화 성능을 유클리디언 거리 최소화 방식과 비교하였다.

Key Words : Correntropy, MCC, Blind Equalizer, PDF, Euclidian Distance, Parzen Window

ABSTRACT

Correntropy is a generalized correlation function that contains higher order moments of the probability density function (PDF) than the conventional moment expansions. The criterion maximizing cross-correntropy (MCC) of two different random variables has yielded superior performance particularly in nonlinear, non-Gaussian signal processing comparing to mean squared error criterion. In this paper we propose a new blind equalization algorithm based on cross-correntropy criterion which uses, as two variables, equalizer output PDF and Parzen PDF estimate of a set of randomly generated symbols that complies with the transmitted symbol PDF. The performance of the proposed algorithm based on MCC is compared with the Euclidian distance minimization.

Ⅰ. Introduction

Blind equalizers are being effectively utilized in most communication systems in order to overcome inter-symbol interference induced by multipath effects. Since blind equalizers do not require a training sequence to start up or to restart after a communications breakdown[1],[2], it has been an increasingly focused topic in broadcasting systems, computer communication networks[3].

Usually the training of adaptive equalizers has been accomplished through the use of mean squared error (MSE) criterion that utilizes error energy. As a different approach to adjusting adap-

* 강원대학교 공학대학 전자정보통신공학부(namyong@kangwon.ac.kr), ** 한국기술교육대학교 정보기술공학부
*** 상명대학교 공과대학 정보통신공학과
논문번호 : KICS2009-09-428, 접수일자 : 2009년 9월 22일, 최종논문접수일자 : 2009년 1월 7일

tive system parameters, information-theoretic learning (ITL) method that has been introduced by Princepe[4] is based on a combination of a non-parametric probability density function (PDF) estimator and a procedure to compute information potential (IP) or entropy.

The estimation of entropy can be calculated by the combination of Renyi's quadratic entropy with the Parzen window method[4] using a Gaussian kernel by computing interactions among pairs of output samples[5]. As another ITL application which is different from the entropy minimization or maximization, Kullback-Leibler (KL) divergence[6] minimization has been studied for training adaptive systems in supervised learning settings using both labeled and unlabeled data[7]. The KL divergence is not quadratic in the PDFs, on the other hand, the Euclidian distance (ED) between two PDFs has been introduced as another measure of divergence[4]. The ED based PDF matching methods have been successfully applied to the classification problem with a real biomedical data set[8] and blind equalization for multipoint communication systems[9].

Entropy, divergence, or distance minimization criteria have led the ITL method to the extension of the fundamental definition of correlation function for random processes. The generalized correlation function is called correntropy[10] which contains higher order moments of the PDF and is much simpler to estimate directly from samples than conventional moments expansions. Auto-correntropy measures one random variable similarity across lags as the autocorrelation, but when averaged across lags, it becomes the entropy of the random variable, hence it has been named as correntropy. Cross-correntropy handles the general case of two arbitrary random variables and is directly related to the probability of how similar two random variables are in the neighborhood of the joint space controlled by the kernel size[11].

In this paper, we introduce the correntropy concept to blind equalization and compare the performance based on correntropy maximization criterion with the performance based on ED minimization criterion.

This paper is organized as follows. In Section II, we briefly describe blind equalization based on ED minimization method using randomly generated symbols at the receiver. The concept of cross-correntropy is introduced concisely in Section III. And then a new blind equalization algorithm based on maximization of cross-correntropy using randomly generated symbols is proposed in Section IV. Section V reports simulation results and discussions. Finally, concluding remarks are presented in Section VI.

Ⅱ. Blind Equalization based on ED Minimization using Randomly Generated Symbols at the Receiver

Unlike traditional trained equalization algorithms, many of the widely employed blind equalization algorithms employ nonlinearity at the equalizer output y_k to generate the error signal for weights updates based on MSE criterion. In case of on-line linear equalization, a tapped delay line (TDL) can be used for input $X_k = [x_k, x_{k-1}, ..., x_{k-L+1}]^T$, output $y_k = W_k^T X_k$, and desired symbol d_k at time k. We assume that M-ary PAM signaling systems are employed and the all M levels are equally likely to be transmitted a priori with a probability $1/M$, and the transmitted levels A_m takes the following discrete values

$$A_m = 2m - 1 - M \ , \ \ m = 1,2,...,M \qquad (1)$$

The Euclidian distance between the transmitted symbol PDF f_d and the equalizer output PDF f_y can be minimized with respect to weight W as

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

If the two distributions are close to each other, the ED cost function (2) minimizes the divergence between the desired symbols and equalizer output samples. In other words, we create desired symbols for the equalizer input signal by utilizing the PDF information of the transmitted symbols. Our initial idea was to generate, at the receiver, random symbols that have the same PDF of the transmitted symbols[9], which is introduced as follows.

Given a set of randomly generated N symbols $D_N = \{d_1, d_2, ..., d_N\}$, the PDF based on Parzen window method can be approximated by

$$f_d(\xi) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma(\xi - d_i)$$
$$= \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp[\frac{-(\xi - d_i)^2}{2\sigma^2}] \quad (3)$$

Under the assumption that all M levels are equally likely, the number of random symbols corresponding to each level A_m is N/M.

The point noticeable here is that the random symbols used for Parzen PDF calculation have the same PDF pattern as the transmitted symbols, but the symbols are randomly generated ones at the receiver, not the exact training symbols. This approach makes blind equalization possible. Then the integrals of the multiplication of two PDFs in (2) become

$$\int f_d^2(\xi)d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(d_j - d_i) \quad (4)$$

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

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

Equation (4) can be the IP of randomly generated symbols, which is denoted as $IP_1(d,d)$ in this paper, where the subscript 1 indicates the method 1. We note that (4) is not a function of weight. By summing the interactions among pairs of output samples we can obtain the $IP_1(y,y)$ as in (5). Equation (6) defined as $IP_1(d,y)$ indicates the interactions between the two different variables d and y. So the cost function can be reduced as

$$P_1 = IP_1(y,y) - 2 \cdot IP_1(d,y) \quad (7)$$

Now a gradient descent method can be applied for the minimization of the cost function (11) with respect to equalizer weight as follows:

$$W_{new} = W_{old} - \mu_1 \frac{\partial P_1}{\partial W} \quad (8)$$

The gradient is evaluated from

$$\frac{\partial P_1}{\partial W_k} = \frac{1}{2N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=k-N+1}^{k} (y_j - y_i)$$
$$\cdot G_{\sigma\sqrt{2}}(y_j - y_i) \cdot (X_i - X_j)$$
$$- \frac{1}{N^2\sigma^2} \sum_{i=k-N+1,}^{k} \sum_{j=1}^{N} (d_j - y_i) \cdot G_{\sigma\sqrt{2}}(d_j - y_i) \cdot X_i \quad (9)$$

This method[9] will be referred to here as minimum ED (MED) algorithm for convenience sake.

Ⅲ. Maximum Cross Correntropy Criterion and Cross Information Potential

Correntropy is a similarity measure that has the analogy with the autocorrelation of two random processes. Let a nonlinear mapping Φ transform the data to an infinite dimensional reproducing kernel Hilbert space F. Then the auto-correntropy function $V_X(t,s)$[10] for random process $X(t)$ is defined as

$$V_X(t,s) = E[< \Phi(X(t), \Phi(X(s)) >_F] \quad (10)$$

where $E[\cdot]$ and $<\cdot,\cdot>_F$ denote statistical expectation and inner product in F, respectively.

Cross-correntropy[11] is a generalized version of similarity measure between two scalar random variables X and Y defined by

$$V(X,Y) = E[<\Phi(X), \Phi(Y)>_F] \quad (11)$$

If two scalar random variables X and Y are statistically independent,

$$V(X,Y) = <E[\Phi(X)], E[\Phi(Y)]>_F \quad (12)$$

From the view point of kernel methods, $f_x(\xi) = E[\Phi(X)]$ and $f_y(\xi) = E[\Phi(Y)]$ are two points in the RKHS, and the cross-correntropy is very similar to the inner product between two PDFs, that is, $\int f_x(\xi) \cdot f_y(\xi) d\xi$. On the other hand, entropy is a scalar quantity that provides a measure for the average information contained in a given PDF. When error entropy is minimized, the error distribution of adaptive systems is concentrated. Renyi's quadratic error entropy which is effectively used in ITL methods is defined as

$$H(e) = -\log(\int f_E(\xi)^2 d\xi) \quad (13)$$

The term $\int f_E^2(\xi) d\xi$ in (13) is referred to as error information potential[12]. Obviously, minimizing the error entropy $H(e)$ is equivalent to maximizing the error information potential. This criterion is referred to as minimization of error entropy (MEE)[12]. From the perspective of information potential, the inner product between two independent PDFs, $\int f_x(\xi) \cdot f_y(\xi) d\xi$ can be defined as cross information potential (CIP)[11]. Therefore, maximizing cross correntropy of two independent PDFs is equivalent to maximization of cross information potential (or the inner product of two independent PDFs).

In the following section, we apply the maximization of cross-correntropy (MCC) concept to blind equalization by maximizing the similarity between the Parzen PDF of equalizer output and the Parzen PDF of randomly generated symbols that matches with the PDF of transmitted symbols.

Ⅳ. Blind Equalization based on MCC Criterion using Randomly Generated Symbols at the Receiver

Normally, modulation schemes are known to receivers. Furthermore most transmitters use independent and identically distributed symbols. Under these considerations, we propose to employ MCC criterion in blind equalization using a set of randomly generated desired symbols.

As described in Section III with a set of randomly generated N symbols $D_N = \{d_1, d_2, ..., d_N\}$, where the number of random symbols corresponding to each level A_m is N/M, the desired PDF based on Parzen window method can be estimated by

$$f_d(\xi) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma(\xi - d_i) \quad (14)$$

Then, the cross information potential $CIP(d,y)$ between the desired PDF and output PDF becomes

$$CIP(d,y) = \int f_d(\xi) f_y(\xi) d\xi = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} G_{\sigma\sqrt{2}}(d_j - y_i) \quad (15)$$

The gradient is calculated from

$$\frac{\partial CIP(d,y)}{\partial W} = \frac{1}{N^2\sigma^2}\sum_{i=1,}^{N}\sum_{j=1}^{N}(d_j - y_i)\cdot G_{\sigma\sqrt{2}}(d_j - y_i)\cdot X_i \quad (16)$$

In on-line systems on a sample-by-sample basis, we can use a small sliding window and then the gradient is evaluated from

$$\frac{\partial CIP(d,y)}{\partial W_k} = \frac{1}{N^2\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{j=1}^{N}(d_j - y_i)\cdot G_{\sigma\sqrt{2}}(d_j - y_i)\cdot X_i \quad (17)$$

Finally the weight update of the proposed method can be obtained as

$$W_{k+1} = W_k + \mu\frac{1}{N^2\sigma^2}\sum_{i=k-N+1,}^{k}\sum_{j=1}^{N}(d_j - y_i)\cdot G_{\sigma\sqrt{2}}(d_j - y_i)\cdot X_i \quad (18)$$

where d_j is one of the elements of D_N. We will refer to this as MCC algorithm in this paper.

It is noticeable in implementation aspect that MED in (9) is computationally cumbersome due to the $O(N^2)$ complexity but that of the proposed algorithm (18) requires decreased complexity by half of the complexity $1/2\cdot O(N^2)$.

V. Simulation Results and Discussion

In the previous work[9], the MED algorithm as a new ITL-based blind equalization technique has shown much better performance than the conventional constant modulus algorithm. In this section we present and discuss simulation results that compare performance of the proposed algorithm with MED algorithm in blind equalization for two linear channels used in [13]. The transmitted signal is composed of 4 level random symbols $\{\pm 3,\pm 1\}$ and the channel impulse response is $h_i = \frac{1}{2}\{1+\cos[2\pi(i-2)/BW]\}$, $i = 1,2,3$. For channel 1, BW =3.1 is used and for channel 2, BW =3.3. The number of taps in the linear TDL equalizer structure is set to L = 11. The zero mean white Gaussian noise is added to the channel output, and its variance is 0.001. As a figure of merit of the algorithms, we use probability density for equalizer errors at that noise variance and MSE learning curves of MED and the proposed algorithm. For MED and the proposed algorithm we set the data-block size $N = 20$, and the convergence parameter $\mu = 0.007$ for both channel 1 and channel 2. The kernel size is chosen as $\sigma = 0.5$ for MED and the proposed algorithm. All the convergence parameters are obtained when the algorithms show the lowest steady-state error performance. For channel 1 in Fig. 1, MSE learning curve of MED algorithm converges in about 6000 samples but the proposed method converges in 2500 samples. Furthermore, the proposed algorithm shows significantly lower minimum MSE by above 5 dB. In the case of channel 2, MED converges in the vicinity of 9000 samples but the proposed method converges in about 3500 samples. Similar to the case of channel 1, the proposed algorithm has 5 dB performance enhancement in steady state MSE. The comparison of probability density for equalizer errors in Fig. 2 reveals that the proposed method has more concentrated distributions to zero for both channel models.

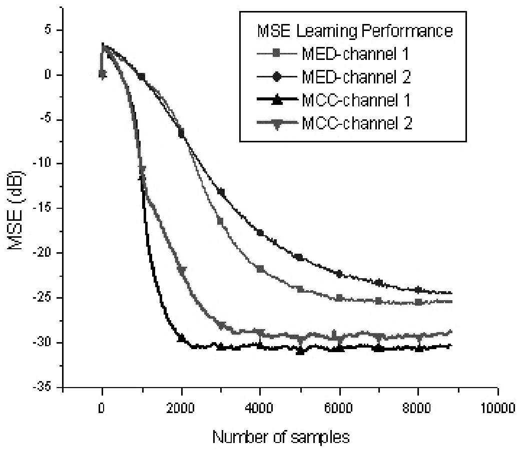


Fig. 1. MSE performance comparison of the proposed and MED algorithm for channel 1 and channel 2

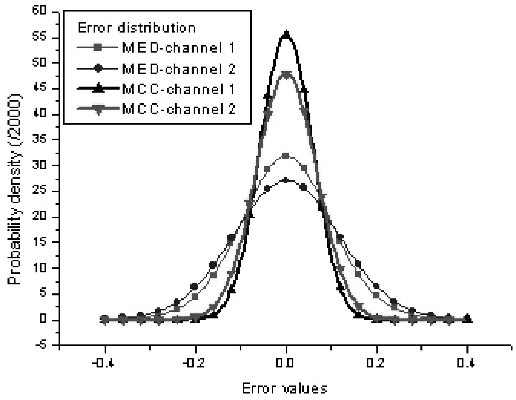


Fig. 2. Probability density for error comparison of the proposed and MED algorithm for channel 1 and channel 2

Ⅵ. Conclusions

Cross-correntropy is a generalized similarity measure of two PDFs. Cross-correntropy can be formulated as the inner product between two independent PDFs. In this paper, we applied the maximization of cross-correntropy concept to blind equalization by maximizing the similarity between the Parzen PDF of equalizer output and the Parzen PDF of randomly generated symbols that matches the PDF of transmitted symbols. Simulation results of comparison between the proposed and MED revealed that the proposed method has more concentrated distributions to zero for channel models, which implies that the equalizer output value satisfyingly approaches the exact transmitted symbol by the proposed blind method based on maximization of cross-correntropy criterion. This makes us conclude that the proposed blind method can be a good candidate for blind equalization applications. Also we can expect that MCC-based adaptive learning will yield superior performance in adaptive signal processing applications.

In future work, research for reduced computational complexity should be carried out for efficient implementation.

References

〔1〕 L. M. Garth, "A dynamic convergence analysis of blind equalization algorithms," IEEE Trans. on Comm., Vol.49, pp.624-634. April, 2001.
〔2〕 F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," IEEE Trans. on Broadcasting, Vol.46, pp.170-176, June, 2000.
〔3〕 W. M. Moh and Y. Chen, "Multicasting flow control for hybrid wired/wireless ATM networks," Performance Evaluation, Vol.40, pp.161-194, Mar., 2000.
〔4〕 J. C. Principe, D. Xu and J. Fisher, Information Theoretic Learning in: S. Haykin, Unsupervised Adaptive Filtering, Wiley, (New York, USA), pp.265-319, 2000.
〔5〕 L. Wasserman, All of Statistics: A Concise Course in Statistical Inference. Springer Texts in Statistics, 2005.
〔6〕 S. Kullback, Information Theory and Statistics, Dover Publications. (New York, USA), 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. pp.v321-v324, 2005.
〔8〕 K. H. Jeong, J. W. Xu, D. Erdogmus, and J. C. Principe, "A new classifier based on information theoretic learning with unlabeled data," Neural Networks, 18, pp.719-726, 2005.
〔9〕 N. Kim and L. Yang, "A New Criterion of Information Theoretic Optimization and Application to Blind Channel Equalization," Journal of Korean Society for Internet Information, Vol.10, No.1, pp.11-17, Feb., 2009.
〔10〕 I. Santamaria, P. P. Pokharel, and J. C. Principe, "Generalized Correlation Function: Definition, Properties, and Application to Blind Equalization," Trans. on Signal Processing, Vol.54, No.6, pp.2187-2198, Jun., 2006.

〔11〕 Weifeng Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and Applications in Non-Gaussian Signal Processing," IEEE Trans. Signal Processing, Vol.55, No.11, pp.5286-5298, Nov., 2007.

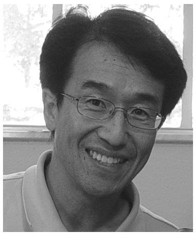
〔12〕 D. Erdogmus, and J.C. Principe, "An Entropy Minimization algorithm for Supervised Training of Nonlinear Systems," IEEE Trans. Signal Processing, Vol.50, pp.1780-1786, July, 2002.

〔13〕 S. Haykin, Adaptive Filter Theory. Prentice Hall. (Upper Saddle River), 4th edition, 2001.

강 성 진 (Sung-Jin Kang) 종신회원

1992년 연세대학교 전자공학과 공학사
1992년 연세대학교 전자공학과 공학석사
1992년 연세대학교 전자공학과 공학박사
1998년~2000년 ETRI무선방송기술연구소
2000년~2002년 (주)이노텔리텍
2002년~2007년 전자부품연구원 통신네트워크연구센터
2007년~현재 한국기술교육대학교 정보기술공학부 조교수
<관심분야> WPAN/WLAN, MODEM SoC

김 남 용 (Namyong Kim) 종신회원



1986년 2월 연세대학교 전자공학과 학사
1988년 2월 연세대학교 전자공학과 석사
1991년 8월 연세대학교 전자공학과 박사
1992년 8월~1998년 2월 관동대학교 전자통신공학과 부교수
1998년 3월~현재 강원대학교 공학대학 정보통신공학과 교수
<관심분야> Adaptive equalization, RBFN algorithms, ITL algorithms, Odor sensing systems

홍 대 기 (Dae-Ki Hong) 종신회원

1997년 광운대학교 컴퓨터공학과 공학사
1999년 연세대학교 전자공학과 공학석사
2003년 연세대학교 전자공학과 공학박사
2003년~2006년 전자부품연구원 통신네트워크 연구센터
2006년~현재 상명대학교 공과대학 정보통신공학과 조교수
<관심분야> 무선통신, 이동통신, WPAN/WLAN