A New Constant Modulus Algorithm based on Maximum Probability Criterion

Namyong Kim* ***Regular Member***

ABSTRACT

In this paper, as an alternative to constant modulus algorithm based on MSE, maximization of the probability that equalizer output power is equal to the constant modulus of the transmitted symbols is introduced. The proposed algorithm using the gradient ascent method to the maximum probability criterion has superior convergence and steady-state MSE performance, and the error samples of the proposed algorithm exhibit more concentrated density functions in blind equalization environments. Simulation results indicate that the proposed training has a potential advantage versus MSE training for the constant modulus approach to blind equalization.

Key Words : Maximum probability criterion; CMA; Blind equalizer; ITL.

Ⅰ. Introduction

Recently, in computer communication networks, including the Internet, the ATM, and the wireless/mobile networks, multipoint communication has been an increasingly focused topic [1]. In those applications, 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 approached 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]. Unlike the mean square error (MSE) criterion that utilizes error energy, ITL method is based on a combination of a nonparametric probability density function (PDF) estimator and a procedure to compute entropy. As a robust ITL algorithm, minimization of error entropy (MEE) has been developed by Principe, Erdogmus and co-workers [5]. 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. ITL-type methods have shown superior performance to MSE-type methods in supervised channel equalization applications[6] and blind channel equalization applications[7]. This implies that approaches using a nonparametric probability density function (PDF) estimator can be an alternative to MSE criterion.

In this paper, instead of being based on MSE criterion for CMA blind equalization[8], we propose to maximize the probability that equalizer output power is equal to the constant modulus of the transmitted symbols, and develop a new algorithm by applying the gradient ascent method to the maximum probability criterion. The proposed algorithm has a faster speed of convergence and lower same steady-state means squared error in comparison with CMA based on MSE criterion. This paper is organized as follows. In Section II, we briefly describe the constant modulus algorithms which is based on MSE criterion. In Section III, the MED algorithm for blind equalization is introduced that minimizes the

* 강원대학교 정보통신공학과 (namyong@kangwon.ac.kr)
논문번호 : Kics2008-12-533, 접수일자 : 2008년 12월 1일, 최종논문접수일자: 2009년 2월 2일

Euclidian distance between two PDFs of randomly generated desired symbols and equalizer output symbols. As a different approach, we derive maximum PDF criterion and propose a gradient accent algorithm based on the maximum PDF criterion in Section IV. Simulation results are presented in Section V, and conclusions are drawn in Section VI.

Ⅱ. CMA based on MSE Criterion[5]

In case of linear equalization, a tapped delay line (TDL) with N taps can be used for input $X_k = [x_k, x_{k-1}, x_{k-2}, \ldots, x_{k-N+1}]^T$ and output $y_k = W_k^T X_k$, where W_k is the weight vector at time k. For training-aided equalization, error sample e_k between the desired training symbol d_k and output y_k are produced by $e_k = d_k - y_k = d_k - W_k^T X_k$. Channel equalization without the aid of a training sequence d_k is referred to as blind channel equalization. Unlike traditional trained equalization algorithms, many of the blind equalization algorithms employ non-linearity at the equalizer output y_k to generate the error signal for weights updates based on mean squared error (MSE) criterion. One of the well known blind equalization algorithms is constant modulus algorithm (CMA) which minimizes the following cost function[5].

$$P_{CMA} = E[(|y_k|^2 - R_2)^2], \quad (1)$$

where $R_2 = E[|d_k|^4]/E[|d_k|^2]$.

The minimization of P_{CMA} with respect to the equalizer coefficients can be performed recursively according to the steepest descent method,

$$W_{new} = W_{old} - \mu_{CMA} \cdot \frac{\partial P_{CMA}}{\partial W}, \quad (2)$$

where μ_{CMA} is the step-size parameter. By differentiating P_{CMA} and dropping the expectation operation we obtain the following LMS-type algorithm for adjusting the blind equalizer coefficients:

$$W_{k+1} = W_k - 2\mu_{CMA} \cdot X_k^* \cdot y_k \cdot (|y_k|^2 - R_2). \quad (3)$$

Ⅲ. Minimum Euclidian Distance (MED) Criterion and Related Algorithms for Blind Equalization

The MED algorithm that has developed recently for blind equalization utilizes the interactions between two information potentials that come out in the process of minimizing the Euclidian distance between two PDFs of randomly generated desired symbols and equalizer output symbols[7]. There exist two information potentials, one is the information potential $IP(y,y)$ of output samples y_k, another is information potential $IP(d,y)$ that is induced from the interactions between output samples y_k and randomly generated desired symbols d_k. A block of randomly generated desired symbols $D_M = \{d_1, d_2, \ldots, d_j, \ldots, d_M\}$ is used in the equalization process regardless of time k. The notation M denotes the number of randomly generated symbols. Then MED cost function P_{MED} becomes

$$P_{MED} = IP(y,y) - 2 \cdot IP(d,y). \quad (4)$$

To minimize the cost function P_{MED}, a gradient descent method is employed with respect to equalizer weight W.

$$W_{k+1} = W_k - \mu_{MED} \frac{\partial P_{MED}}{\partial W_k}. \quad (5)$$

The gradient becomes

$$\frac{\partial P_{MED}}{\partial W_k} = \frac{1}{2M^2\sigma^2} \sum_{i=k-M+1,}^{k} \sum_{j=k-M+1}^{k} (y_j - y_i) \cdot G_{\sigma\sqrt{2}}(y_j - y_i) \cdot (X_i - X_j)$$
$$- \frac{1}{M^2\sigma^2} \sum_{i=k-M+1,}^{k} \sum_{j=1}^{M} (d_j - y_i) \cdot G_{\sigma\sqrt{2}}(d_j - y_i) \cdot X_i, \quad (6)$$

where $G_\sigma(\cdot)$ is typically a zero-mean Guassian kernel with standard deviation σ.

Ⅳ. Maximum PDF Criterion and the Proposed Algorithm for Blind Equalization

In this section we develop a new algorithm that tries to create a concentration of constant modulus error $e_{CMA} = |y_k|^2 - R_2$ near zero using a different approach from the CMA. Aiming at this goal, we propose to maximize the following PDF f_E.

$$\max_{W} f_E(e_{CMA} = 0) \quad (7)$$

To obtain $f_E(\cdot)$ non-parametrically, we need the Parzen estimator [4] using Gaussian kernel as follows

$$f_X(x) \cong \frac{1}{M}\sum_{i=1}^{M} G_\sigma(x - x_i) \quad (8)$$

Inserting $e_{CMA} = |y_k|^2 - R_2$ into (8) and using a block of past output samples $Y_k = \{y_k, y_{k-1}, \ldots, y_{k-M+1}\}$, we have

$$f_E(e_{CMA}) = \frac{1}{M}\sum_{i=0}^{M-1} G_\sigma(e_{CMA} - [|y_{k-i}|^2 - R_2])$$
$$= \frac{1}{M}\sum_{i=0}^{M-1} \frac{1}{\sigma\sqrt{2\pi}} \exp[\frac{-(e_{CMA} - [|y_{k-i}|^2 - R_2])^2}{2\sigma^2}] \quad (9)$$

Letting e_{CMA} be zero, the probability $f_E(e_{CMA} = 0)$ reduces to

$$f_E(e_{CMA} = 0) = \frac{1}{M}\sum_{i=0}^{M-1} G_\sigma(-[|y_{k-i}|^2 - R_2]) \quad (10)$$

Now we derive a gradient ascent method for the maximization of the cost function (5).

$$W_{k+1} = W_k + \mu \cdot \frac{2}{\sigma^2 M}\sum_{i=0}^{M-1} G_\sigma(|y_{k-i}|^2 - R_2)$$
$$\cdot (R_2 - |y_{k-i}|^2) \cdot y_{k-i} \cdot X_{k-i}^*, \quad (11)$$

where μ is the step-size for convergence control of the proposed algorithm .

We assume that L-ary PAM signaling systems are employed and the all L levels are equally likely to be transmitted a priori with a probability $1/L$, and the transmitted levels A_l takes the following discrete values

$$A_l = 2l - 1 - L, \quad l = 1, 2, \ldots, L. \quad (12)$$

Then the constant modulus R_2 becomes

$$R_2 = E[|A_l|^4] / E[|A_l|^2]. \quad (13)$$

To verify that (10) plays the role of a cost function in adaptive equalization and it has a single maximum (global optimum) with respect to the equalizer weights, the value of the cost function was tested for a simple case with an all pass channel having a delta function $\delta(k)$ as its impulse response and an equalizer with a single weight. In Fig. 1, the normalized cost surface is plotted with varying weight values.

Fig. 1 shows that the proposed cost function has a single maximum point, so it justifies that the proposed probability maximization criterion can be a reasonable choice for blind equalization.

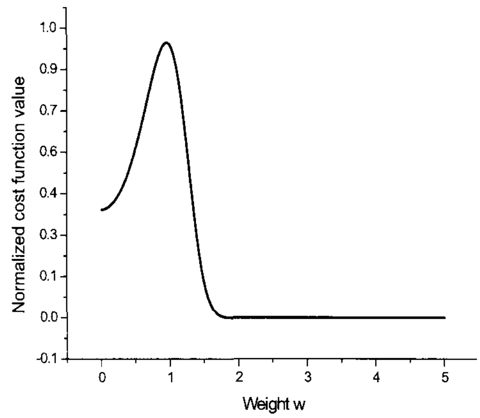

Fig. 1. Normailzed cost function output as function of a single equalizer weight

Ⅴ. Simulation Results and Discussion

In this section the comparative performance of the proposed, MED, and CMA in blind equalization is presented for two linear channels and simulation results are discussed. The 4 level ($L=4$) random signal (4-ary PAM) is transmitted to the channel and the impulse response, h_i of the channel model in [9] is

$$h_i = \frac{1}{2}\{1+\cos[2\pi(i-2)/BW]\},\quad i=1,2,3. \qquad (14)$$

The parameter BW determines the channelbandwidth and controls the eigenvalue spread ratio (ESR) of the correlation matrix of the inputs in the equalizer[9].

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 N=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 (the difference between the actual transmitted symbol and the output) of CMA, MED and the proposed. The convergence parameters for CMA which have shown the lowest steady-state MSE are 0.00001 and 0.0000005 for CH1 and CH2, respectively. We used a data-block size $M=20$ for ITL algorithms. For MED, the kernel size $\sigma=0.5$ in (6) and the convergence parameter $\mu_{MED}=0.007$ is used. For the proposed algorithm, we use the kernel size $\sigma=6$ in (11) (s=6 in Fig. 7) and the convergence parameter $\mu=0.02$ (u=0.02 in Fig. 6). Performance comparison for the varying parameter values for the proposed is presented later in Fig. 6 and 7.

We have studied the PDF of steady-state errors and MSE learning curves of the proposed, MED and CMA as a figure of merit. We see in Figs. 2 and 4 for error-performance that increasing the ESR has the effect of increasing the steady-state error of CMA.

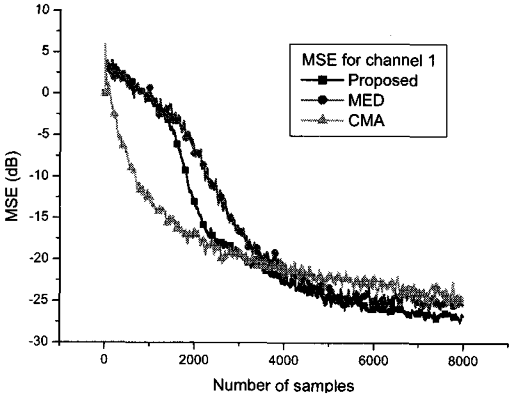

Fig. 2. MSE convergence performance in channle 1.

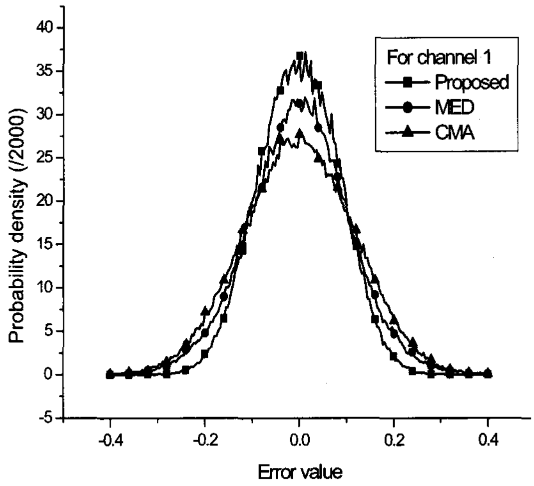

Fig. 3. Probability density for errors in channel 1.

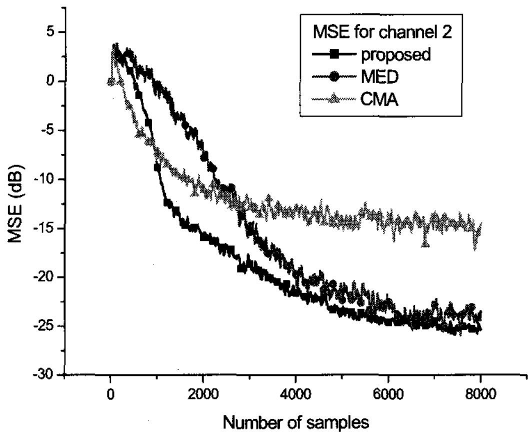

Fig. 4. MSE convergence performance in channle 2.

In case of channel 1 with ESR=11.12, the error-performance in Fig. 2 show that the proposed has a slightly enhanced performance in comparison with MED and CMA. In Fig. 3, the error PDF estimates are shown. Clearly, the error distribution of the proposed is more concentrated around zero. In the severer channel model, channel 2, whose ESR is 21.71, CMA shows severe performance degradation in Fig. 4. On the other hand, the steady-state error-performance of the proposed and MED show similar performance to that in channel 1, so the ITL-type algorithms

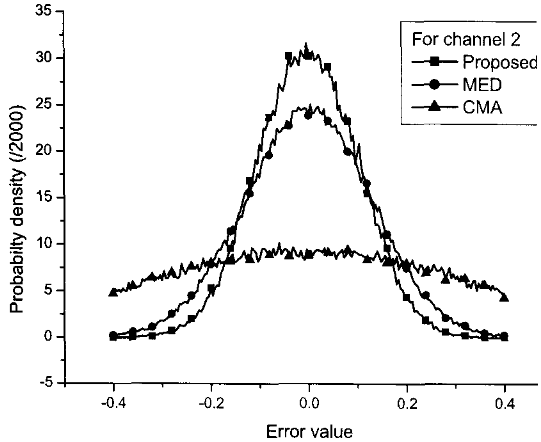

Fig. 5. Probability density for errors in channel 2.

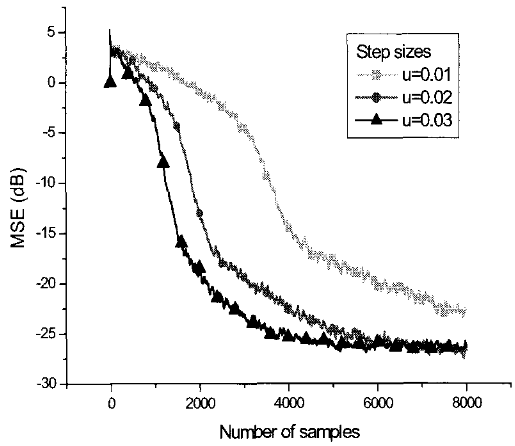

Fig. 6. MSE convergence comparison with varying step-size of the proposed algorithm where u= μ in (11).

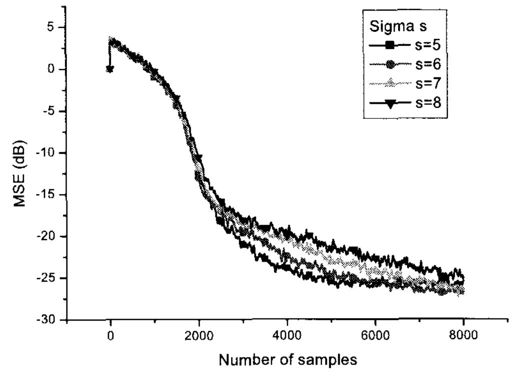

Fig. 7. MSE convergence comparison with varying kernel-size s of the proposed algorithm where s= σ in (11)

can be considered relatively insensitive to ESR variations compared to CMA based on MSE criterion. Fig. 5 depicts the estimated probability densities of the algorithms in channel 2. Their performance differences are shown more clearly. The error values of CMA appear not to be gathered well around zero, but the proposed and MED produce error distribution still concentrated around zero and especially the proposed algorithm yields superior error PDF performance.

To choose proper parameters for the proposed, MSE performance comparison for the varying parameter values is presented in Fig. 6 and 7.

In Fig. 6, clearly, small step-size makes the performance slow and large step-size induces fast learning speed but increased minimum MSE. In this simulation, 0.02 is chosen as the best step-size. The kernel size of the proposed algorithm is also shown to have effect on MSE convergence speed and steady state MSE value. Large kernel size (s=7 or 8 in Fig. 7) decreases the convergence speed and small kernel size (s=5 in Fig. 7) of the proposed algorithm reveals to increase steady state MSE value. As the best kernel size, 6 is chosen from simulation, and it is noticeable that a more in-depth research on the effect of kernel size on the proposed algorithm is needed.

Ⅵ. Conclusions

In this paper, as an alternative to constant modulus algorithm based on MSE criterion for blind equalization, we have proposed to maximize the probability that equalizer output power is equal to the constant modulus of the transmitted symbols, and develop a new algorithm by applying the gradient ascent method to the maximum probability criterion. The proposed criterion has a global maximum showing a single maximum point with respect to varying equalizer weight values. This enables us to use PDF maximization for blind equalization. The optimal performance obtained by CMA based on MSE,

MED and the proposed based on information theoretic learning have been compared in terms of MSE learning curves and the error distributions for two different multipath channels. These analyses demonstrated that the proposed algorithm has a faster speed of convergence, lower same steady-state means squared error in comparison with CMA based on MSE criterion and the error samples of the proposed algorithm exhibit a more concentrated density function. These results indicate the potential advantage of nonparametric PDF training versus MSE training as to the constant modulus approach to blind equalization.

In future work, it is considered for applications to QAM to be studied. A research for reduced computational complexity of the proposed method 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, pp.161-194, Mar. 2000.

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

〔3〕 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.

〔4〕 J.C. Principe, D. Xu and J. Fisher, Information Theoretic Learning, in: S. Haykin, Unsupervised Adaptive Filtering, Wiley, New York, Vol.I, pp.265-319, 2000.

〔5〕 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.

〔6〕 D. Erdogmus, Y. Rao and J. C. Principe, "Supervised Training of Adaptive Systems with Partially Labeled Data," Proceedings of the International Conference on ASSP, pp.v321-v324, Apr. 2005.

〔7〕 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 (to be published).

〔8〕 J. R. Treichler and B. Agee. 1983. "A new approach to multipath correction of constant modulus signals," IEEE Trans. Acoust., Speech, Signal Process. Vol.ASSP-31, pp.349-372, Nov. 1983.

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

김 남 용 (Namyong Kim) 정회원

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 정보통신공학과 교수

<관심분야> Adaptive equalization, RBFN algorithms, Odor sensing systems