Theoretical Derivation of Minimum Mean Square Error of RBF based Equalizer

Jung-Sik Lee* ***Regular Member***

ABSTRACT

In this paper, the minimum mean square error(MSE) convergence of the RBF equalizer is evaluated and compared with the linear equalizer based on the theoretical minimum MSE. The basic idea of comparing these two equalizers comes from the fact that the relationship between the hidden and output layers in the RBF equalizer is also linear. As extensive studies of this research, various channel models are selected, which include linearly separable channel, slightly distorted channel, and severely distorted channel models. In this work, the theoretical minimum MSE for both RBF and linear equalizers were computed, compared and the sensitivity of minimum MSE due to RBF center spreads was analyzed. It was found that RBF based equalizer always produced lower minimum MSE than linear equalizer, and that the minimum MSE value of RBF equalizer was obtained with the center spread which is relatively higher(approximately 2 to 10 times more) than variance of AWGN. This work provides an analytical framework for the practical training of RBF equalizer system.

Key Words : equalizer, linear channel, RBF, neural network

I. Introduction

The most widely known equalizer is an adaptive linear transversal equalizer, in which the output signal is compared to the expected signal and the tap coefficients are updated in accordance with the error between the desired signal and actual filter output. For more than decade, there has been much attention given to applying neural networks to the digital communication areas, including channel equalization[9].

Multilayer perceptrons(MLP) equalizer is able to equalize non-minimum phase channels without the introduction of any time delay; and it is less susceptible than a linear equalizer to the effects of high levels of additive noise[1-3]. However, the network architecture and training algorithm of the MLP equalizer is much more complex than the linear equalizer. Also, the RBF network has received a great deal of attention by many researchers because of its structural simplicity and more efficient learning[4-8]. Chen et al. applied RBF network to channel equalization problem to get the optimal Bayesian solution[4]. Although many of studies, mentioned above, claims that RBF based equalizers are superior to conventional linear equalizer due to both RBF network's structural linearity(or simplicity) and efficient training, none of them tried to compare those two from the theoretical minimum mean square error(MSE) point of view. The basic idea of comparing these two equalizers comes from the fact that once input domain of RBF network is transformed to another domain through Gaussian basis functions, the relationship between the hidden and output layers in the RBF equalizer is also linear.

In this paper, the theoretical minimum mean square error(MSE) for various channels for both RBF and the linear equalizers were evaluated and compared. Also, the sensitivity of minimum MSE due to RBF center spreads was analyzed.

※ This work was financially supported by the Kunsan National University's Long-term Overseas Research Program for Faculty Member in the year 2004.

* School of Electronics & Information Eng., Kunsan National University(leejs@kunsan.ac.kr)

논문번호 : KICS2006-05-242, 접수일자 : 2006년 5월 29일, 최종논문접수일자 : 2006년 8월 11일

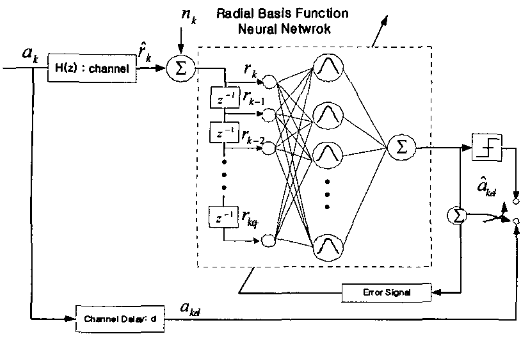


Fig. 1. The block diagram of RBF equalizer

II. Background of RBF Equalizer

A digital information source is to transmit independent and equi-probable binary symbols, designated as a_k, that are either -1 or +1. Dispersion in the digital channel may be represented by the transfer function

$$H(z) = h_0 + h_1 z^{-1} + ... + h_d z^{-d} + ... + h_p z^{-p} \qquad (1)$$

where p is the channel order, and d is the channel delay. The channel output at the time, k, may be written

$$r_k = h_0 a_k + h_1 a_{k-1} + ... h_d a_{k-d} ... + h_p a_{k-p} + n_k = \hat{r}_k + n_k \qquad (2)$$

where the vector $\mathbf{a}_k = [a_k, a_{k-1}, ..., a_{k-p}]$ is a length $p+1$ sequence of the transmitted data which affect the kth decision, and n_k is additive zero-mean white Gaussian noise(AWGN), assumed to be independent from one decision to the next. For a conventional RBF equalizer, the input is a sequence of the channel outputs, or a vector

$$\mathbf{r}_k = [r_k, r_{k-1}, ..., r_{k-q}] \qquad (3)$$

where, q is the RBF equalizer order(the number of tap delay elements in RBF equalizer). A radial bans function is a three-layer neural network structure as shown in Fig.1

The output layer forms a weighted sum of the outputs from the internal hidden layer. Each node in the hidden layer is a RBF center vector with dimension $q+1$. The following equation is the output of RBF equalizer.

$$F = \sum_{i=1}^{M} w_i \exp\left(\|\mathbf{r}_k - \mathbf{g}_i\|^2 \Big/ 2\sigma^2 \right) \qquad (4)$$

where M is the number of RBF centers, the w_i are the output layer weights, the $\mathbf{g}_i$ are Gaussian basis function center vectors, and the σ^2 are RBF center spread parameters. Clustering techniques are commonly used to determine the desired RBF centers from the number of noisy centers[4].

III. Derivation of Minimum Mean Square Errors of RBF Equalizer

The main objective of this paper is to derive the theoretical minimum mean square error of RBF based equalizer and compare it with that of the linear equalizer(finite transversal filter). The basic idea of comparing these two equalizers comes from the fact the relationship between the hidden and output layers in the RBF equalizer is also linear. To do this, first the general theroy of minimum MSE for linear equalizer was reviewed[10-11]. The theoretical minimum mean square error of linear filter is as follows:

$$\xi_{\min} = E\left[{a_{k-d}}^2 \right] - \alpha^{*T} \Phi^{-1} \alpha \qquad (5)$$

where Φ and α are autocorrelation and cross-correlation matrix. Then, the theoretical minimum MSE for RBF based equalizer was derived using equations (5) and finally compared with linear equalizer. As shown in[4], the maximum number of Gaussian centers in the RBF equalizer is

$$M = 2^{p+q+1} \qquad (6)$$

Accordingly, M number of Gaussian basis function outputs are generated from the hidden layer in the RBF equalizer systems.

$$\left.\begin{aligned} F_i^k &= \exp\left[-\frac{\|\mathbf{r}_k - \mathbf{g}_i\|^2}{2\sigma^2} \right] \\ \mathbf{g}_i &= \left[g_{i,0}, g_{i,1}, ..., g_{i,q} \right]^T, i = 1, 2, ..., M \end{aligned}\right\} \qquad (7)$$

where, F_i^k, and $\mathbf{g}_i$ denote the Gausssian basis function output in the i th hidden unit of the RBF equalizer and the desired Gaussian basis function vector. Then the autocorrelation matrix based on Gaussian basis function output is represented as

$$\Phi = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,M} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,M} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \phi_{M,1} & \phi_{M,2} & \cdots & \phi_{M,M} \end{bmatrix}_{M \times M} \quad (8)$$

where

$$\phi_{i,j} = E\left[F_i^k . F_j^k\right] = E\left[e^{-\frac{\left(\|\mathbf{r}_k - \mathbf{g}_i\|^2 + \|\mathbf{r}_k - \mathbf{g}_j\|^2\right)}{2\sigma^2}} \right], i, j = 1, 2, ..., M \quad (9)$$

The cross-correlation matrix between the desired symbol and the basis function outputs is

$$\alpha = E\begin{bmatrix} a_{k-d} . e^{-\frac{\|\mathbf{r}_k - \mathbf{g}_1\|^2}{2\sigma^2}} \\ a_{k-d} . e^{-\frac{\|\mathbf{r}_k - \mathbf{g}_2\|^2}{2\sigma^2}} \\ \cdot \\ \cdot \\ a_{k-d} . e^{-\frac{\|\mathbf{r}_k - \mathbf{g}_M\|^2}{2\sigma^2}} \end{bmatrix}_{M \times 1} \quad (10)$$

Ⅳ. Simulation Studies by Theoretical Approach

The following channel model($p = 1$) is selected as an example for calculating and comparing the $\xi_{\min}$ for both RBF and linear equalizer

$$H(z) = 0.5 + z^{-1} \quad (11)$$

The equalizer order, q, is assumed to be 1. As mentioned earlier, the transmitted sequences are assumed to be independent and equi-probable binary symbols, designated as a_k, that are either +1 or -1. Then the maximum number of RBF centers, M is equal to 8($M = 2^{p+q+1}$). To determine $\mathbf{g}_i = \left[g_{i,0}, g_{i,1}\right]^T$, the noise-free channel output, $\hat{r}_k$ and $\hat{r}_{k-1}$ are provided in Table 1. From (3), the $\mathbf{r}_k$ is represented as

$$\begin{aligned} \mathbf{r}_k &= \left[r_k, r_{k-1}\right]^T \\ &= \left[0.5a_k + a_{k-1} + n_k, 0.5a_{k-1} + a_{k-2} + n_{k-1}\right]^T \end{aligned} \quad (12)$$

Table 1. Input and Desired Channel States
$p = 1, q = 1, M = 8, \ H(z) = 0.5 + z^{-1}$

RBF Center	a_k	a_{k-1}	a_{k-2}	$\hat{r}_k$	$\hat{r}_{k-1}$
$\mathbf{g}_1$	-1	-1	-1	-1.5	-1.5
$\mathbf{g}_2$	-1	-1	1	-1.5	0.5
$\mathbf{g}_3$	-1	1	-1	0.5	-0.5
$\mathbf{g}_4$	-1	1	1	0.5	1.5
$\mathbf{g}_5$	1	-1	-1	-0.5	-1.5
$\mathbf{g}_6$	1	-1	1	-0.5	0.5
$\mathbf{g}_7$	1	1	-1	1.5	-0.5
$\mathbf{g}_8$	1	1	1	1.5	1.5

The estimated auto correlation matrix is as follows(refers to [7])

$$\Phi = \begin{bmatrix} 0.044 & 0 & 0 & 0 & 0.003 & 0 & 0 & 0 \\ 0 & 0.044 & 0 & 0 & 0 & 0.003 & 0 & 0 \\ 0 & 0 & 0.044 & 0 & 0 & 0 & 0.003 & 0 \\ 0 & 0 & 0 & 0.044 & 0 & 0 & 0 & 0.003 \\ 0.003 & 0 & 0 & 0 & 0.044 & 0 & 0 & 0 \\ 0 & 0.003 & 0 & 0 & 0 & 0.044 & 0 & 0 \\ 0 & 0 & 0.003 & 0 & 0 & 0 & 0.044 & 0 \\ 0 & 0 & 0 & 0.003 & 0 & 0 & 0 & 0.044 \end{bmatrix}_{8 \times 8} \quad (13)$$

The estimated cross-correlation matrix is

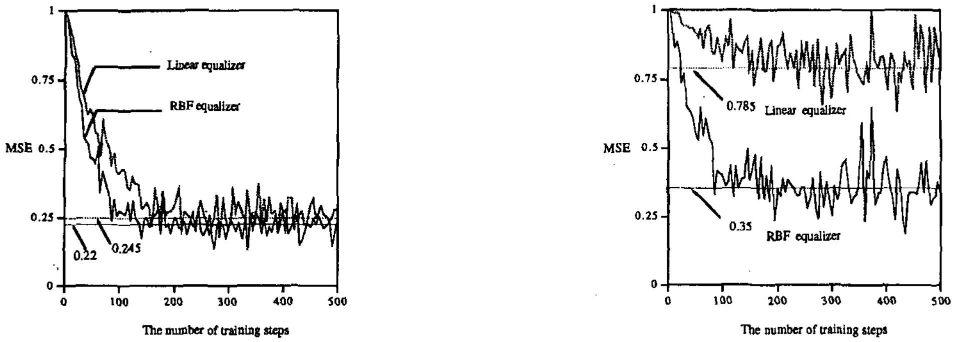
$$\begin{aligned} \alpha_{d=1} &= \left[-0.067,-0.067,0.067,0.067,-0.067,-0.067,0.067,0.067\right]^T \\ \alpha_{d=0} &= \left[-0.055,-0.055,-0.055,-0.055,0.055,0.055,0.055,0.055\right]^T \end{aligned} \quad (14)$$

Based on the equations (5), (13), and (14), the minimum MSE values are calculated as

$$\left.\begin{aligned} \xi_{\min}^{d=1} &= E\left[\left|a_{k-1}\right|^2\right] - \hat{\alpha}_{d=1}{}^T \hat{\Phi}^{-1} \hat{\alpha}_{d=1} = 0.22 \\ \xi_{\min}^{d=0} &= E\left[\left|a_k\right|^2\right] - \hat{\alpha}_{d=0}{}^T \hat{\Phi}^{-1} \hat{\alpha}_{d=0} = 0.35 \end{aligned}\right\} \quad (15)$$

For the purpose of comparison, theoretical minimum MSE for the linear equalizer is described as follows. By considering equations (12), the following results come out

$$\left.\begin{aligned} \Phi &= E\left[\mathbf{r}_k . \mathbf{r}_k^T\right] = \begin{bmatrix} 1.25 + \sigma_n^2 & 0.5 \\ 0.5 & 1.25 + \sigma_n^2 \end{bmatrix} \\ \alpha_{d=1} &= E\left[a_{k-1} . \begin{pmatrix} \mathbf{r}_k \\ \mathbf{r}_{k-1} \end{pmatrix}\right] = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \\ \alpha_{d=0} &= E\left[a_k . \begin{pmatrix} \mathbf{r}_k \\ \mathbf{r}_{k-1} \end{pmatrix}\right] = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} \end{aligned}\right\} \quad (16)$$



(a) desired symbol, $a_{k-d} = a_{k-1}$ (b) desired symbol, $a_{k-d} = a_k$ (no channel delay)

Fig. 2. Comparison of MSE convergence $H(z) = 0.5 + z^{-1}$, $\sigma_n^2 = 0.1$

where $\sigma_n^2 = E[n_k n_k]$. By substituting $\sigma_n^2 = 0.1$ into (16), the result is

$$\left.\begin{aligned} \xi_{\min}^{d=1} &= E\left[|a_{k-1}|^2\right] - \alpha_{d=1}^T \Phi^{-1} \alpha_{d=1} = 0.245 \\ \xi_{\min}^{d=0} &= E\left[|a_k|^2\right] - \alpha_{d=0}^T \Phi^{-1} \alpha_{d=0} = 0.785 \end{aligned}\right\} \quad (17)$$

From equations (15) and (17), we can see that the minimum MSE of the RBF equalizer is a little less than that of the linear equalizer when a_{k-1} is used as the desired symbol. In addition, the difference becomes more distinct when channel delay is not introduced. The low minimum MSE of the RBF equalizer without channel delay implies its capability for solving the non-linear separable problem. Fig. 2 shows the comparison of MSE convergence of both linear and RBF equalizers, which was obtained by stochastic gradient LMS learning[10]. As shown in 2(a), the minimum MSE of the RBF equalizer with the introduction of proper channel delay is a little lower than the linear equalizer. For the case of not introducing channel delay, the minimum MSE of the RBF equalizer is much lower than the linear equalizer, as shown in Fig. 2(b).

Thus far, what has been described above are results with training noise variance $\sigma_n^2 = 0.1$. Table 2 shows how differently noise variances affect the minimum MSE.

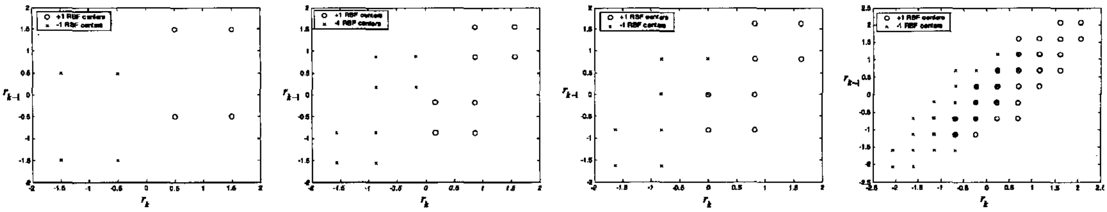
Other channel impulse responses were selected for simulation as an extension to this research study[7,12], which are the linearly separable, slightly distorted, and severely distorted channel models. Lee[7] shows that RBF equalizer outperforms linear equalizer for those three channel

Table 2. Comparison of minimum MSE $H(z) = 0.5 + 1.0z^{-1}$

Noise variance (σ_n^2)	Minimum MSE			
	$\xi_{\min}^{d=1}$		$\xi_{\min}^{d=0}$	
	RBF	Linear	RBF	Linear
0.1	0.22	0.245	0.35	0.785
0.2	0.27	0.291	0.59	0.804
0.3	0.327	0.332	0.732	0.82

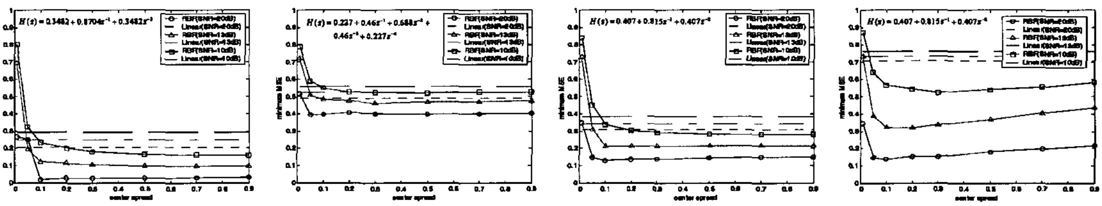
models. As mentioned in[3], distorted channel models contain deep nulls in their frequency spectrum which result in high MSE, and as a consequence a high bit error rate probability. Fig.3 shows the center distribution of four different kinds of linearly dispersive channel models. For the purpose of graphical illustration, it is assumed that the equalizer order, q, is one.

Fig. 3(a) and (b) show the distribution of noise free-RBF center, $\hat{\mathbf{r}}_k$, for linearly separable channels, (c) for slightly distorted, and (d) severely distorted channel model. Through this research, the minimum MSE of RBF equalizer for the channel models (b),(c), and (d) in Fig. 3 are investigated and compared with that of the linear transversal equalizer. For the purpose of the practical training of RBF equalizer, RBF centers used in simulation were estimated using supervised k-means clustering[4]. Also, both auto-correlation and cross correlation matrix are obtained from randomly generated 50,000 training samples, by statistically averaged values. Fig. 4 shows the minimum MSE of RBF equalizer and compares with that of linear equalizer. This simulation study is performed with different center spread parameters, σ^2.



(a) $\mathbf{h} = [0.5\ 1.0]$ (b) $\mathbf{h} = [0.3482\ 0.8704\ 0.3482]$ (c) $\mathbf{h} = [0.407\ 0.815\ 0.407]$ (d) $\mathbf{h} = [0.227\ 0.460\ 0.688\ 0.460\ 0.688]$

Fig. 3. Comparison of noise free-center distribution:



(a) desired symbol, $a_{k-d} = a_{k-1}$ (b) desired symbol, $a_{k-d} = a_{k-2}$ (c) desired symbol, $a_{k-d} = a_{k-1}$ (d) desired symbol, $a_{k-d} = a_k$ (no delay)

Fig. 4 Comparison of minimum MSE of RBF and linear equalizers for different SNR

Fig. 4 (a)-(c) show the approximate minimum MSE of RBF equalizer when channel delay is introduced in training. As shown in Fig. 4, the approximate minimum MSE of RBF equalizer is always less than that of linear equalizer. It was also found that the lowest value of approximate minimum MSE value of RBF equalizer for each channel model are obtained usually when the center spread values are approximately two to ten times more than the variances of AWGN. For example, in the Fig. 4 (c), the approximate minimum MSE of RBF equalizer for SNR=20dB(corresponding to 0.01 values of AWGN) was obtained when the center spread value is equal to 0.1, which is ten times more than the variance of introduced AWGN(equal to 0.01). For SNR= 13dB(corresponding to 0.05). the approximate minimum MSE of RBF equalizer was obtained when the center spread value is approximately equal to 0.1. Fig. 4(b) shows the approximate minimum MSE of RBF equalizer for the severely distorted channel. Although the approximate minimum MSE value of RBF equalizer for this model is still less than that of linear equalizer, their difference is relatively smaller than the case of either linearly separable or slightly distorted channels. In other words, the difference of approximate minimum MSE between RBF and linear equalizer for the linearly separable channel cases tend to be greater than the cases of distroted channels. On the other hand, Fig. 4(d) shows the minimum MSE for RBF equalizer when proper channel delay is not introduced. As shown in Fig. 4(c), the minimum MSE of RBF equalizer with 20dB of SNR is almost equal to the case where channel delay is introduced. Also, it was found that the difference of approximate minimum MSE between RBF and linear equalizer, when not considering channel delay, is much greater over all range of SNR than when the channel delay is introduced. This properties prove that the RBF based equalizer recover the transmitted symbols successfully without introducing proper channel delay, because the lower value of MSE usually leads to the good bit error rate performance.

V. Conclusion

Traditional adaptive algorithms for channel equalizers are based on the criterion of minimizing the mean square error between the desired filter output and the actual filter output. In this paper, the theoretical minimum MSE of a RBF equalizer was evaluated and compared with that of a linear equalizer. The procedure of computing theoretical minimum MSE of RBF equalizer was derived using the same concepts of finding minimum MSE for linear equalizer, based on the fact that the relationship between the hidden and output layers in the RBF equalizer is also linear. For the purpose of theoretically exact minimum MSE for both RBF and linear equalizer, a linear time dispersive channel, whose order is one, is selected. As extensive studies of this research, various channel models are selected for simulation which include fairly good channel (i.e., linearly separable), slightly distorted channel, and severely distorted channel models. Through the simulation studies with various channel models, it was found that RBF based equalizer with introduction of proper channel delay always produced lower minimum MSE than linear equalizer, and their difference varies with channel models. When channel is linearly separable, their difference goes high, when channel is worse, it becomes small. On the other hand, their difference, when the channel delay is not introduced, even goes higher than when channel delay is introduced. This property proves that RBF equalizer has a capability of making nonlinear decision boundary which eventually makes it possible to recover the transmitted symbols. In addition, it was found that the minimum MSE value of RBF equalizer was obtained with the center spread parameter which is relatively higher (approximately 2 to 10 times more) than variance of AWGN. This work provides the analytical framework for the practical training RBF equalizer system.

References

〔1〕 G.J. Gibson, S.Siu, and C.F.N. Cowan, "Application of Multilayer Perceptrons as Adaptive Channel Equalizers," *IEEE Int. Conf. Acoust. Speech, Signal Processing*, Glasgow, Scotland, pp. 1183-1186, 1989

〔2〕 P. R. Chang and B. C. Wang "Adaptive Decision Feedback Equalization for Digital Channels using Multilayer Neural Networks," *IEEE J. Selected Areas Commun.*, Vol.13, pp.316-324, Feb. 1995.

〔3〕 K. A. Al-Mashouq, I. S. Reed, "The Use of Neural Nets to Combine Equalization with Decoding for Severe Intesymbol Interference Channels," *IEEE Trans. Neural Networks*, Vol.5, pp.982-988, Nov. 1994.

〔4〕 S. Chen, B. Mulgrew, and P. M. Grant, "A Clustering Technique for Digital Communication Channel Equalization using Radial Basis Function Networks," *IEEE Trans. Neural Networks*, Vol.4, pp.570-579, Jul. 1993.

〔5〕 B. Mulgrew, "Applying Radial Basis Functions," *IEEE Sig. Proc. Mag.*, pp.50-65, Mar. 1996.

〔6〕 S. K. Patra and B. Mulgrew, "Computational Aspects of Adaptive Radial Basis Function Equalizer Design," *IEEE Int. Symposium on Circuits and Systems*, Hong Kong, pp.521-524, Jun. 1997.

〔7〕 J. Lee, A Radial Basis Function Equalizer with Reduced Number of Centers, Ph.D. dissertation, Florida Institute of Technology, 1996.

〔8〕 J. Lee, C.B. Beach, and N. Tepedelenlioglu, "A Practical Radial Basis Function Equalizer," *IEEE Trans. Neural Networks*, Vol.10, pp.450-455, Mar. 1999.

〔9〕 M. Ibnkahla, "Applications of Neural Networks to Digital Communications- a Survey," *Signal Processing*, pp. 1185-1215, Mar. 1999.

〔10〕 E. Lee and D. Messerschmitt, Digital Communication, second edition, Springer, 1993

〔11〕 S. Haykin, Adpative Filter Theory, third edition, Prentice Hall, 1996.

〔12〕 J. G. Proakis, Digital Communications, third edition, McGraw Hill, 1995

이 정 식 (Jung-Sik Lee) 정회원

1983년 한양대학교 전자통신과 공학사

1990년 Florida Institute of Tech., Elec. Eng. M.S.

1996년 Florida Institute of Tech., Elec. Eng. Ph.D.

1997년~현재 군산대 전자정보공학부 부교수

2004년~2006년 미국 플로리다 주립대(Univ. Of South Florida) 방문 교수

1986년-1988년 OPC 중앙연구소 연구원

<관심분야> 적응신호처리, 디지털 통신, 신경망응용