ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 267-272

Blind Equalization with a Linear Feedforward Neural Network

*Xi-Ren Cao, Jie Zhu**

Department of Electrical & Electronic Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong

Jennie Si†

Department of Electrical Engineering
Arizona State University
Tempe, AZ 85287-7606

Abstract. In this paper, we introduce a linear feedforward neural network for blind equalization in digital communications. The approach is based on a fundamental theorem, which makes the training procedure of the neural network very simple. The training is equivalent to a stochastic approximation algorithm and can be implemented recursively every time a sample data is received. As usual, the received signal is oversampled so that the channel can be described by a full-column rank matrix; the neural network searches for the inverse of the matrix. A simulation example is given to illustrate the performance of the neural network.

1. Problem Formulation and Related Work

Consider a simplified model of a multipath digital communication system illustrated in Figure 1. A sequence of signals $\{\alpha_k\}$, $k = 0, 1, 2, \cdots$, is sent over a communication channel with T time apart. The transfer function of the channel is denoted as a continuous function $h(t)$. The signals may be corrupted by noise $w(t)$. Denoting the received signal (baseband) as a continuous function $y(t)$, we have

$$y(t) = \sum_{i=-\infty}^{+\infty} \alpha_i h(t - iT) + w(t). \qquad (1)$$

Blind equalization is to estimate the input symbol sequence $\{\alpha_i,\ i = 1, 2, \cdots\}$ from the received signal $y(t)$ only. Let the sampling frequency be N/T where

*The work by Cao and Zhu was partly supported by Hong Kong RGC under grant HKUST 690/95E.

†The work by Si was partly supported by VSF under grant ECS9553202 and by Motorola.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 267-272

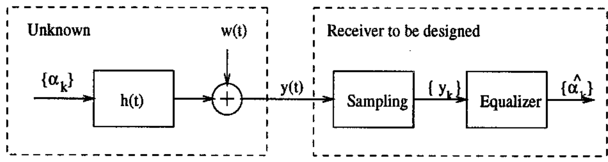

Figure 1: A Schematic Diagram of a Digital Communication Equalization

N is any positive integer, $N > 1$. After every duration $\delta = T/N$, one sample of the received signal is obtained by

$$y_k = y(k\delta) = \sum_{i=-\infty}^{+\infty} \alpha_i h((\frac{k}{N} - i)T) + w_k,$$

where we denote $w(k\delta)$ as w_k for convenience. Moreover, we assume that

$$h(t) = \begin{cases} h^c(t) & 0 \leq t < MT \\ 0 & else \end{cases}$$

where $h^c(t)$ is the channel with a time-span of MT. Therefore,

$$y_{k+nN} = \sum_{i=0}^{M-1} \alpha_{n-i} h^c((\frac{k}{N} + i)T) + w_{k+nN}. \qquad (2)$$

Let

$$\begin{aligned} \vec{o}(n) &= [y_{nN}, \cdots, y_{N-1+nN}, y_{(n+1)N}, \cdots, y_{N-1+(n+\lambda-1)N}]^\dagger, \\ \vec{s}(n) &= [\alpha_{n-M+1}, \alpha_{n-M+2}, \cdots, \alpha_n, \alpha_{n+1}, \cdots, \alpha_{n+\lambda-1}]^\dagger, \\ \vec{w}(n) &= [w_{nN}, \cdots, w_{N-1+nN}, w_{(n+1)N}, \cdots, w_{N-1+(n+\lambda-1)N}]^\dagger, \end{aligned}$$

where the superscript $\dagger$ stands for *transpose* and λ is the smallest integer which is not less than $(M-1)/(N-1)$. (2) can then be represented in vector matrix form as

$$\vec{o}(n) = A\vec{s}(n) + \vec{w}(n) \qquad (3)$$

where A is a $(\lambda N) \times (M + \lambda - 1)$ matrix defined by

$$A = \begin{bmatrix} \vec{h}_1 & \vec{h}_2 & \cdots & \vec{h}_M & 0 & 0 & \cdots & 0 \\ 0 & \vec{h}_1 & \vec{h}_2 & \cdots & \vec{h}_M & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \vec{h}_1 & \vec{h}_2 & \cdots & \vec{h}_M \end{bmatrix},$$

where $\vec{h}_j$ for $j = 1, 2, \cdots, M$ is a column vector and its i-th element is

$$h_{i,j} = h^c((\frac{i-1}{N} - j + M)T), \qquad \text{for } i = 1, 2, \cdots, N.$$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 267-272

We call A the *channel matrix.*

With the formulation represented in (3), the blind equalization without using higher order statistics has received an increasing interest during the last years because it requires a relatively short sequence of symbols. A number of algorithms have been published [1, 2, 3, 4, 5]; some of them employ the cyclostationarity of the input sequence, and the others are based on the exploitation of the *Toeplitz* structure of the channel matrix and the *Hankel* property of the input vector sequence $\vec{s}(n)$, $n = 1, 2, \cdots$. Most of them are implemented by matrix decomposition thereby belong to the category of batch algorithms; i.e., the input sequence is estimated once after a block of the samples has been received. As we know, batch algorithm suffers from a large amount of computation and may not be implemented on-line.

In this paper, we propose a blind equalizer based on the Hankel property of the input vector samples. The equalizer is a linear feedforward neural network with no hidden layers. For the sake of simplicity, we let $\vec{w}(n) = 0$ and make the following assumptions throughout this paper:

- The time-span of the channel over T, M, is known;
- A is an $N' \times M'$ full-column rank matrix (with $N' = \lambda N$ and $M' = M + \lambda - 1$);
- The input symbols are driven from a set of *i.i.d.* random variables with a finite sample set. The mean of the input sequence is zero.

The paper is organized as follows. In section 2, we prove a fundamental theorem which is the base of our approach; In section 3 we first propose a linear feedforward neural network as a channel equalizer, and then we develop its temporal learning rule; the simulation result is presented in section 4; finally the conclusion is given in section 5.

2. A Fundamental Theorem

An $m \times n$ matrix $H_{m\times n}$ is said to be a Hankel matrix if its (i, j)-th element takes the value of α_{i+j-1} where $\alpha_1, \alpha_2, \cdots$, is a sequence of arbitrary complex numbers. More precisely,

$$H_{m\times n} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots & \alpha_{n+1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_m & \alpha_{m+1} & \alpha_{m+2} & \cdots & \alpha_{m+n-1} \end{bmatrix}.$$

Theorem 1 *Let* $\{\alpha_k, \quad k = 1, 2, \cdots\}$ *be a sequence of uncorrelated random variables and* $S = [\alpha_{i+j-1}]$ *be a* $p \times q$, $p \leq q$, *Hankel matrix. Suppose* $\Gamma = DS$ *and* D *is a* $p \times p$ *non-singular matrix. If* Γ *is also a Hankel matrix, denoted as* $\Gamma = [\beta_{i+j-1}]$, *then* $D = \rho I$ *where* ρ *is any nonzero scalar and* I *is the identity matrix.*

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 267-272

Let S be such a matrix that its n-th column is $\vec{s}(n)$ for $n = 1, 2, \cdots$, then S is a Hankel matrix with M' rows. Moreover, let $\vec{e}(n) = W\vec{o}(n)$ be the n-th column of a matrix Γ for $n = 1, 2, \cdots$. By (3) with $\vec{w}(n) = 0$, we have $\Gamma = WAS$. Since S is a Hankel matrix, and by the previous assumption its elements are uncorrelated, the fundamental theorem claims that if Γ is a Hankel matrix then $WA = \rho I$ with $\rho \neq 0$. When $WA = \rho I$, the first row of Γ can be considered as the input symbol sequence. As a result, the problem becomes how to find an $M' \times N'$ matrix W, based on the received sequence $\vec{o}(n)$, $n = 1, 2, \cdots$, such that Γ is a Hankel matrix. Such a matrix W apparently exists if the channel matrix A is full column rank. One example of the suitable W is the pseudo-inverse of A, $W = (A^H A)^{-1} A^H$. The superscript ^{H} is *Hermitian transpose.*

3. The Linear Feedforward Network

A linear feedforward neural network without hidden units is depicted in Figure 2. It consists of N' input nodes, M' output nodes. The weight w_{ji} associated with

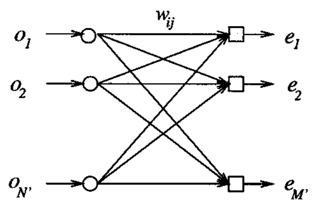

Figure 2: A Linear Feedforward Network without Hidden Layers

the connection between the i-th input node and j-th output node is a complex number. We use $W = \{w_{ji}\}$ to denote the weight matrix. Thus, in matrix notation we have $\vec{e}(n) = W\vec{o}(n)$ for $n = 1, 2, \cdots$.

In order to determine whether the output $[\vec{e}(1), \vec{e}(2), \cdots]$ forms a Hankel matrix, we propose the following criterion:

$$f(W) = \sum_{j=1}^{M'-1} |e_j(n) - e_{j+1}(n-1)|^2. \tag{4}$$

When the neural network is fully trained, $f(W)$ reaches its minimum zero. The learning rule of the neural network is simply

$$W(n) = W(n-1) - \eta_n \nabla f(W)|_{W=W(n-1)} \tag{5}$$

where $W(n)$ is the value of W after n-th update, η_n the learning step size used in the n-th update, $\nabla f(W)$ the steepest gradient matrix of $f(W)$. The (i,j)-th element of $\nabla f(W)$ at $W = W(n-1)$ denoted by $\nabla f_{ij}(W(n-1))$, can be obtained by (4) as follows

$$\nabla f_{ij}(W(n-1)) = \begin{cases} \Delta_1(n)o_i^*(n), & j=1, \\ \Delta_j(n)o_i^*(n) + \Delta_{j-1}(n)o_i^*(n-1), & 2 \leq j \leq M'-1, \\ \Delta_{M-1}(n)o_i^*(n-1), & j = M', \end{cases}$$

for $i = 1, 2, \cdots, N'$, where the superscript * is the *complex conjugate* operator and

$$\Delta_j(n) = e_j(n) - e_{j+1}(n-1).$$

Note that by (5) the network may trap to $W = 0$. There are a lot of techniques to prevent $W = 0$. In our simulations the last row of W is normalized after each update so that its Euclidean norm is unit. It is worth while to mention that the criterion function in (4) is by no means the only one that are minimized when the output has a Hankel formed. By the design of different criterion functions, other algorithms can be developed, focusing on their implementation complexity, convergent speed, and so on.

4. A Simulation Example

In this section we investigate the convergence property of the neural network. The symbols were drawn from the QPSK signal constellation with a uniform distribution. The Intersymbol Interference(ISI) is employed as a performance measure of the estimated sequence. We simulated a three-ray multipath channel presented in [5]. This channel was truncated up to 4 symbols, $M = 4$. We want to design a linear feedforward neural network with 4 input nodes and 4 output nodes. The initial phase of the neural network are set as W(0)=I. The ISI with respect to the number of received sample vectors is illustrated in Figure 3, where SNR=20dB, $\eta_n = 0.015$ for all n.

5. Conclusions

In this paper, we established a fundamental theorem for blind equalization. Based on this theorem we developed a linear feedforward neural network to implement an equalizer. Compared with the neural network approach presented in [6], the training rule in our approach is simpler and hence is easier to be implemented. However, the convergence speed of our approach is slower than [6]. The fundamental theorem provides a very simple optimization criterion. Other training rules based on this criterion may be developed. The research for faster algorithms is under way.

References

[1] L.Tong, G.Xu, and T.Kailath,"Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach," *IEEE Trans. on Information Theory,* vol.40, no.2, pp340-349, Mar.1994.

[2] D.Slock, "Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction," *Proceeding of IEEE ICASSP*, Adelaide, IV:585-588, May 1994.

[3] Z. Ding and Y. Li,"On Channel Identification Based on Second Order Cyclic Spectra", *IEEE Trans. on Signal Processing*, vol.42, no.5, pp1260-1264, May 1994.

[4] E.Moulines, P.Duhamel, J-F.Cardoso and S.Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. on Signal Processing*, vol.43, no.2, pp516-525, Feb. 1995.

[5] H. Liu, G. Xu, "Closed-Form Blind Symbol Estimation in Digital Communications," *IEEE Trans. on Signal Processing*, vol.43, no.11, pp2714-2723, Nov. 1995.

[6] G.Dong and R.Liu,"A Globally Convergent Temporal Neural Network for Blind Channel Identification," private communication.

[7] S.Haykin,"Neural Networks: A Comprehensive Foundation," New York: Macmillan College Publishing Company, 1994.

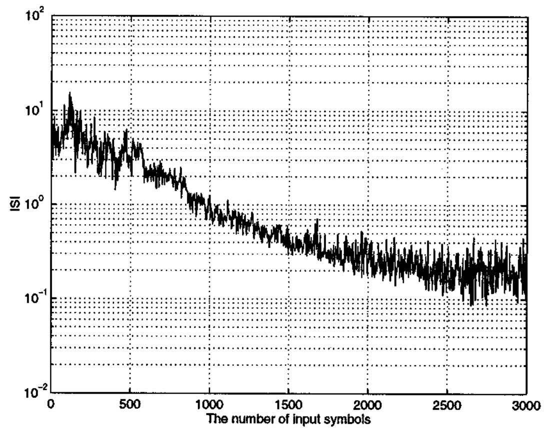

Figure 3: ISI vs. the number of input symbols under SNR=20dB