GDRNN을 이용한 비선형 채널 등화
Nonlinear channel equalization using GDRNN

김 용운, 박 동조
(Yong-Woon Kim, Dong-Jo Park)

305-701 대전광역시 유성구 구성동 373-1
한국과학기술원, 전기 및 전자공학과
(Department of Electrical Engineering, KAIST)
Tel) +82-42-869-8038 Fax) +82-42-869-8038
E-mail) ywkim@tercel.kaist.ac.kr, djpark@eekaist.kaist.ac.kr

ABSTRACT

이 논문에서는 비선형 Channel의 등화기를 설계하기 위해 새로운 구조를 갖는 신경회로망을 제안하였다. 비선형 Channel의 동적 특성을 제대로 학습하기 위해 새로운 신경회로망은 은닉층 노드의 출력이 은닉층의 입력으로 되먹임되는 구조를 갖는다. 또한 이 논문에서는 제안한 신경회로망의 구조에 알맞는 학습 알고리즘을 제안하였다. 제안한 신경회로망과 학습 알고리즘의 성능은 Computer simulation을 통해 보였고, 그 결과는 기존의 Channel 등화기를 사용했을 경우보다 나은 결과를 보여 주었다.

I. 서론

적응 Channel equalization은 디지털 통신에 있어서 중요한 문제 중의 하나이다. 기존의 Channel 등화기들은 선형 필터들을 결합하여 주어진 Channel의 역모델링이나 분류화 문제를 다루었다. 이들 선형 필터들의 계수들은 주어진 Channel의 특성에 알맞게 학습이 가능하다. 하지만 Channel 왜곡이나 Channel에 주어지는 잡음의 강도가 심해지면 선형 필터의 성능은 현저히 떨어진다.

이와 같은 선형 필터들의 결합을 보완하기 위해 비선형 필터들이 많이 소개되었다. 그 중에서 신경회로망을 이용한 Channel 등화기의 설계에 많은 관심이 집중되어 왔다. 하지만 신경회로망을 이용한 Channel 등화기를 설계함에 있어서도 단순한 Feedforward 구조를 가지게 함으로써 많은 가중치 요소들을 필요로 하게 하였고, 이는 계산상의 많은 문제점을 안겨 주었다. Feedforward 구조를 가지면서 Channel의 Delay 현상에 대한 동적 특성을 제대로 학습할 수가 없다. 이를 해결하기 위해 전체 폐회로를 구성하여 신경회로망의 출력을 다시 신경회로망의 입력으로 사용하는 방법들을 많이 사용하였다. 이 또한 주어진 Channel의 Delay가 얼마인지를 미리 알고 있어야 한다는 단점을 지니고 있고, 이를 해결하기 위해 많은 입력 노드들를 필요로 하게 되었다. 이는 계산량의 엄청난 증가를 의미한다.

이와 같은 가중치의 크기와 계산량을 고려하여 이 논문에서는 새로운 구조를 가지는 신경회로망을 제안하였다. 제안한 신경회로망은 Generalized diagonal recurrent neural networks (GDRNN)로서, 은닉층 노드는 각각 자기 자신의 출력을 다시 입력으로 받아 들이

계 된다. 즉 은닉층의 입력은 입력층에서 들어오는 외부 입력과 자체의 Feedback에 의한 내부 입력으로 구성된다. 따라서 제안한 신경회로망의 은닉층의 기존의 비선형 IIR 필터를 일반화한 구조인 것이다. 여기에서는 입력층과 출력층을 갖는 2층 구조 신경회로망을 고려한다.

II. 제안한 신경회로망 및 학습 알고리즘

2.1 신경회로망의 구조

제안한 신경회로망의 수학적 모델은 다음과 같이 쓸 수 있다.

$$y(k) = f(O(k)) \tag{1}$$

$$O(k) = W_O^T X(k) \tag{2}$$

$$X(k) = f(S(k)) \tag{3}$$

$$S(k) = W_I^T U(k) + \sum_p W_{Dp}^T X(k-p) \tag{4}$$

이고, 여기에서 $f(x)$는 기존의 신경회로망에서와 같은 Sigmoid 함수를 의미한다. 식 (1)에서 $y(k)$는 신경회로망의 출력을 나타내고, $O(k)$는 신경회로망의 출력층의 Sigmoid 함수로 들어오는 은닉층의 가중치 합을 의미한다. 식 (3)에서 $X(k)$는 신경회로망 은닉층의 출력을 의미하고, $S(k)$는 은닉층으로 들어오는 가중치 합을 나타낸다. 앞에서 설명한 바와 같이 $S(k)$는 외부 입력의 가중치 합과 내부 입력 즉 Feedback된 자체 입력의 가중치 합으로 구해진다.

또한 식 (2)에서 $W_O \in R^{N_O \times N_H}$는 신경회로망 출력층의 가중치를 나타내고, 식 (4)에서 $W_I \in R^{N_H \times N_I}$는 입력층의 가중치를 의미하고, $W_D \in R^{N_H \times N_D}$는 은닉층이 가지고 있는 Feedback된 입력들이 가지는 가중치 행렬이다. 여기에서 N_I는 입력층 노드들의 개수, N_H는 은닉층 노드들의 개수, N_D는 은닉층의 Feedback된 출력의 개수, 그리고 N_O는 출력층 노드들의 개수이다. 이때 j번째 은닉층의 구조를 그림으로 나타내면 그림 1과 같다.

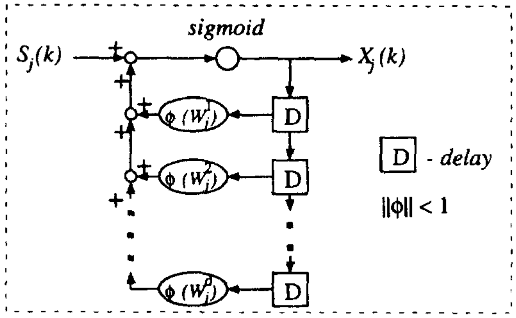

그림 1. 제안한 신경회로망의 j번째 은닉층의 구조

2.2 학습 알고리즘

우선 신경회로망의 출력층 가중치 W_O를 학습하는 알고리즘부터 알아본다. 학습 알고리즘에 필요한 최적화 방법은 다음과 같다.

$$\textbf{Minimize} \quad \frac{1}{2}\|W_O(k+1) - W_O(k)\|^2 \tag{5}$$

$$\textbf{Subject to} \quad f^{-1}(\tilde{d}(k+1)) - \tilde{O}(k+1) = 0$$

이고, 여기에서 $\tilde{d}(k+1) = d(k)$는 원하는 신경회로망 출력인 A posteriori 출력값이고, $\tilde{O}(k+1) = W_O^T(k+1)X(k)$는 신경회로망의 A posteriori 출력값을 말한다. 즉 이는 주어진 제한적 최소화 문제 (Constrained minimization problem)를 풀어야 함을 알 수 있다. 따라서 주어진 문제를 비제한적인 문제 (Unconstrained problem)로 바꾸기 위해서는 Lagrangian λ_O를 도입하여 Hamiltonian을 구하면 된다. Hamiltonian H_O를 구하면 다음과 같다.

$$H_O = \frac{1}{2}\|W_O(k+1) - W_O(k)\|^2 + \lambda_O(f^{-1}(\tilde{d}(k+1)) - \tilde{O}(k+1)) \tag{6}$$

이다. 이때, 신경회로망의 출력층 가중치 $W_O(k+1)$와 Lagrangian λ_O에 대해 위의 Hamiltonian을 최소화하는 필요조건은 $\frac{\partial H_O}{\partial W_O(k+1)} = 0$와 $\frac{\partial H_O}{\partial \lambda_O} = 0$가 된다.

위의 필요 조건들과 $\frac{\partial \tilde{O}(k+1)}{\partial W_O(k+1)} = X(k)$

라는 관계를 이용하면, 신경회로망 가중치 W_O의 학습 방법은

$$W_O(k+1) = W_O(k) + \lambda_O \frac{\partial \tilde{O}(k+1)}{\partial W_O(k+1)} = W_O(k) + \lambda_O X(k) \quad (7)$$

이 된다. 이때, 주어진 λ_O에 대한 필요조건에서 $\frac{\partial H_O}{\partial \lambda_O} = f^{-1}(\tilde{d}(k+1) - \tilde{O}(k+1) = 0$이므로

$$\lambda_O = \frac{f^{-1}(\tilde{d}(k+1) - W_O^T(k)X(k)}{X^T(k)X(k)} \quad (8)$$

가 된다. 따라서 신경회로망 가중치 W_O를 학습시키는 알고리즘은

$$W_O(k+1) = W_O(k) + \frac{f^{-1}(\tilde{d}(k+1) - W_O^T(k)X(k)}{X^T(k)X(k)} X(k) \quad (9)$$

와 같이 구할 수 있다.

Hamiltonian H_O를 구하는 방법과 같은 방법으로 신경회로망의 가중치 W_D를 구할 수 있다. j번째 은닉층의 가중치 W_j^p에 대해 Hamiltonian을 H_{jp}라 하면,

$$H_{jp} = \frac{1}{2}\|W_j^p(k+1) - W_j^p(k)\|^2 + \lambda_D(f^{-1}(\widetilde{d_{hj}}(k+1)) - \tilde{S}_j(k+1))$$

(10)

과 같이 쓸 수 있다. 여기에서 $\widetilde{d_{hj}}(k+1) = d_{hj}(k)$는 원하는 내부 출력값 $d_h(k)$의 j번째 A posteriori 출력값을 의미하고, $\tilde{S}_j(k+1)$은 $S_j(k)$의 A posteriori 값을 나타낸다. 여기에서 원하는 내부 출력값 $d_h(k)$는 $d(k) = f(W_O^T d_h(k))$라는 가정에서 얻을 수 있다. 따라서 Moore-Penrose의 Pseudo-inverse에 의해 $d_{h(k)} = W_O^+ f^{-1}(d(k))$와 같이 구할 수 있고, 여기에서 +는 주어진 행렬의Pseudo-inverse를 의미한다. 신경회로망 가중치 W_O를 구하는 방법과 같은 방법에 의해 Lagrangian을 구하면, 가중치 W_j^p를 구하는 학습 알고리즘은

$$W_j^p(k+1) = W_j^p(k) + \frac{\varepsilon(k)}{P_j^p(k)X_j(k-p)} \cdot P_j^p(k) \quad (11)$$

이고, 여기에서 $P_j^p(k) = \frac{\partial \tilde{S}_j(k+1)}{\partial W_j^p}$와 같이 정의되고, A posteriori 오차인 $\varepsilon(k)$는 다음과 같이 정의된다.

$$\varepsilon(k) = f^{-1}(\widetilde{d_{hj}}(k+1)) - \sum_d W_j^d(k)X_j(k-d) - (W_j^I)^T(k)U(k) \quad (12)$$

그리고, 간단한 계산과 Real-Time Recurrent Learning (RTRL) 방법을 이용하면,

$$P_j^p(k) = f(S_j(k))(X_j(k-p) + \sum_d W_j^p(k)P_j^p(k-d)) \quad (13)$$

를 얻을 수 있다.

위와 같은 방법으로 신경회로망 가중치 W_I를 구하면,

$$W_j^I(k+1) = W_j^I(k) + \frac{\varepsilon(k)}{Q_j^T(k)U(k)} \cdot Q_j(k) \quad (14)$$

를 얻을 수 있고, 여기에서 $Q_j(k) = \frac{\partial \tilde{S}_j(k+1)}{\partial W_j^I}$와 같이 정의되며, 다음과 같이 구해진다.

$$Q_j(k) = f(S_j(k))(U(k) + \sum_d W_j^d(k)Q_j(k-d)) \quad (15)$$

III. Computer Simulation Results

위에서 제안한 신경회로망과 그의 학습 알고리즘을 이용한 Channel 등화기의 구조는 그림 2와 같다.

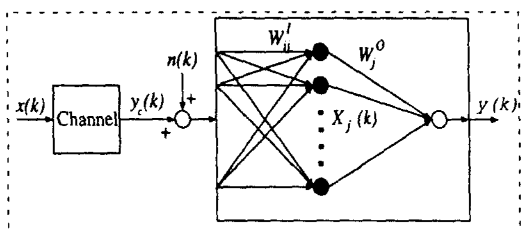

그림 2. 제안한 신경회로망을 이용한 Channel 등화기의 구조

그리고, 주어진 Channel의 수학적 모델은

$$\begin{aligned} y_c(k) &= 0.2y_c^2(k-1)+0.2y_c^2(k-2) \\ &+0.4\sin(0.5y_c(k-1)y_c(k-2)) \\ &\cdot\cos(0.5y_c(k-1)y_c(k-2))+x(k) \\ &+0.5x(k-1)+0.25x(k-2)+n(k) \end{aligned} \quad (16)$$

과 같다고 가정하고, 여기에서 $y_c(k)$는 Channel의 출력을 말하고, $x(k)$는 Channel의 입력을 의미하고, $n(k)$는 Channel의 출력에 더해지는 White gaussian 잡음을 의미한다. 주어진 Channel과 제안한 신경회로망 및 그의 학습 알고리즘을 사용하여 Computer simulation을 행한 결과가 그림 3이다.

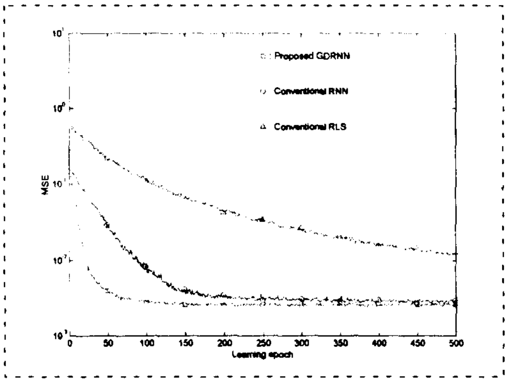

그림 3. 제안한 신경회로망의 MSE 수렴성

그림 3에서는 제안한 신경회로망을 이용한 방법과 기존의 신경회로망을 이용한 방법, 선형 필터를 이용한 방법을 비교하였다. Computer simulation에서, 제안한 신경회로망은 1개의 입력 노드, 5개의 은닉 노드와 각각 1개의 Delay 가중치, 그리고 1개의 출력 노드를 사용하였다. 또한 선형 필터는 5개의 Tapped delay로 구성되어 있다. 그림에서 보는 바와 같이 제안한 신경회로망은 1개의 입력 노드를 가지면서 주어진 동적 Channel의 특성을 잘 학습하고 있는 것을 볼 수 있다.

IV. 결론

이 논문에서는 비선형 Channel 등화기를 설계하기 위해 새로운 구조를 갖는 신경회로망을 제안하였다. 또한 제안한 신경회로망을 위한 학습 알고리즘을 Constrained 최적화 방법에 의해 도출하였다. Computer simulation에서도 볼 수 있듯이 주어진 신경회로망은 최소한의 가중치를 가지고서 동적 시스템의 특성을 잘 모사할 수 있다는 것을 볼 수 있다.

참고문헌

[1] T. Adali, X. Liu, and M. K. Sonmez, "Conditional Distribution Learning with Neural networks and Its Application to Channel Equalization," *IEEE Trans. on Signal Processing*, Vol. 45, No. 4, Apr 1997, pp.1051-1064.

[2] C. C. Ku and K. Y. Lee, "Diagonal recurrent neural networks for dynamic systems control," *IEEE Trans. on Neural Networks*, Vol. 6, No. 1, Jan 1995, pp.144-156.

[3] S. Ong, C. You, S. Choi and D. Hong, "A decision feedback recurrent neural equalizer as an infinite impulse response filter," *IEEE Trans. on Signal Processing*, Vol. 45, No. 11, Nov 1997, pp.2851-2858.

[4] R. Parsi, E. D. Di Claudio, G. Orlandi, and B. D. Rao, "Fast adaptive digital equalization by recurrent neural networks," *IEEE Trans. on Signal Processing*, Vol. 45, No. 11, Nov 1997, pp.2731-2739.