

Hybrid 리커런트 신경망을 이용한 시스템 식별

System Identification Using Hybrid Recurrent Neural Networks

최한고*, 고일환*, 김종인*

Han-Go Choi*, Il-Whan Go*, Jong-In Kim*

요약

동적 신경망은 temporal 신호처리가 요구되는 여러 분야에 사용되어 왔다. 본 논문에서는 다층 리커런트 신경망(RNN)의 동특성을 더 향상시키기 위해 지역 케환 신경망(LRNN)과 광역 케환 신경망(GRNN)으로 구성된 합성 신경망을 사용하여 시스템 식별을 다루고 있다. 합성 신경망의 구조는 LRNN으로 IIR-MLP를, GRNN으로 Elman RNN을 결합하고 있다. 합성신경망은 선형과 비선형 시스템 식별을 통해 평가되었으며 상대적인 성능평가를 위해 Elman RNN과 IIR-MLP 신경망과 비교하고 있다. 시뮬레이션 결과에 의하면 합성 신경망은 학습속도와 정확도에서 더 우수하게 동작하였으며, 이러한 사실은 비선형 시스템 식별에 있어서 합성 신경망이 기존의 다층 리커런트 신경망보다 더 효과적인 신경망이 될 수 있음을 보여주었다.

Abstract

Dynamic neural networks have been applied to diverse fields requiring temporal signal processing. This paper describes system identification using the hybrid neural network, composed of locally(LRNN) and globally recurrent neural networks(GRNN) to improve dynamics of multilayered recurrent networks(RNN). The structure of the hybrid network combines IIR-MLP as LRNN and Elman RNN as GRNN. The hybrid network is evaluated in linear and nonlinear system identification, and compared with Elman RNN and IIR-MLP networks for the relative comparison of its performance. Simulation results show that the hybrid network performs better with respect to the convergence and accuracy, indicating that it can be a more effective network than conventional multilayered recurrent networks in system identification.

Key words : IIR-MLP, Recurrent neural networks, Hybrid RNN, System identification

I. 서론

최근에 동적 신경망은 제어, 음성신호처리, 통신, 영상처리, 의공학 등에 많이 사용되고 있는데, 구체적으로 시스템 식별, 플랜트 제어, 패턴인식, 시계열 예측, 비선형 필터링과 같이 temporal 신호처리가 요구되는 분야에 많이 적용하고 있다[1-4]. 이 중 시스템 식별(identification)이란 미지의(unknown) 변수를 갖는 plant나 process를 모델링하는 것으로서, 신경망을 사용할 경우 plant의 입출력 데이터를 수집한 후 반복적인 학습을

통해 미지 시스템을 실험적으로 구한다. 신경망은 회로망 내부에 비선형적인 처리능력을 갖춘 유닛으로 구성되어 있고 주위 환경에 적응할 수 있는 학습능력을 갖고 있으므로 비선형 및 비정적 시스템 식별에 적절한 모델이라 할 수 있다. 그러나 미지 시스템의 비선형적 특성이 강할수록 만족스런 식별을 위해 동특성이 우수한 신경망이 요구되는데, 이러한 동적 신경망은 기존의 정적 신경망에 동특성을 추가하기 위해 신경망 내에 버퍼(buffer)를 두어 과거 입력에 대한 메모리로 사용하거나 케환(feedback)을 사용함으로써 신경망 내부에 temporal 동특성을 갖도록 하였다.

신경망 구조에서 가장 일반적으로 사용되고 있는 다층 신경망을 이용하여 지금까지 제안되어 구현되고 있는 동특성 신경망은 크게 두 부류로 대별된다. 첫 번째 구조로는 신경망 내에 버퍼를 사용하는 구조로 신경망 입력에

*금오공과대학 전자공학부

접수 일자 : 2005. 1. 11 수정 완료 : 2005. 1. 24

논문 번호 : 2005-1-2

※본 연구는 금오공과대학교 학술연구비에 의하여 연구된 논문입니다.

TDL(tapped delay line) 구조의 버퍼를 사용한 buffered 다층 퍼셉트론(MLP)과 각 뉴런의 입력이 FIR(finite impulse response) 필터 시냅스(synapse)를 갖는 FIR 신경망 혹은 시지연 신경망(Time delay neural networks)이 있다[5-7]. FIR 신경망 입력의 내부버퍼가 신경망의 외부버퍼로 구현될 수 있기 때문에 buffered MLP와 FIR 신경망은 이론적으로 동일하다.

두 번째로는 케환구조의 신경망인데 지역케환 신경망(locally recurrent neural networks, LRNN)과 광역케환 신경망(globally recurrent neural networks, GRNN)이 있다. 광역케환 신경망으로는 한 층의 모든 뉴런 출력들이 다른 층의 모든 뉴런 입력들과 연결된 구조, 즉, fully connected 리커런트 신경망(RNN)으로서 가장 대표적인 신경망으로 Elman과 Jordan 신경망이 있다[2,8-10]. 지역케환 구조로는 최근에 제안된 구조로서 뉴런과 뉴런 사이에 이산시간 내부 메모리를 갖는 시냅스를 사용하고 있다. 즉, 기존의 시냅스를 FIR 필터나 IIR(infinite impulse response) 필터로 대체한 구조로서 이러한 FIR-MLP 혹은 IIR-MLP는 기존의 선형 FIR 혹은 IIR 필터의 비선형 확장으로 생각할 수 있다[11]. 특히, IIR-MLP 구조는 Back-Tosi에 의해 제안되었는데 기존 정적 신경망 시냅스를 ARMA(Autoregressive moving average) 모델의 IIR 적응필터로 대체한 것이다[12]. 이러한 구조의 동적 다층 신경망은 기존의 정적(static) 다층 신경망이나 GRNN의 MLP나 FIR-MLP에 비해 시스템 식별과 같이 비선형 혹은 비정적 신호처리에서 더 우수한 성능을 보여주고 있다.

본 논문에서는 LRNN의 특성을 유지하면서 신경망의 동특성을 더욱 더 향상시키기 위해 LRNN과 GRNN이 합성된 구조의 hybrid 신경망(locally and globally RNN, LGRNN)을 제안하였으며, LGRNN을 사용한 시스템 식별이 기존의 다층 동적 신경망보다 더 우수한 식별을 수행할 수 있음을 확인하고자 한다. Hybrid 신경망인 LGRNN의 구조는 LRNN의 대표적인 신경망인 Tsoi의 IIR-MLP 구조와 GRNN으로 많이 사용하고 있는 Elman RNN 구조로 구성되어 있다. 학습방법으로는 LGRNN과 IIR-MLP에 대해서는 Campolucci가 제안한 on-line RBP(recursive backpropagation) 알고리즘을, GRNN에 대해서는 기존의 오차 역전파(error backpropagation) 학습 알고리즘을 사용하였다[11].

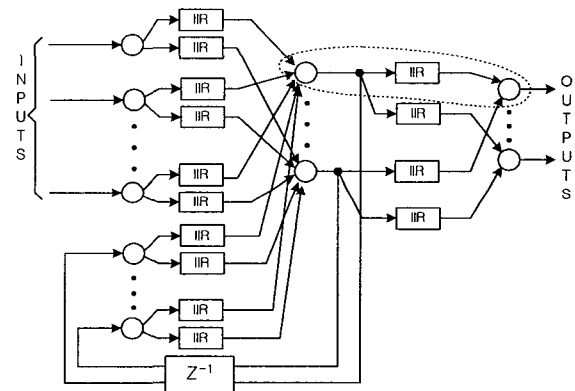
LGRNN 신경망 특성은 선형과 비선형 시스템 식별을 통해서 조사하였으며, 상대적인 성능평가를 위해 GRNN으로 기존의 Elman RNN과 LRNN으로 IIR-MLP와의 식별결과를 상호 비교함으로써 LGRNN 신경망의 우수성을 확인하고자 한다. 달리 언급하지 않으면 지금부터 본 논문에서는 GRNN은 Elman 신경망을, 그리고 LRNN은 IIR-MLP를 각각 나타낸다. 시뮬레이션 결과를 살펴보면 LGRNN 신경망에 의한 시스템 식별은 LRNN과 GRNN보다 수렴특성과 시스템 안정성, 즉, 시스템 오차면에서 더 우수한 성능을 보여주었다.

II. Hybrid 신경망

2.1 신경망 구조

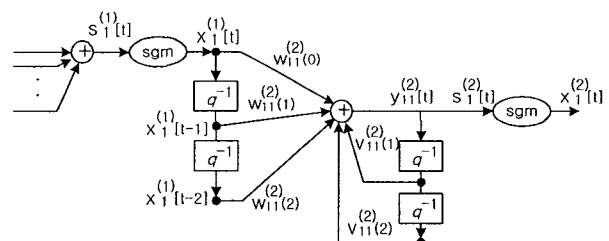
Hybrid LGRNN은 Elman의 RNN과 Tsoi의 IIR-MLP이 합성된 다층 신경망이다. 그림 1은 LGRNN 구조로서 그림 1(a)에서와 같이 입력층이 외부 입력과 히든 층의 모든 뉴런으로부터의 케환입력으로 구성되어 있으며, 입력층과 히든층의 모든 뉴런 출력이 IIR 시냅스를 통과하여 히든층과 출력층의 모든 뉴런으로 연결되어 있는 fully connected 구조이다.

그림 1(b)는 그림 1(a)의 점선 부분, 즉, 층간의 연결계수로 사용하고 있는 시냅스 구조를 확대한 것으로서 LGRNN의 입력층과 히든층, 그리고 히든층과 출력층간에 연결된 모든 시냅스의 내부 연결계수들을 자세히 도시한 것이다. 시냅스는 AR(autoregressive)과 MA(moving average) 구조를 갖는 IIR 적응필터를 사용하고 있으며 그림 1(b)에서는 MA 차수가 3, AR 차수가 2인 경우이다. 그림 1(a)에서 히든층으로부터 입력층으로의 케환이 없는 구조가 IIR-MLP 신경망이며, IIR 시냅스 대신 한 개의 연결계수만으로 이루어진 구조가 Elman RNN 신경망이다. 이와 같이 층간의 케환과 시냅스 내부에서의 케환 구조는 케환이 없는 다층 신경망 구조보다 동특성을 더욱 향상시킬 수 있다.



(a) 전체 신경망 구조

(a) Structure of neural networks



(b) 확대된 IIR 시냅스

(b) Enlarged IIR synapse

그림 1. LGRNN 신경망 구조

Fig. 1. Structure of LGRNN networks

입력층은 입력신호의 처리없이 단순히 다음 층으로의 신호전달 역할만 수행하며, 히든층 뉴런의 내부 전달함수는 양극 시그모이드(bipolar sigmoid)함수를 사용하였고, 출력층에서는 이득이 1인 선형 전달함수를 사용하였다. 히든층의 j 번째 뉴런 출력 $x_j(t)$ 과 출력층의 k 번째 뉴런 출력 $x_k(t)$ 는 각각 다음과 같다.

$$x_j(t) = \text{sgm}\{s_j(t)\} = \frac{1 - e^{-\alpha \cdot s_j(t)}}{1 + e^{-\alpha \cdot s_j(t)}} \quad (1)$$

$$x_k(t) = s_k(t) \quad (2)$$

여기서 α 는 양극 시그모이드 함수의 기울기이며, $s_j(t)$ 와 $s_k(t)$ 는 히든층의 j 번째와 출력층의 k 번째 뉴런의 입력을 나타낸다. 상대비교를 위한 Elman RNN에서도 같은 활성화함수를 사용하였다.

2.2 신경망 학습

LRGNN의 IIR 시냅스 학습 알고리즘은 LRNN의 학습과 동일하다. LRNN에서의 학습은 recursive backpropagation(RBP) 알고리즘을 제안하고 있는데, on-line 구현이 가능하도록 알고리즘을 수정한 causal RBP과 학습속도를 향상시키기 위해 이를 근사화한 instantaneous recursive backpropagation (IRBP)를 사용하고 있다[11]. 본 논문에서 제안된 LGRNN 신경망의 학습 알고리즘으로는 IRBP 알고리즘을 적용함으로써 on-line 학습이 수행되도록 하였다. IRBP 학습 알고리즘을 요약하면 다음과 같다.

$$\begin{aligned} \Delta w_{nm}^{(l)}[t+1] &= -\frac{\eta}{2} \frac{\partial E}{\partial s_n^{(l)}[t]} \frac{\partial s_n^{(l)}[t]}{\partial w_{nm}^{(l)}} \\ &= \eta \delta_n^{(l)}[t] \frac{\partial s_n^{(l)}[t]}{\partial w_{nm}^{(l)}} \end{aligned} \quad (3)$$

$$\begin{aligned} \Delta v_{nm}^{(l)}[t+1] &= -\frac{\eta}{2} \frac{\partial E}{\partial s_n^{(l)}[t]} \frac{\partial s_n^{(l)}[t]}{\partial v_{nm}^{(l)}} \\ &= \eta \delta_n^{(l)}[t] \frac{\partial s_n^{(l)}[t]}{\partial v_{nm}^{(l)}} \end{aligned} \quad (4)$$

여기서 $\Delta w_{nm}^{(l)}[t+1]$ 와 $\Delta v_{nm}^{(l)}[t+1]$ 는 시냅스 내부의 MA와 AR의 연결계수 증가량을, η 는 학습율을, E 는 학습샘플에 대한 자승에러를, $s_n^{(l)}[t]$ 는 l 층의 n 번째 뉴런에서 전달함수 이전의 내부 시냅스의 합을 나타낸다. 그리고,

$$\frac{\partial s_n^{(l)}[t]}{\partial w_{nm}^{(l)}} = x_m^{(l-1)}[t-p] + \sum_{r=1}^{l-1} w_{nm}^{(r)} \frac{\partial s_n^{(l)}[t-r]}{\partial w_{nm}^{(l)}} \quad (5)$$

$$\frac{\partial s_n^{(l)}[t]}{\partial v_{nm}^{(l)}} = y_{nm}^{(l-1)}[t-p] + \sum_{r=1}^{l-1} v_{nm}^{(r)} \frac{\partial s_n^{(l)}[t-r]}{\partial v_{nm}^{(l)}} \quad (6)$$

이며, 여기서 $x_m^{(l-1)}[t-p]$ 는 시간 $(t-p)$ 에서 l 층의 m 번째 뉴런 출력이며 특별히 $m=0$ 는 바이어스 입력을

나타낸다. $y_{nm}^{(l-1)}[t-p]$ 는 시간 $(t-p)$ 에서 $l-1$ 층의 n 번째 뉴런의 m 번째 입력과 관련된 시냅스 필터의 출력을 의미한다. $w_{nm}^{(l)}$ 과 $v_{nm}^{(l)}$ 은 내부 시냅스의 MA 부분과 AR 부분의 연결계수를 각각 나타낸다. 그리고

$$\delta_n^{(l)}[t] = e_n^{(l)}[t] \cdot \text{sgm}'(s_n^{(l)}[t]) \quad (7)$$

이며, $e_n^{(l)}[t]$ 는 다음과 같다.

$$e_n^{(l)}[t] = \begin{cases} e_n[t], & \text{for } l=M \\ \sum_{q=1}^{N_{l+1}} \delta_q^{(l+1)}[t] w_{nm}^{(l+1)}, & \text{for } 1 \leq l < M \end{cases} \quad (8)$$

$e_n(t)$ 는 원하는 출력과 실제 출력과의 차이를 나타내며, $\text{sgm}'(s_n^{(l)}[t])$ 은 뉴런 내부의 시냅스 합에 대한 시그모이드 함수의 미분치를 나타내며, M 은 신경망의 층의 개수이며 l 은 인덱스로 $l=0$ 과 $l=M$ 은 각각 입력층과 출력층을 나타낸다.

참고로 RNN의 학습 방법으로는 오차 역전파 알고리즘을 사용하였으며, 학습속도를 향상시키기 위해서 모멘텀항을 추가하였다.

$$\begin{aligned} \Delta w[t+1] &= -\eta \frac{\partial E}{\partial w[t]} + \alpha \Delta w[t] \\ &= \eta \delta[t] y(t) + \alpha \Delta w[t] \end{aligned} \quad (9)$$

여기서 η 는 학습율을, α 는 모멘텀항을 나타낸다. $\delta(t)$ 는 뉴런 내부상태를 나타내는데, 출력층의 $\delta_k(t)$ 와 히든층의 $\delta_j(t)$ 는 각각 다음과 같다.

$$\delta_k[t] = \frac{\partial E}{\partial h_k[t]} = -e_k[t] f'(h_k[t]) \quad (10)$$

$$\delta_j[t] = \frac{\partial E}{\partial h_j[t]} = f'(h_j[t]) \sum_k \delta_k[t] w_{kj}[t] \quad (11)$$

여기서 $e_k[t]$ 는 실제 출력과 원하는 출력과의 차이이며 $f(\cdot)$ 는 뉴런 활성화함수의 미분치이고, $h_j[t]$ 와 $h_k[t]$ 는 히든층 뉴런과 출력층 뉴런의 내부 상태를 나타낸다.

III. 실험결과

시스템 식별을 위한 방법으로는 state-space 모델링이나 input-output 모델링을 사용하여 구현하고 있다. 그림 2는 input-output 모델링 방식으로서 신경망을 사용하여 시스템 식별을 수행하기 위한 구성도이다. 신경망 출력 $\hat{d}(k)$ 은 실제 시스템 출력 $d(k)$ 의 추정치이며, 이들 간의 오차 $e(k)$ 는 다음과 같다.

$$\hat{d}(k) = f\{s(k-1), \dots, s(k-N), h(k-1), \dots, h(k-M)\} \quad (12)$$

$$e(k) = d(k) - \hat{d}(k) \quad (13)$$

여기서 N 는 미지 시스템 차수이며, $h(n)$ 은 신경망에서 히든 층이나 출력 층으로부터 입력층으로의 궤환입력이다. 그림에서 미지의 시불변(time invariant) 시스템과 신

경망에 인위적으로 발생시킨 랜덤 신호 $s(k)$ 를 가하여, 아래와 같은 입력과 출력 쌍을 생성한 후 오차 $e(k)$ 가 최소가 되도록 신경망을 학습하여 최적의 신경망 계수가 결정된다.

$$[\{s(k), s(k-1), \dots, s(k-N)\}, d(k)] \quad (14)$$

$$k=1, 2, \dots, L$$

여기서 L 은 학습 데이터 구간을 나타낸다. 제안된 신경망의 성능을 평가하기 위해 3.1과 3.2에서 기술된 선형 및 비선형 시스템 모델에 대해 실험하였으며, 상대적인 성능평가를 위해 광역계환 구조인 GRNN과 지역계환 구조인 LRNN에 의한 식별결과와 상호 비교하였다.

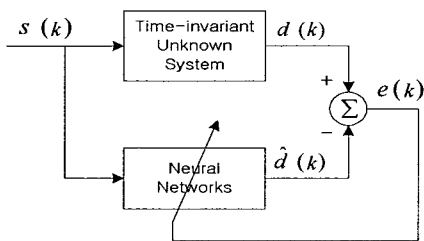


그림 2. 신경망을 사용한 시스템 식별

Fig. 2. System identification using neural networks

3.1 선형 시스템 식별

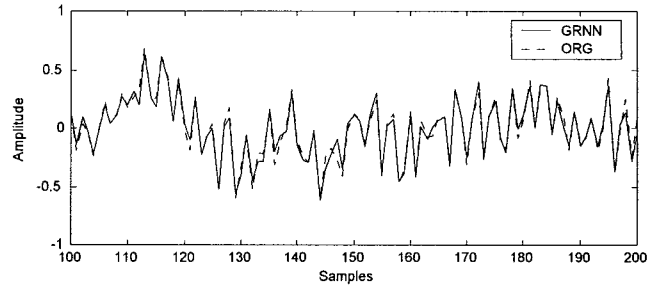
미지의 선형 시스템 식별로 필터와 같이 다음 식으로 정의되는 시불변 시스템을 사용하였다.

$$d(k) = 0.921s(k) + 0.107s(k-1) + 0.336s(k-2) + s(k-3) + 0.336s(k-4) + 0.107s(k-5) + 0.921s(k-6) \quad (15)$$

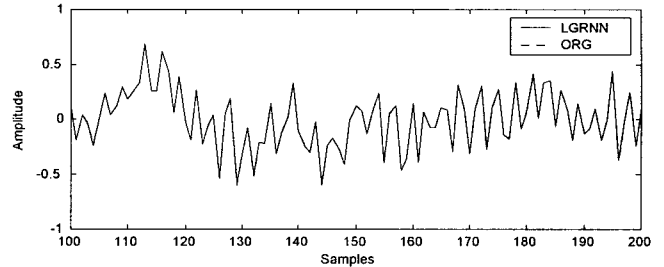
사용된 신경망의 구조는 GRNN의 경우 신경망 입력층 유닛은 4개, 히든층과 출력층의 유닛개수는 각각 8개와 1개로 (4x8x1)로 구성하였으며, 입력층에 바이어스 항을 추가하여 연결강도를 갱신하였다. LRNN의 경우 (4x2x1)개의 뉴런으로 구성되었고, 시냅스 내부 계수는 MA=5, AR=6이다. 제안된 LGRNN의 경우에는 (4x2x1)개의 뉴런으로 구성되었고, 시냅스 내부 계수는 MA=4, AR=4이다. LRNN과 LGRNN에서 시냅스에 +1의 상수 바이어스가 사용되었으며, 바이어스항의 계수는 갱신하지 않았다.

따라서 실험에 사용된 각 신경망 모델의 전체 계수 수는 GRNN이 112개, LRNN이 110개, LGRNN이 112로서 신경망 계수의 수를 비슷하게 설정하였으며, 각 신경망의 변수 값은 시스템 오차가 최소가 되도록 실험을 통하여 결정하였다. 입력 신호 $s(k)$ 는 평균이 영이고 1.0의 표준편차를 갖는 백색 가우시안(Gaussian) 랜덤 신호이며, 학습과정에서 1000 샘플에 대해 안정된 수렴이 이루어질 때까지 1000번 학습을 수행하였다. 학습결과를 살펴보면 지역계환 신경망이 광역계환 신경망 보다 빠르게 수렴하였으며, 제안된 LGRNN은 LRNN보다 더 안정된 시스템 오차를 보여주었다.

그림 3과 4는 GRNN과 LGRNN 모델에 대해 샘플구간 100-200과 주파수 구간 0-0.5(rad)에서 시스템 식별 결과를 확대한 것으로서, 시간영역과 주파수 영역에서 선형 시스템 출력(점선으로 표시)과 각 신경망 출력(실선으로 표시)과의 차이를 확인할 수 있다.

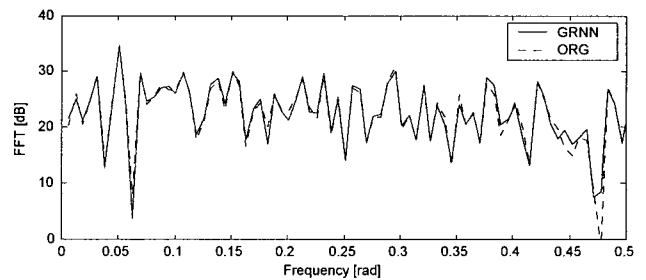


(a) GRNN

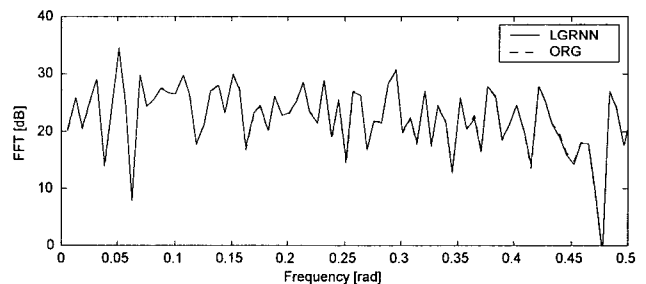


(b) LGRNN

그림 3. 샘플구간 100-200에서의 선형 시스템 식별결과
Fig. 3. Identification results of linear system in sample 100-200



(a) GRNN



(b) LGRNN

그림 4. 주파수 구간 0-0.5[rad]에서 선형시스템 식별결과

Fig. 4. Identification results of linear system in frequency 0-0.5[rad]

그림을 통해서 LGRNN 신경망이 GRNN 보다 우수함을 확인할 수 있으나 시스템 식별결과에 대한 정확한 성능차이를 정량적으로 조사하기 위해 다음 식을 사용하여 시간과 주파수 영역에서 오차를 상호 비교하였다.

$$t_err = \frac{1}{N} \sum_{k=N_1}^{N_2} |d(n) - \hat{d}(n)|^2 \quad (16)$$

$$f_err = \frac{1}{N} \sum_{k=N_1}^{N_2} ||fft(d(n)) - |fft(\hat{d}(n))||^2 \quad (17)$$

여기서 $d(n)$ 와 $\hat{d}(n)$ 는 미지 시스템과 신경망의 출력을 각각 나타낸다. 표 1은 각 신경망에 대해 위 식을 사용하여 계산된 결과로서 시스템의 계수 초기치가 서로 다른 조건에서 5번을 반복하여 평균한 값들이다. 시간영역에서 살펴보면 LGRNN에 의한 시스템 식별은 LRNN보다 1.228배 정도 우수하며, GRNN에 비해 16.019배 정도 우수하며, LRNN은 GRNN에 비해 13.043배 정도 우수하였다. 주파수 영역에서는 LGRNN은 LRNN보다 1.375배 정도 우수하며, GRNN에 비해 12.927배 정도 우수하며, LRNN은 GRNN에 비해 9.40배 우수하였다. 따라서 선형 시스템 식별에서 시간과 주파수 영역 모두 지역궤환 구조가 광역궤환 구조에 비해 훨씬 우수하며, 그리고 지역궤환 구조 신경망에서 LRGNN은 LRNN 보다 다소 우수한 성능을 보여주었다.

표 1. 선형 시스템 식별에서 시간영역과 주파수 영역에서 신경망 모델간의 오차 비교

Table 1. Error comparison of time and frequency domains among network models in linear system identification

신경망	오 차	
	t_err	f_err
GRNN	0.003300	1.2953
LRNN	0.000253	0.1378
LGRNN	0.000206	0.1002

3.2 비선형 시스템 식별

미지의 비선형 시스템 식별을 위해 아래와 같은 차분 방정식으로 정의되는 서로 다른 두 시스템을 사용하였는데, 첫 번째 시스템은 Back-Tsoi가 사용한 시스템이며, 두 번째 시스템은 선형 시스템 식별에 사용하였던 시스템에 비선형 특성을 추가하여 사용하였다.

비선형 시스템 #1 :

$$z(k) = 0.0154s(k) + 0.0462s(k-1) + 0.0462s(k-2) + 0.0154s(k-3) + 1.99z(k-1) - 1.572z(k-2) + 0.4583z(k-3) \quad (18a)$$

$$d(k) = \sin(z(k)) \quad (18b)$$

비선형 시스템 #2 :

$$d(k+1) = 0.15d(k) + 0.3d(k-1) + 0.1s^3(k) + 0.18s^2(k) + 0.921s(k) + 0.107s(k-1) + 0.336s(k-2) + s(k-3) + 0.336s(k-4) + 0.107s(k-5) + 0.921s(k-6) \quad (19)$$

비선형 시스템 식별에 사용된 신경망 구조와 계수의 수, 시스템 변수 값, 입력신호, 그리고 학습방법들은 선형 시스템 식별의 경우와 동일하게 사용하였다.

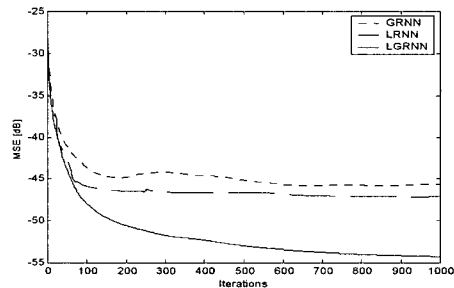
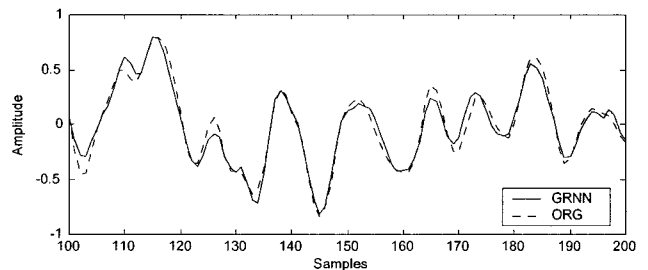


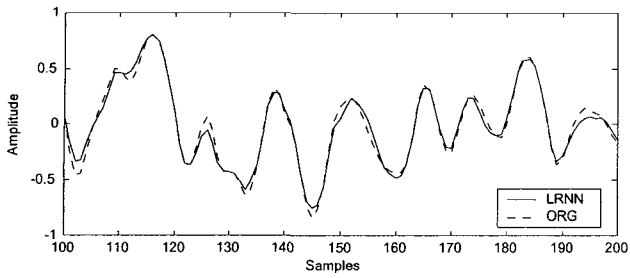
그림 5. 신경망 모델간의 학습오차 비교

Fig. 5. Comparison of learning error among neural network models

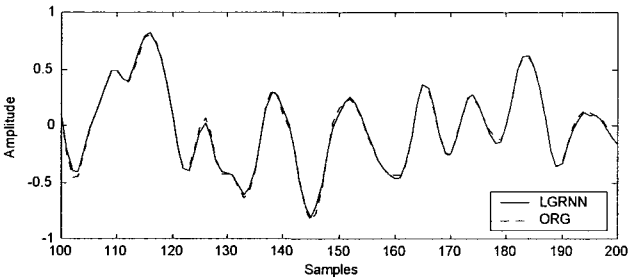
그림 5는 비선형 시스템 #1의 모델링에 대해 각 신경망의 학습오차 변화과정을 보여주고 있는데, LGRNN이 다른 리커런트 신경망보다 더 빠른 수렴과 안정된 시스템 오차를 보여주고 있다. 그림 6과 7은 GRNN, LRNN, 그리고 LGRNN 신경망 모델에 대해 시간영역과 주파수 영역에서 비선형 시스템 #1 식별 결과의 일부(시간영역에서의 샘플구간 100-200와 주파수 구간 0-0.5[rad])를 보여주고 있는데, 비선형 시스템 출력과 각 신경망 출력이 각각 점선과 실선으로 표시되어 있으므로 각 모델의 식별능력과 모델간의 식별차이를 확인할 수 있다.



(a) GRNN



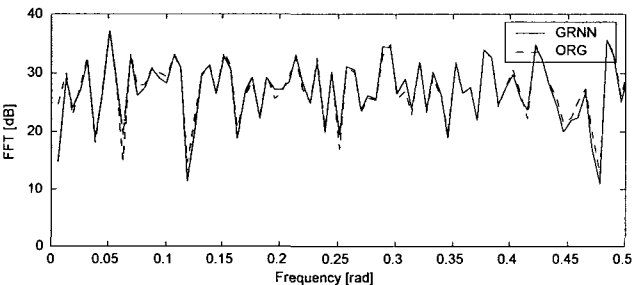
(b) LRNN



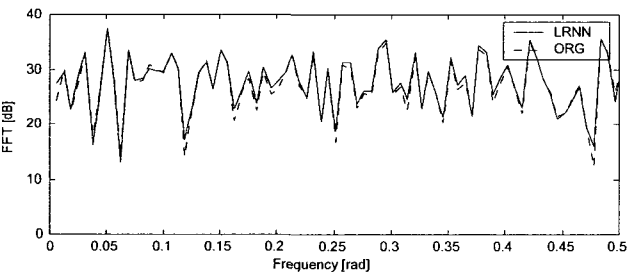
(c) LGRNN

그림 6. 샘플구간 100-200에서 비선형시스템 #1 식별결과

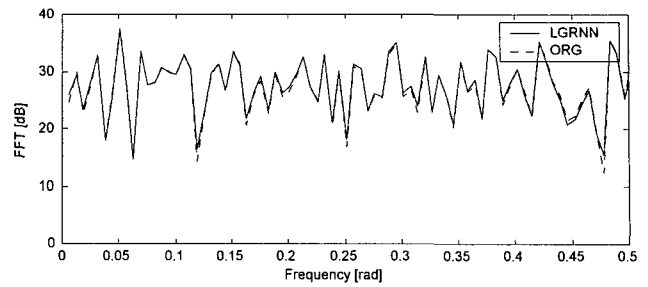
Fig. 6. Identification results of nonlinear system #1 in samples 100-200



(a) GRNN



(b) LRNN



(c) LGRNN

그림 7. 주파수 구간 0-0.5(rad)에서 비선형 시스템 #1 식별 결과

Fig. 7. Identification results of nonlinear system #1 in frequency 0-0.5[rad]

위의 그림으로부터 LGRNN이 더 우수한 시스템 모델링 성능을 보여주고 있음을 예상할 수 있으나 정확한 성능차이는 확인할 수 없다. 따라서 정량적 평가를 위해서 시간영역과 주파수 영역에서 식 (16)과 식 (17)을 사용하여 각 신경망에 대해 계산한 결과를 상호 비교하였다.

표 2. 비선형 시스템의 시간영역과 주파수 영역에서 신경망 모델간의 오차 비교

Table 2. Error comparison of time and frequency domains among network models in nonlinear system identification

신경망	비선형 시스템 #1		비선형 시스템 #2	
	t_{err}	f_{err}	t_{err}	f_{err}
GRNN	0.0052	3.8348	0.0055	2.5892
LRNN	0.0044	1.8113	0.0036	1.5213
LGRNN	0.0019	0.6894	0.0023	0.8041

표 2는 두 시험 모델에 대한 결과를 비교하고 있는데, 표에 나타난 수치는 시스템 계수의 초기치가 서로 다른 조건에서 5번 반복 수행 후 평균하였다. 시험 모델 #1에 대해 시간영역에서 LGRNN의 식별 성능은 GRNN과 LRNN에 비해 2.737배와 2.316배 정도 우수하며, 주파수 영역에서는 5.563배와 2.627배 정도 더 우수하였다. 그리고 LRNN의 식별 성능은 GRNN보다 시간영역에서 1.182배, 주파수 영역에서 2.117배 정도 우수하였다.

시험 모델 #2에 대해 시간영역에서 LGRNN의 식별 성능은 GRNN과 LRNN에 비해 2.391배와 1.565배 정도 우수하며, 주파수 영역에서는 3.220배와 1.892배 정도 더 우수하였다. 그리고 LRNN의 식별 성능은 GRNN보다 시간영역에서 1.528배, 주파수 영역에서 1.702배 정도 우수하였다. 선형 시스템 식별 경우처럼 지역 케환 구조를 갖는 시스템, 즉 LRNN과 LGRNN은 광역케환 구조를 갖는 GRNN에 비해 더 우수한 비선형 시스템 식별을 수행하였으며, 지역케환과 광역케환 구조를 모두 갖는 LGRNN이 LRNN보다 더 정확하게 비선형 시스템을 모델

링하였다.

3.3 시뮬레이션 결과

결과를 요약하면 선형과 비선형 시스템 식별에 대해 지역케환 신경망이 광역케환 신경망 보다 성능이 우수하였고, 지역케환 중에서도 LGRNN이 LRNN보다 더 우수하였다. 즉, 서로 다른 시험 모델에 대해 제안된 LGRNN은 시스템 오차면에서 GRNN과 LRNN 신경망 보다 더 우수한 모델링 성능을 갖고 있음을 확인하였다. 특히, 신경망간의 상대적인 오차 비교에서 선형 시스템 식별의 경우 비선형 시스템에 비해 지역케환이 광역케환보다 훨씬 더 우수한 성능을 보여주었으며, 비선형 시스템 식별의 경우 선형 시스템에 비해 LGRNN이 LRNN 보다 더 우수한 식별을 수행하였다.

이와 같은 결과는 시간영역과 주파수 영역 모두에서 비슷한 성능 차이를 보여주었다. 신경망의 오차수렴 속도 면에서도 LGRNN이 LRNN과 GRNN에 비해 전반적으로 더 안정되고 빠른 수렴특성을 보여주었다. 따라서 제안된 신경망은 선형 시스템 뿐만 아니라 비선형 시스템의 식별을 수행하는데 케환 구조를 갖는 기존의 다른 다층 신경망보다 더 우수한 신경망 모델이라 할 수 있다.

실험과정을 통해서 나타난 내용으로는 비선형 시스템 식별에 대해서 광역케환 구조의 GRNN은 시험 모델에 따라서 지역케환 구조의 LGRNN 보다 시스템 성능에서 더 민감한 결과를 보여 주었다. 그리고 신경망 학습과정에서 GRNN은 수렴도중에 진동이 발생하거나 학습이 진행됨에 따라서 발산하는 경우가 있었으며, LRNN의 경우 진동과 발산현상은 존재하지 않았다. 그리고 LGRNN의 경우 미세한 진동이 존재하였으나 회수와 크기는 GRNN에 비해 크게 작았으며 발산 현상은 발생하지 않았다.

IV. 결론

본 논문에서는 기존의 다층 리커런트 신경망보다 동특성이 향상된 hybrid 신경망인 LGRNN을 사용하여 시스템 식별을 수행하였다. Hybrid 신경망은 뉴런간의 케환 시냅스 구조를 갖는 지역케환 신경망과 층간의 케환구조를 갖는 광역케환 신경망을 결합한 합성구조 신경망으로서, 지역케환 신경망으로는 IIR-MLP를, 광역케환 신경망으로는 Elman RNN 신경망 구조를 각각 사용하였다.

LGRNN의 성능을 조사하기 위해 선형 및 비선형 시스템 식별을 수행하였으며, 다른 신경망과의 상대적인 평가를 위해 Elman RNN과 IIR-MLP 신경망과 상호 비교하였다. 시뮬레이션 결과에 의하면 제안된 신경망은 학습 수렴속도와 시스템 정확도(안정성)에서 기존의 다층 리커런트 신경망보다 더 우수한 시스템 식별능력을 보여줌으로써 LGRNN 신경망은 시스템 식별에 효과적인 모델로 사용할 수 있음을 확인하였다. LGRNN 신경망은 시스템의 구조가 다소 복잡해지는 단점이 있지만 시스템 동특

성이 기존의 다층 케환 신경망에 비해 향상된 신경망이므로 더 높은 차수의 비선형 및 비정적 시변 신호처리뿐만 아니라 temporal 동특성이 요구되는 다른 분야에서도 사용할 수 있는 다층 리커런트 신경망이라 할 수 있다.

참고 문헌

- [1] R. Grino, G. Cembrano, and C. Torras, "Nonlinear system identification using additive dynamic neural networks - two on-line approaches", IEEE Trans. on circuits and systems a publication of the IEEE Circuits and System Society. I, Fundamental theory and applications, vol. 47, no. 2, pp. 150-165, 2000.
- [2] A. Yazdizadeh and K. Khorasani, "Adaptive time delay neural network structures for nonlinear system identification", Neurocomputing, vol. 47, no. 1/4, pp. 207-240, 2002.
- [3] P. M. Ferreira, E. A. Faria, and A. E. Ruano, Yazdizadeh and K. Khorasani, "Neural network models in greenhouse air temperature prediction", Neurocomputing, vol. 43, no. 1/4, pp. 51-75, 2002.
- [4] H. G. Choi and E. B. Shim, "Enhancement of QRS complex using a neural network based ALE", 22nd Int. Conf. of the IEEE EMB, World Congress on Medical Physics and Biomed. Eng., vol. 27, no. 6, Chicago, U.S.A., Jul. 2000.
- [5] S. Haykin, Neural Networks: A Comprehensive Foundation, IEEE Press, New York, 1994.
- [6] A. D. Back and A. C. Tsoi, "FIR and IIR synapses, a new neural network architecture for time series modeling", Neural Comput., vol. 3, pp. 375-385, 1991.
- [7] E. A. Wan, "Temporal backpropagation for FIR neural networks", in Proc. Int. Joint Conf. Neural Networks, vol. 1, pp. 575-580, 1990.
- [8] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks", Neural Comput., vol. 1, pp. 270-280, 1989.
- [9] J. L. Elman, "Finding structure in time", Cognitive Science, vol. 14, pp. 179-211, 1990.
- [10] M. Jordan, "Serial order: A parallel distributed processing approach", Univ. California San Diego, Inst. Cognitive Science, ICS Rep. 8604, 1986.
- [11] P. Campolucci, A. Uncini, F. Piazza, and B. D. Rao, "On-line learning algorithm for locally recurrent neural networks", IEEE Trans. Neural

Networks, vol. 10, pp. 253-271, 1999.

- [12]A. C. Tsoi and A. D. Back, "Locally recurrent globally feedforward neural networks: A critical review of architectures", IEEE Trans. Neural Networks, vol. 5, pp. 229-239, 1994.



최 한 고 (Han-Go Choi)
1979년 경북대학교 전자공학과 (공학사)
1988년 Univ. of Florida, USA
전기 및 전자공학과 (공학석사)
1992년 Univ. of Florida, USA
전기 및 전자공학과 (공학박사)

2001년 MIT, USA (Research Scientist)
1979년 ~ 1986년 국방과학연구소
1993년 ~ 현재 금오공과대학교 전자공학부 교수
관심분야: 디지털 및 적응신호처리, 패턴인식, 신경회로망



고 일 환 (Il-Whan Go)
2002년 금오공과대학교
전자제어공학과 (공학사)
2004년 금오공과대학교
전자제어공학과 (공학석사)
2004년 ~ 현재 삼성전자 무선개발팀
연구원

관심분야: 디지털 신호처리, 회로설계, 신경회로망



김 중 인 (Jong-In Kim)
2000년 금오공과대학교
전자제어공학과 (공학사)
2004년 금오공과대학교
전자제어공학과 (공학석사)
2004년 ~ 현재 금오공과대학교
전자제어공학과 박사과정중

관심분야: 디지털 및 적응신호처리, NN & Biometrics
Computer Network
