

論文2001-38SC-2-2

시간-주파수 분석을 이용한 모듈라 웨이블릿 신경망의 최적 구조 설계

(On Designing Optimal Structure of Modular Wavelet Neural Network with Time-Frequency Analysis)

徐在龍*, 金鏞宅*, 趙鉉讚**, 全洪兌**

(Jae-Yong Seo, Yong-Taek Kim, Hyun-Chan Cho and Hong-Tae Jeon)

요약

본 논문에서는 새로운 구조의 모듈라 시스템의 최적구조 설계 알고리즘을 제안하였다. 모듈라 시스템은 구조의 단순화와 시간-주파수 분석법을 이용하기 위해 웨이블릿 신경망으로 구성하였다. 제안한 최적구조 설계 알고리즘을 이용하여 근사화 대상함수의 시간-주파수 특성을 분석하여 모듈의 개수와 부-시스템의 노드의 개수를 결정할 수 있다. 제안한 최적 구조 설계 알고리즘은 시스템의 특성을 분석하여 모듈라 웨이블릿 신경망의 최적구조를 설계할 수 있는 방법론을 제공할 수 있다. 제안한 새로운 구조와 최적 구조 설계 알고리즘을 근사화 문제에 적용하여 우수성을 검증하였다.

Abstract

In this paper, we propose the new algorithm which can design on the optimal structure of modular system. This system is composed to the wavelet neural network in order to simplify the structure of modular system and use the time-frequency analysis. We will determine the number of module and node of each sub-system using the proposed algorithm. This algorithm provides the methodology, which we will design optimal structure of modular wavelet neural network through analyzing the character of system. We apply the proposed new structure and algorithm to approximation problem and evaluate the effectiveness of the proposed system and algorithm.

I. 서론

일반적으로 신경망은 새로운 환경에 적응하기 위해서 이전에 학습한 매핑 결과를 잊어버리는 파멸간섭(catastrophic interference) 현상이 발생한다. 이러한 현

상이 발생하는 이유는 신경망이 학습을 수행할 때 매우 연관된 전역 매개변수를 갱신하기 때문이다. 이러한 현상을 방지하기 위해서 신경망의 연관성과 비학습성을 줄이거나 지역화된 응답 특성을 갖는 모듈화된 신경망을 사용한다^[1,4,5].

모듈화된 시스템은 복잡한 문제(complex problem)를 단순한 부-문제(sub-problem)로 나누어 해결하는 방식이다. 따라서 단순한 부-문제를 해결하기 때문에 학습 속도가 향상되고, 사용되는 파라미터가 감소하며 표현 능력을 증가시키고, 구조의 단순화로 하드웨어 구현이 용이하다. 또한 모듈화는 전체 시스템에서 계산의 지역화를 달성함으로써 전체 시스템의 행동에서 부-시스템의 역할을 정의할 수 있다. 그리고 모듈화 설계를 통해 신경망의 다른 활성화 함수나 다른 망 구조를 사용할

* 正會員, 中央大學校 電子工學科

(Dept. of Electronic Eng. Chung-Ang Univ.)

** 正會員, 韓國技術教育大學校 電子工學科

(Dept. of Electronic Eng. Korea Univ. of Tech. and Edu.)

※ 본 연구는 과학기술부의 뇌과학연구 프로그램의 지원 하에 수행되었습니다.

接受日字:2000年6月26日, 수정완료일:2001年1月29日

수 있고, 다른 종류의 학습 알고리즘을 사용할 수 있는 장점을 가지고 있다^[1,4].

현재 모듈화된 시스템에 대한 이론적인 토대를 마련하고 실제 시스템에 적용하기 위한 응용사례가 많이 등장하고 있다^[6~8]. 특히 Jacobs와 Jordan^[1,9~11]이 제안한 혼합형 모듈라 신경망(Mixed Modular Neural Network; MMNN)은 작업재구성(task decomposition)과 학습(learning), 재결합(recombination)의 관점에서 탁월한 성능을 발휘하고 있으며, 현재 가장 많이 적용되고 있는 모듈화된 시스템이다^[12,13]. 그러나 단일 신경망과 유사하게 모듈화된 부-시스템의 크기나 파라미터의 초기화 문제에 대한 연구는 아직도 미흡한 실정이다. 지금까지 대부분의 연구는 모듈라 신경망의 모듈을 신경망 이외의 다른 시스템으로 구성^[6,7,14], 작업재구성^[15], 학습 알고리즘^[16]과 재결합 방식^[17]에 대해 집중되어 있었다. Viswanath^[18]는 모듈화된 망의 최적구조를 얻기 위해 성장(growing)과 전지(pruning) 알고리즘을 제안했지만 이 알고리즘은 입출력 쌍을 분석하거나 시스템의 특성을 분석하는 이론적인 접근방식보다는 시행착오 방식을 이용하였다.

따라서 본 논문에서는 새로운 구조인 모듈라 웨이블릿 신경망을 제안하고, 최적 구조 설계 알고리즘을 개발하고자 한다. 모듈라 웨이블릿 신경망은 각 모듈을 웨이블릿 신경망으로 구성하여 시간-주파수 분석이 용이하도록 설계한다. 웨이블릿 신경망의 최적의 노드 개수와 기저함수를 선택하기 위해 근사화 대상 함수나 시스템의 입출력 데이터를 이용하여 시간-주파수 분석(time-frequency analysis)을 수행한다. 선행처리(pre-processing) 단계로서 부-시스템인 모듈을 구성하기 위한 웨이블릿 신경망의 노드 개수뿐만 아니라 원하는 작업에 가장 적합한 모듈의 개수를 결정할 수 있다. 제안한 새로운 구조의 모듈라 웨이블릿 신경망과 최적구조 설계 알고리즘을 함수 근사화 문제에 적용하여 새로운 구조의 우수성과 최적구조 설계 알고리즘의 효율성을 입증하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 모듈라 신경망의 새로운 구조인 모듈라 웨이블릿 신경망을 소개하고, 근사화 대상 함수나 시스템의 시간-주파수 특성을 분석하여 모듈라 웨이블릿 신경망의 최적구조를 설계하기 위한 알고리즘을 제안한다. 3장에서는 제안한 새로운 모듈라 웨이블릿 신경망과 최적구조 설계 알고리즘을 함수근사화 문제에 적용하여 효율성을 검토하

고 마지막으로 결론에 대해서 기술한다.

II. 모듈라 웨이블릿 신경망의 최적 구조 설계

모듈라 웨이블릿 신경망의 구조는 그림 2.1과 같다.

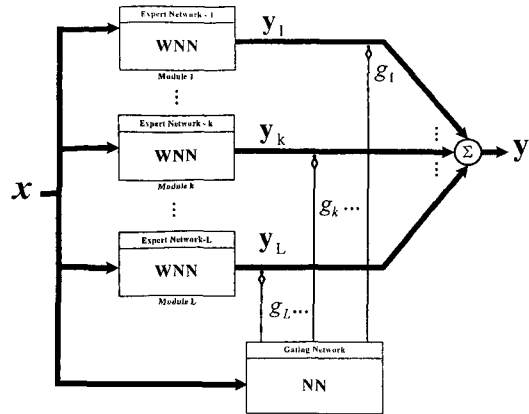


그림 1. 모듈라 웨이블릿 신경망
Fig. 1. Modular Wavelet Neural Network.

각 모듈을 구성하는 엑스퍼트 네트워크는 최적구조를 설계하기 위한 이론적인 배경을 제공하는데 용이하도록 다층 신경망 대신에 웨이블릿 신경망으로 구성하였다. 그러나 게이팅 네트워크는 기존의 구조인 신경망을 이용하여 구성하였다. 엑스퍼트 네트워크와 게이팅 네트워크의 최적구조를 동시에 구하는 알고리즘을 설계할 경우, 각각의 기능 차이 때문에 문제의 복잡도가 증가하고, 최적의 구조를 구할 가능성이 감소하게 된다. 따라서 본 논문에서는 모듈라 네트워크의 대부분의 크기를 차지하고 있는 엑스퍼트 네트워크의 최적구조를 설계하는 알고리즘만 연구하였으며, 게이팅 네트워크는 단일신경망과 다층 신경망으로 구성하여 수행결과를 비교한다.

1. 모듈라 웨이블릿 신경망(MWNN)

1) k번째 엑스퍼트 네트워크

MWNN에서 차원이 P인 입력벡터는 $x = [x_1 \ x_2 \ \dots \ x_P]^T$ 이고, MWNN의 최종출력인 $Q \times 1$ 벡터는 $y = [y^1 \ y^2 \ \dots \ y^Q]^T$ 이다. 또한 k번째 엑스퍼트 네트워크 출력벡터는 $y_k = [y_k^1 \ y_k^2 \ \dots \ y_k^{Q_k}]^T$ 이고, 원하는 응답 벡터는 $d = [d^1 \ d^2 \ \dots \ d^Q]^T$ 이다. l번째 노드의 활성

화 함수로 사용되는 웨이블릿 함수는 식 (2.1)과 같다.

$$\psi_{k, m, l}^l(\mathbf{x}) = 2^{\frac{m}{2}} \phi_k^l(\mathbf{ax} - n_l) \quad (2.1)$$

여기서 $m, n_l \in \mathbf{Z}$, m 은 팽창이고, n_l 은 l 번째 이동이며, 팽창 벡터 $\mathbf{a} = [2^m 2^m \dots 2^m]$ 이다.

k 번째 엑스퍼트 네트워크의 출력은 아래의 식과 같이 가중치와 웨이블릿 함수의 곱의 합으로 얻을 수 있다.

$$y_k^j = \sum_{m=1}^M w_{k, m}^j \psi_{k, m, m_l}^l(\mathbf{x}), \quad \begin{cases} j = 1, 2, \dots, Q \\ k = 1, 2, \dots, L \end{cases} \quad (2.2)$$

2) 게이팅 네트워크

일반적으로 게이팅 네트워크는 신경망으로 구성한다. 적용 대상에 따라 단일 신경망(SNN), 출력층에 활성화 함수를 사용하지 않고 은닉층과 출력층의 가중치와 은닉층의 출력의 곱의 합인 선형조합으로 출력층을 구성한 다층 신경망(OLCMLNN), 출력층에 비선형 활성화 함수를 사용하는 다층 신경망(OAFMLNN) 등으로 구성한다.

게이팅 네트워크의 i 번째 뉴런의 출력 u_k 는 가중치 벡터 $\mathbf{a}_k = [a_k^1 a_k^2 \dots a_k^L]$ 와 입력 벡터 \mathbf{x} 의 곱의 합으로 표현된다.

$$u_k = \sum_{i=1}^L a_k^i x_i = \mathbf{a}_k \mathbf{x} \quad (2.3)$$

또한 게이팅 네트워크의 최종 활성화 출력인 g_k 는 식 (2.2)와 같이 사전확률(priori probability)로 고려한다^{9, 10, 11}. 모듈라 웨이블릿 신경망의 최종 출력은 아래와 같다.

$$\mathbf{y} = \sum_{k=1}^K g_k \mathbf{y}_k \quad (2.4)$$

2. 확률적인 강화 학습 알고리즘 유도

통계적인 강화학습의 목표는 주어진 학습패턴 집합의 분포를 모델링하는 것이다. 즉, 시스템의 입출력 쌍이나 근사화 대상함수의 샘플 데이터를 이용하여 원하는 응답함수를 모듈화된 웨이블릿 신경망이 학습하도록 가우스 혼합형 밀도함수를 비용함수로 정의하고, 강화학습 알고리즘을 이용하는 것이다.

1) 엑스퍼트 네트워크 학습알고리즘

웨이블릿 신경망으로 구성된 엑스퍼트 네트워크의 미지의 매개변수는 은닉층과 출력층 사이의 가중치 벡터 \mathbf{w}_k 이다. 따라서 비용함수에 대한 \mathbf{w}_k 의 변화정도인 민감도를 구하여 강화 학습에 이용할 수 있다. 식 (2.5)는 비용함수에 대한 k 번째 모듈의 j 번째 가중치 $w_k^j = [w_{k1}^j w_{k2}^j \dots w_{kM}^j]^T$ 에 대한 변화량을 체인룰을 적용하여 나타낸 것이다.

$$\frac{\partial l(\mathbf{w}, \mathbf{g})}{\partial \mathbf{w}_k^j} = \frac{\partial l(\mathbf{w}, \mathbf{g})}{\partial y_k^j} \cdot \frac{\partial y_k^j}{\partial \mathbf{w}_k^j} \quad (2.5)$$

기울기 상승(gradient ascent) 방법을 이용한 미소조정 값은 식 (2.6)과 같다.

$$\Delta \mathbf{w}_k^j = \eta \frac{\partial l(\mathbf{w}, \mathbf{g})}{\partial \mathbf{w}_k^j} \quad (2.6)$$

k 번째 모듈의 웨이블릿 신경망의 시냅스 가중치는 식 (2.7)에 의해서 매 샘플링 시간마다 갱신된다.

$$\mathbf{w}_k^j(t+1) = \mathbf{w}_k^j(t) + \eta \frac{h_k}{\sigma_k} e_k^j \psi_{k, m, m_l}^l(\mathbf{x}) \quad (2.7)$$

2) 게이팅 네트워크 학습 알고리즘

기존의 연구된 결과는 단일 신경망으로 엑스퍼트와 게이팅 네트워크를 구성^{1,9}하거나, 다층 신경망으로 구성할 경우 활성화 함수로 tanh 함수를 사용하였다^{13, 15, 18}.

다음은 게이팅 네트워크를 다층신경망을 구성하였을 경우를 고려한다. OLCMLNN의 경우, 다층 신경망의 입력층과 은닉층의 가중치 $a_m^i(t)$ 갱신규칙은 식 (2.8)과 같고, 은닉층과 출력층의 가중치 $a_k^m(t)$ 의 갱신규칙은 식 (2.9)과 같이 얻을 수 있다.

$$a_m^i(t+1) = a_m^i(t) + \eta \cdot \delta_m^g \cdot x_i \quad (2.8)$$

$$\begin{cases} m = 1, 2, \dots, N \\ i = 1, 2, \dots, P \end{cases}$$

$$a_k^m(t+1) = a_k^m(t) + \eta \cdot \delta_k^g \cdot f(net_m^g) \quad (2.9)$$

$$\begin{cases} k = 1, 2, \dots, L \\ m = 1, 2, \dots, N \end{cases}$$

여기서 활성화 함수를 sigmoid 함수를 사용한 경우와 tanh 함수를 사용한 두 가지 경우로 구분하여 은닉층의 출력과 출력층의 출력을 정의할 수 있다. sigmoid 함수를 은닉층의 활성화 함수로 사용한 경우의 은닉층

의 출력은

$$\delta_k^g = (h_k - g_k) \cdot u_k \cdot (1 - u_k) \quad (2.10)$$

이고, tanh 함수를 사용한 경우 은닉층의 출력은

$$\delta_k^g = (h_k - g_k) \cdot (1 - (u_k)^2) \quad (2.11)$$

이다. 또한 출력층의 출력은 다음과 같이 정의할 수 있다.

$$\delta_m^g = \delta_k^g \cdot a_k^m \quad (2.12)$$

그리고 OAFMLNN의 경우(출력층에 활성화 함수 사용)에도 위와 유사한 방법을 이용하여 가중치 갱신규칙을 얻을 수 있다.

3. 시간-주파수 분석^[2,3]

시간축 상에서 신호의 지역적인 주파수 행동을 분석하기 위해 신호를 시간과 주파수에 한정(time-frequency localization)된 함수로 표현하고, 시간과 주파수 특성을 동시에 고려하여 신호를 분석한다. 시간 한정 중심 $t_c(f)$ 와 폭 σ_t 는 다음 식 (2.13), (2.14)와 같이 정의된다.

$$t_c(f) = \frac{1}{\|\hat{f}\|^2} \int_{-\infty}^{\infty} t|\hat{f}(t)|^2 dt \quad (2.13)$$

$$\sigma_t = \frac{1}{\|\hat{f}\|} \left\{ \int_{-\infty}^{\infty} (t - t_c(f))^2 |\hat{f}(t)|^2 dt \right\}^{\frac{1}{2}} \quad (2.14)$$

또한 주파수 한정의 중심 $\omega_c(\hat{f})$ 와 폭 σ_ω 는 식 (2.15), (2.16)과 같이 정의된다.

$$\omega_c(\hat{f}) = \frac{1}{\|\hat{f}\|^2} \int_0^{\infty} \omega|\hat{f}(\omega)|^2 d\omega \quad (2.15)$$

$$\sigma_\omega = \frac{1}{\|\hat{f}\|} \left\{ \int_0^{\infty} (\omega - \omega_c(\hat{f}))^2 |\hat{f}(\omega)|^2 d\omega \right\}^{\frac{1}{2}} \quad (2.16)$$

척도구성 함수, 웨이블릿 함수와 웨이블릿 프레임의 시간-주파수 원자는 다음과 같이 표현되며,

$$W = [t_c - \sigma_t, t_c + \sigma_t] \times [\omega_c - \sigma_\omega, \omega_c + \sigma_\omega] \quad (2.17)$$

이산화된 경우의 시간-주파수 원자는 다음 식 (2.24)과 같다.

$$W_{m,n} = [2^{-m}(t_c + nb - \sigma_t), 2^{-m}(t_c + nb + \sigma_t)] \times [2^m(\omega_c - \sigma_\omega), 2^m(\omega_c + \sigma_\omega)] \quad (2.18)$$

4. 시간-주파수 분석법을 이용한 최적구조 설계

주어진 학습 데이터나 시스템의 입출력 쌍의 집합에 대한 시간-주파수 분석을 수행하여 문제를 해결하는데 적합한 해를 구하기 위한 엑스퍼트 네트워크의 최적 구조를 구한다. 즉, 엑스퍼트 네트워크를 구성하는 웨이블릿 신경망의 활성화 함수인 웨이블릿 함수의 팽창과 이동을 시스템의 시간-주파수 분석을 이용하여 결정한다. 또한 문제해결에 사용될 최적의 모듈개수를 결정한다.

근사화 대상함수나 시스템의 시간-주파수 분석은 다음과 같은 단계로 이루어진다.

[단계 1] 엑스퍼트 네트워크를 방사기저함수망과 유사한 웨이블릿 신경망(WNN)으로 구성하고, 원하는 작업에 적합한 엑스퍼트 네트워크의 크기를 결정하기 위해 입출력 데이터의 시간-주파수 분석을 수행한다. 입출력 데이터의 시간-주파수 윈도우 Q_f 는 데이터의 시간과 푸리에 변환을 이용하면 다음 식 (2.19)와 같이 얻을 수 있다.

$$Q_f = [t_{\min}, t_{\max}] \times [\omega_{\min}, \omega_{\max}] \quad (2.18)$$

여기서 t_{\min} 과 t_{\max} 는 함수나 시스템의 시간축 상의 한정된 영역의 최소, 최대 값을 의미하고, ω_{\min} 과 ω_{\max} 는 주파수축 상의 한정된 영역의 최소, 최대 값을 나타낸다.

[단계 2] 웨이블릿 신경망의 활성화 함수를 웨이블릿 함수로 선택하고, 웨이블릿 함수에 해당되는 시간-주파수 원자를 식 (2.20)를 이용하여 얻는다. 식 (2.20)은 웨이블릿 프레임 함수($a_0=2, b_0=b$)를 선택하고 푸리에 변환의 팽창 특성(dilation property)인 $\hat{f}(ax) \xrightarrow{F} a^{-1} \hat{f}(a^{-1}\omega)$ 을 이용하고, $\psi_{m,n}$ 을 팽창, 이동의 값을 변화시켜 얻은 시간-주파수 원자를 의미한다.

$$Q_{m,n} = [2^{-m}(t_c + nb - \sigma_t), 2^{-m}(t_c + nb + \sigma_t)] \times [2^m(\omega_c - \sigma_\omega), 2^m(\omega_c + \sigma_\omega)] \quad (2.19)$$

[단계 3] 식 (3.34)를 이용하여 얻은 Q_f 의 시간-주파수 윈도우와 식 (2.20)을 이용하여 얻은 시간-주파수 원자 $Q_{m,n}$ 을 동일한 시간-주파수 평면에 작도하고, Q_f 윈도우에 속하는 모든 $Q_{m,n}$ 를 선택하여 웨이블릿 함수 군을 선택한다.

[단계 4] 단계 3에서 선택한 웨이블릿 함수군의 팽창 지수(dilation index)의 개수만큼 모듈을 구성한다. 즉, 시간-주파수 분석결과 선택된 웨이블릿 함수군의 팽창 계수가 $m=2,3,4$ 인 경우, 3개의 모듈을 구성한다. 그리고 각 모듈을 구성하는 웨이블릿 신경망의 노드는 해당모듈(특정 팽창계수)에 대해 시간축 상에서 결정된 이동계수의 인덱스 집합을 이용하여 결정한다.

위의 4단계를 통해 선행처리(preprocessing) 단계로 엑스퍼트 네트워크의 개수인 모듈의 개수를 결정하고, 엑스퍼트 네트워크에 사용되는 웨이블릿 신경망의 노드 수를 원하는 작업에 적합하도록 최적의 개수로 선택한다. 즉, 최적해를 구할 수 있는 모듈라 웨이블릿 신경망의 최적구조를 결정할 수 있다.

III. 모의실험 및 검토

모의실험에서는 비주기적인 근사화 대상 함수를 사용했다. 근사화 대상 함수는 식 (3.1)과 같으며 시간축 상에서 $[0, 1]$ 을 고려했다.

$$f(t) = \sin(t) \cos(5t) \sin(10t) \cos(30t) \sin(50t) \quad (3.1)$$

시간-주파수 분석 [단계 1]의 과정을 통해 근사화 대상 함수의 시간-주파수 평면의 분포를 구한다. 먼저 근사화 대상 함수를 푸리에 변환을 통해 주파수 상의 분포를 분석한다. 본 논문에서는 양의 주파수 영역만을 고려한다. 근사화 대상 함수의 유효 주파수 분포는 정규화된 진폭(amplitude)의 63%로 설정했으며, $[\omega_{min}, \omega_{max}] = [6.2832, 43.9823] \times [69.1150, 106.8142]$

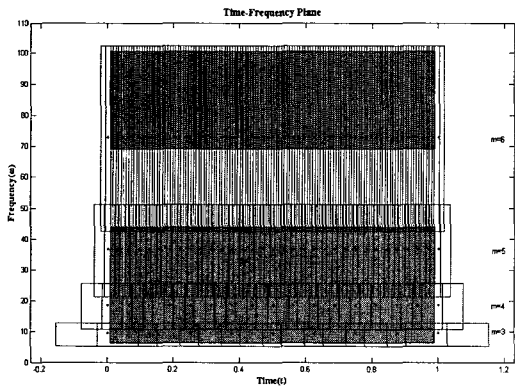


그림 2. 근사화 대상 함수와 Frame2의 시간-주파수 표현
Fig. 2. Time-frequency representation of approximation function and Frame2.

이다. 또한 시간상의 근사화 대상 함수의 분포는 $[t_{min}, t_{max}] = [0, 1.0]$ 이다.

그림 2는 근사화 대상 함수의 시간-주파수 표현과 Frame2의 시간-주파수 표현을 동일한 시간-주파수 평면에 나타낸 것이다.

표 1은 그림 2의 결과를 이용하여 모듈라 웨이블릿 신경망을 구성한 예를 나타낸 것이다. 웨이블릿 프레임이나 척도구성 함수를 사용할 경우 4개의 모듈로 구성할 수 있고, 웨이블릿 함수를 사용할 경우 5개의 모듈로 구성할 수 있다.

표 1. 모듈라 웨이블릿 신경망의 구성 예
Table 1. Example of constructing MWNN (Frame1, Frame2, Sinc Scaling).

	Module 1	Module 2	Module 3	Module 4
팽창(m)	3	4	5	6
이동(n)	0~7	0~15	0~31	0~63
노드 개수	8	16	32	64

표 2. 모듈라 웨이블릿 신경망의 구성 예 (Mexicanhat Wavelet)
Table 2. Example of constructing MWNN (Mexicanhat Wavelet).

	Module 1	Module 2	Module 3	Module 4	Module 5
팽창(m)	2	3	4	5	6
이동(n)	0~3	0~7	0~15	0~31	0~63
노드 개수	4	8	16	32	64

위와 같은 방법의 선행처리 단계로 모듈라 웨이블릿 신경망의 최적 구조를 결정한 후 대상 함수의 근사화를 수행하기 위해 확실적인 강화 학습 방법으로 모의 실험을 수행했다. 본 모의실험은 웨이블릿 신경망의 활성화 함수로 4종류의 웨이블릿 함수를 사용하여 수행 능력을 비교해 보았으며, 게이팅 네트워크를 SNN, OLCMLNN과 OAFMLNN으로 구분하여 구성하고 수행결과를 확인하였다. 또한 동일한 노드 수나 2배 정도의 노드수로 구성된 다층 신경망의 수행결과와 비교하여 모듈라 웨이블릿 신경망의 근사화 능력의 우수성을 검증하였다.

그림 3은 모듈라 웨이블릿 신경망과 동일한 크기(노

드 수 : 128)로 구성된 다층 신경망으로 함수 근사화를 수행한 결과이다. OLCMLNN과 OAFMLNN의 결과를 비교하여 나타낸 것이다. 학습률은 시행착오를 거쳐 가장 우수한 수행결과를 보이는 값을 선택하여 $\eta = 0.025$ 로 설정했다. 학습률을 모듈라 웨이블릿 신경망을 이용한 모의실험 값과 동일하게 설정 ($\eta = 0.001$)할 경우 함수근사화에 실패하였다.

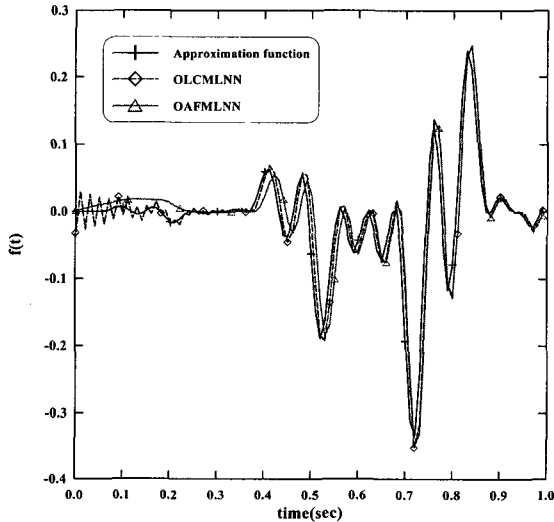


그림 3. 다층 신경망의 근사화 결과
Fig. 3. Approximation results of MLNN.

표 3은 2000번 반복수행 동안에 얻은 평균제곱근 에러의 최소 값을 나타낸 것이다. 다층 신경망의 노드 수를 증가시켜도 수행 결과(RMSE)가 크게 향상되지 않음을 확인할 수 있다.

표 3. 다층 신경망의 수행 결과(RMSE)
Table 3. Learning results of MLNN(RMSE).

		OLCMLNN	OAFMLNN
RMS	a number of node : 128	0.0276	0.0360
최소값	a number of node : 200	0.0273	0.0359

그림 4~5는 웨이블릿 시간-주파수 분석법에 의해 최적 구조로 구성된 모듈라 웨이블릿 신경망을 이용하여 근사화를 수행한 결과를 나타낸 것이다. 엑스퍼트 네트워크는 4종류의 다른 활성화 함수를 사용하였고, 게이팅 네트워크는 각각 OLCMLNN(그림 4), OAFMLNN(그림 5)를 사용하여 모듈라 웨이블릿 신경망을 구성하고 모의실험을 수행한 것이다. 수행조건은

다층 신경망과 동일하며, 학습률은 구성된 신경망의 구조에 따라 적합한 값이 다르게 설정되기 때문에 다층 신경망을 이용한 모의실험에서 사용한 값과 다른 값을 사용했다. 시행착오를 거쳐 비교적 수행능력이 우수한 값을 선택하여 $\eta = 0.001$ 로 설정하였다.

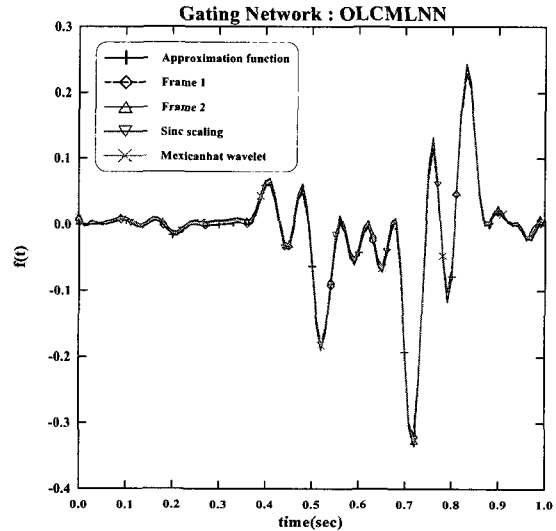


그림 4. 모듈라 웨이블릿 신경망의 학습 결과
Fig. 4. Learning results of MWNN.

표 4는 2000번 반복수행 동안에 얻은 평균제곱근 에러의 최소 값을 나타낸 것이다. 위의 결과와 표 4에서 확인할 수 있듯이 수행결과가 가장 우수한 경우는 모듈라 웨이블릿 신경망의 활성화 함수로 Frame1을 사용하고 게이팅 네트워크를 출력층이 비선형인 다층 신경망으로 구성한 경우이다.

표 4. 4종류의 다른 활성화 함수를 사용할 경우 MWNN의 RMS 최소값
Table 4. Minimum RMS of MWNN using 4 class of activation function.

		게이팅 네트워크 종류	Frame1	Frame2	Sinc Scaling	Mexicanhat Wavelet
RMS 최소값	SNN		0.0015	0.0059	0.0030	0.0043
	OLCMLNN		0.0029	0.0059	0.0030	0.0043
	OAFMLNN		0.0012	0.0065	0.0031	0.0049

IV. 결 론

본 논문에서는 모듈라 신경망의 새로운 구조인 모듈

라 웨이블렛 신경망(MWNN)을 제안하였다. 제안한 모듈라 웨이블렛 신경망은 혼합형 모듈라 신경망과 유사한 구조로 엑스퍼트 네트워크에 해당하는 각 모듈을 웨이블렛 신경망으로 구성하고 게이팅 네트워크는 단일 신경망과 다층 신경망으로 구성하였다. 그리고 통계적인 강화학습 알고리즘을 유도하였다. 또한 제안한 모듈라 웨이블렛 신경망의 최적구조를 설계할 수 있는 알고리즘을 제안하였다.

제안한 최적구조 설계 알고리즘은 웨이블렛 신경망의 최적의 노드 개수와 기저함수를 선택하기 위해 근사화 대상 함수나 시스템의 입력력 데이터를 이용하여 시간-주파수(time-frequency) 분석을 수행한다. 이 알고리즘은 선행처리(pre-processing) 단계로서 부-시스템인 모듈을 구성하기 위한 웨이블렛 신경망의 노드 개수뿐만 아니라 원하는 작업에 가장 적합한 모듈의 개수를 결정할 수 있는 장점이 있다. 또한 시행착오(trial and error) 방식이 아닌 시스템의 분석을 통한 최적구조를 설계할 수 있는 방법론(methodology)을 제공할 수 있다.

새로운 구조인 모듈라 웨이블렛 신경망과 최적구조 설계 알고리즘을 2가지 다른 종류의 함수근사화 문제에 적용하여 모의실험을 수행하였고, 모의실험 결과에서 근사화 수행능력이 우수함을 검증할 수 있었다. 또한 일반 다층 신경망과 수행능력을 비교하여 모듈라 웨이블렛 신경망의 우수성을 확인하였다.

참 고 문 헌

- [1] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company Inc., 1994.
- [2] Y. C. Pati and P. S. Krishnaprasad, "Analysis and synthesis of feedforward neural network using discrete affine wavelet transformations," *IEEE Trans. on Neural Network*, vol. 4, pp. 73-85, Jan. 1993.
- [3] Ingrid Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trnas. on Information Theory*, Vol. 36, No. 5, pp. 961-1005, 1990.
- [4] Tomas Hrycej, *Modular Learning in Neural Network*, John Wiley & Sons Inc., 1992.
- [5] Gasser Auda and Mohamed Kamel, "Modular Neural Network Classifiers: A Comparative Study," *Journal of Intelligent and Robotic Systems*, Vol. 21, pp. 117-129, 1998.
- [6] G. Auda and M. Kamel, "CMNN: Cooperative Modular Neural Networks for pattern recognition," *Pattern Recognition Letters*, V.18 pp. 11-13, 1997.
- [7] Eric Ronco, Henrik Gollee and Peter J. Gawthrop, "Modular Neural Network and Self-Decomposition," *Centre for Systems and Control*, Technical Report CSC-96012, 1996.
- [8] Ke Chen, Liping Yang, Xiang Yu and Huisheng Chi, "A self-generating modular neural network architecture for supervised learning," *Neurocomputing*, Vol. 16, pp. 33-48, 1997.
- [9] R. A. Jacobs, M. I. Jordan, s. J. Nowlan, and G. E. Hinton, "Adaptive Mixtures of Local Experts," *Neural Computation*, Vol. 6, pp.181-214, 1994.
- [10] Michael I. Jordan and Robert A. Jacobs, "Hierarchical Mixtures of Experts and the EM Algorithm," *Neural Computation*, Vol. 6, No. 1, pp. 181-214, 1994.
- [11] Jacobs R. A. and Jordan M. I., "Learning Piecewise Control Strategies in a Modular Neural Network Architecture," *IEEE Trans. Sys. Man and Cybernetics*, Vol. 23, No. 2, pp. 337-345, 1993.
- [12] M. N. Dailey and G. W. Cottrell, "Organization of face and object recognition in modular neural network models," *Neural Networks*, Vol. 12, pp. 1053-1073, 1999.
- [13] Ath. Keagias and Vas. Petridis, "Predictive Modular Neural Networks for Time Series Classification," *Neural Networks*, Vol. 10, No. 1, pp. 31-49, 1997.
- [14] Wassim S. Chaer, Robert H. Bishop and Joydeep Ghish, "A Mixture-of-Experts Framework for Adaptive Kalman Filtering," *IEEE Trans. on Systems, Man, and*

- Cybernetics-Part B : Cybernetics*, Vol. 27, No. 3, pp. 452-464, 1997.
- [15] Andreas S. Weigend, Morgan Mangeas, and Ashok N. Srivastava, "Nonlinear Gated Experts for Time Series: Discovering Regimes and Avoiding Overfitting," *International Journal of Neural Systems*, Vol. 6., pp.373-399, 1995.
- [16] Masumi Ishikawa, "Learning of modular structured networks," *Artificial Intelligence*, Vol. 75, pp. 51-62, 1995.
- [17] Bao-Liang Lu and Masami Ito, "Task Decomposition and Module Combination Based on Class Relations : A Modular Neural Network for Pattern Classification," *IEEE Trans. on Neural Networks*, Vol. 10, No. 5, pp. 1244-1256, 1999.
- [18] Viswanath Ramamuri and Joydeep Ghosh, "Structurally Adaptive Modular Networks for Nonstationary Environments," *IEEE Trans. on Neural Networks*, Vol. 10, No. 1, pp.152-160, 1999.

저 자 소 개

徐 在 龍(正會員) 第37卷 SC編 第4號 參照
 1969年 7月 21日生. 1996년 2월 중앙대학교 전자공학과 졸업(학사). 1998년 2월 중앙대학교 대학원 전자공학과 졸업(석사). 1998년 8월~현재 : 중앙대학교 대학원 전자공학과 박사과정 재학중. 주관심분야 : 퍼지, 유전알고리즘, 신경망, FNN 등임

全 洪 兌(正會員)
 현재 : 중앙대학교 전자공학과 교수

金 鏞 宅(正會員)
 정회원, 1967年 2月 23日生. 1992년 2월 중앙대학교 전자공학과 졸업(학사). 1994년 8월 중앙대학교 대학원 전자공학과 졸업(석사). 1996년 9월~현재 : 중앙대학교 대학원 전자공학과 박사과정 재학중. 주관심분야 : 신경망, 로보틱스, 퍼지 등임

趙 鉉 讚(正會員)
 현재 : 한국기술교육대학교 전자공학과 교수