

論文2002-39CI-6-5

직교 기저함수 기반의 혼합 신경회로망 구조

(Structure of the Mixed Neural Networks Based On Orthogonal Basis Functions)

金成柱*, 徐在龍**, 趙鉉讚**, 金成炫***, 全洪兌*

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

요약

웨이블릿 함수의 경우 스케일링 함수에서 비롯되었으며, 스케일과 중심을 결정함으로써 신경회로망의 노드로 구성된다. 본 논문에서는 웨이블릿 함수를 이용하여 망을 구성하는 과정에 스케일링 함수를 은닉층의 노드로 복합 구성한 구조를 제안하고자 한다. 제안한 구조의 특징은 스케일링 함수를 이용하여 대강 근사(rough approximation)를 행한 다음, 웨이블릿 함수를 이용하여 미세 근사(fine approximation)를 행하도록 신경회로망의 은닉층을 복합 구성하는 데 있다. 또한, 복합 신경회로망을 구성하는 과정에서 미세 근사에 필요한 웨이블릿 함수의 개수를 유전 알고리즘을 이용하여 결정하는 초기 구조의 최적화를 도모하고자 한다.

Abstract

The wavelet functions are originated from scaling functions and can be used as activation function in the hidden node of the network by deciding two parameters such as scale and center. In this paper, we would like to propose the mixed structure. When we compose the WNN using wavelet functions, we propose to set a single scale function as a node function together. The properties of the proposed structure is that while one scale function approximates the target function roughly, the other wavelet functions approximate it finely. During the determination of the parameters, the wavelet functions can be determined by the global search algorithm such as genetic algorithm to be suitable for the suggested problem. Finally, we use the back-propagation algorithm in the learning of the weights.

Keywords: Mixed Neural Networks, wavelet function, Scaling function, Orthogonal Basis functions.

* 正會員, 中央大學校 電子電氣工學部

(School of Electrical and Electronics Eng. Chung-Ang Univ.)

** 正會員, 韓國技術教育大學校 情報技術工學部

(School of Info. Tech. Eng. Korea Univ. of Tech. and Edu.)

*** 正會員, 東元大學 電子科

(Dept. of Electronics, Tongwon College)

※ 본 연구는 과학기술부의 뇌과학연구 프로그램의 지원하에 수행되었습니다.

接受日字:2001年11月23日, 수정완료일:2002年9月17日

1. 서 론

신경회로망(Neural Network)은 적응 능력과 빠른 병렬 처리 능력을 특징으로 한다^[1]. 그렇지만, 종종 바람직하지 못한 지역적 최소치로 수렴하기도 하고 때로는 대단히 느린 수렴 속도를 보이기도 한다. 이러한 다층 구조 신경회로망의 대안으로 방사 기준 함수(RBF - Radial Basis Function)가 등장하게 된다. RBF 회로망은 학습 속도 측면에서 다층 전방향 신경회로망보다 훨씬 빠른 성능을 보인다^[2]. RBF 회로망은 지역적 변화와 불연속을 갖는 함수를 학습하는 데 적합하다. 그렇지만 RBF 회로망은 일반적으로 기준 함수들이 직교하지 않는다는 이유 때문에 필요 이상의 기준 함수들을 사용하게 되고 이것은 곧 효율적이고 고유한 망 구성이 이루어지지 않을 수 있음을 의미한다. 이러한 문제들을 극복하기 위해 기존의 신경회로망과 웨이블릿 이론을 결합한 웨이블릿 신경회로망(Wavelet Neural Network)이 제안되었다^[3]. 이와 같이 구성된 웨이블릿 신경회로망은 RBF 신경회로망의 대부분의 장점을 유지하고 알려지지 않은 함수에 대하여 효율적이고 고유한 망 구성을 제공한다. 그러나 웨이블릿 함수들은 강한 직교 조건을 만족해야 하며 이러한 직교 조건은 망의 유연성을 저해하는 요소로 작용하고, 또한 시간-스케일의 지역화에는 낮은 성능을 나타낸다^[3-4].

함수 근사의 문제에 있어서 근사 대상 함수를 포함하는 전체 영역은 스케일링 함수와 웨이블릿 함수로 표현됨을 알 수 있다. 또한 이 과정에서 웨이블릿 함수의 파라미터를 적절히 구성할 때, 더욱 효율적으로 구성할 수 있다. 이에 본 논문에서는 일반적인 웨이블릿 구조의 은닉층(Hidden Layer)에 적절히 선택된 하나의 스케일 함수를 이용하여 대상 함수를 대강 근사(rough approximation)하며 몇 개의 웨이블릿 함수를 이용하여 대상 함수를 미세 근사(fine approximation)할 수 있는 스케일링 함수와 웨이블릿 함수로 구성된 복합 구조를 제안하며 각 함수의 파라미터를 적절히 구성하고 학습시키는 최적 구조 설계 알고리즘을 제안한다.

제안한 스케일링-웨이블릿 복합 신경회로망의 최적 구조 설계 알고리즘은 다음과 같은 세 과정으로 구성되어 있다.

첫 번째 과정은 대상 함수를 최소 분해할 수 있는 스케일링 함수를 선택하며 파라미터를 결정한다. 두 번

째 과정은 최적화 알고리즘으로 널리 알려진 유전 알고리즘을 사용하여 웨이블릿 신경회로망의 기준 함수들을 최적으로 구성한다. 세 번째 과정은 문제 해결에 적합한 최적의 기준 함수를 이용하여 스케일링-웨이블릿 복합 신경회로망을 구성하고 기존의 역전파 학습으로 은닉층과 출력층 사이의 가중치를 조절한다.

제안한 스케일링-웨이블릿 복합 신경회로망의 최적 구조 설계 알고리즘을 함수 근사화 문제에 적용하여 그 결과를 기존의 다층 전방향 신경회로망과 직교 웨이블릿 신경회로망의 결과와 비교하여 우수성을 검증하고자 한다.

II. 스케일링 함수, 웨이블릿 함수 그리고 다중 분해 분석

일반적으로 근사 대상 함수에 대한 분석에 있어 단일한 분해도를 적용하여 분석하는 경우보다 다중 분해 분석을 이용하여 분석할 경우, 낮은 분해도의 분석과 높은 분해도의 분석이 가능하여 대상 함수의 특성을 다각적으로 분석할 수 있음을 알 수 있다.

다중분해분석(Multi-Resolution Analysis : MRA)을 이해하기 위하여, 먼저 스케일링(scaling) 함수를 정의하고, 이로부터 웨이블릿(wavelet) 함수를 정의하기로 한다. 이차원 스케일링 함수의 집합은 다음과 같은 정수 팽창(dilation) 변수와 이동(translation) 변수에 의하여 정의된다.

$$\varphi_{m,n}(x) = \varphi(2^m x - n) \quad (1)$$

여기서 $\varphi(x) \in L^2(\mathbb{R})$ 는 유한한 에너지를 갖는 함수, m, n 은 정수이며 이동에 의하여 서로 직교한다. 이 함수에 의하여 표현할 수 있는 $L^2(\mathbb{R})$ 의 부분공간은 다음과 같이 정의하면

$$V_m = \overline{\text{Span}_n\{\varphi_{m,n}(x)\}} \quad (2)$$

이는 곧 임의의 함수 $f(x) \in V_m$ 을 다음과 같이 나타낼 수 있음을 의미한다.

$$f(x) = \sum_n a_n \varphi(2^m x - n) \quad (3)$$

$m > 0$ 일 때, m 이 커질수록 $\varphi_{m,n}(x)$ 의 폭은 좁아지고 작은 간격으로 이동되기 때문에 더욱 정밀한 표현이

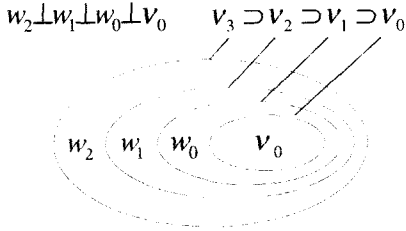


그림 1. 스케일링 함수와 웨이블릿 벡터 공간
Fig. 1. Scaling Function And Wavelet Vector Space.

가능하고 그 부분공간 또한 커진다.

$L^2(R)$ 공간의 부분공간 V_{m+1} 과 V_m 의 직교 웨이블릿 여공간 W_m 은 다음과 같이 정의된다.

$$V_m \cap W_m = 0 \tag{4}$$

$$V_m \oplus W_m = V_{m+1} \tag{5}$$

여기서, 부분공간 W_m 은 웨이블릿 함수 $\psi_{m,n}(x) = \psi(2^m x - n)$ 에 의하여 표현되는 공간이다. 이를 이용하면 $L^2(R)$ 공간은 다음과 같이 분해될 수 있다.

$$L^2 = V_0 \oplus W_0 \oplus W_1 \oplus \dots \tag{6}$$

$$= \dots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \dots$$

<그림 1>은 다른 스케일 m 에 대한 스케일링 함수 공간 V_m 과 웨이블릿 공간이 어떻게 나누어지는가를 보여준다.

$L^2(R)$ 공간의 임의의 함수 $f(x)$ 가 다음과 같이 두 가지로 분해됨을 알 수 있다.

$$f(x) = \sum_n \langle f, \phi_{m_0, n} \rangle \phi_{m_0, n}(x) + \sum_{m \geq m_0, n} \langle f, \psi_{m, n} \rangle \psi_{m, n}(x) \tag{7}$$

달리 표현하면,

$$f(x) = \sum_n c(n) \phi_{m_0, n}(x) + \sum_{m=m_0}^{\infty} \sum_n d_m(n) \psi_{m, n}(x) \tag{8}$$

여기서, $\langle \cdot, \cdot \rangle$ 는 내적을 표시하고, $c(n), d_m(n)$ 은 각각 스케일링 함수와 웨이블릿 함수의 적절한 계수이며, m_0 는 가장 낮은 분해능(resolution)을 표시하는 임의의 정수이다.

1. 웨이블릿 신경회로망

앞에서 살펴보았듯이 기존의 신경회로망의 각 뉴런

에서는 다음과 같은 입력-출력 관계를 갖는다.

$$S_{out} = \alpha \sum_m W_m S_{in, m} + T \tag{9}$$

여기서, $S_{in, m}$ 은 m 번째 입력, W_m 은 가중치(weight), T 는 임계값(threshold) 그리고 $\sigma(\cdot)$ 는 활성화 함수로 일반적으로 시그모이드 함수와 같은 비선형 함수이다. 활성화 함수를 웨이블릿 함수로 대체하고, 가중치 및 임계값을 각각 웨이블릿 함수의 팽창과 이동으로 생각한다면 다음과 같이 웨이블릿 신경회로망의 출력식을 구성할 수 있다.

$$f(x) = \sum_i w_i \phi_a(a_i x - t_i) \tag{10}$$

여기서, $x \in R^d$ 는 각 은닉층 뉴런의 입력 벡터이고, $a_i \in R^d, t_i \in R^d$ 는 각각 i 번째 은닉층 뉴런의 팽창과 이동 벡터이며, w_i 는 가중치를 나타낸다. 웨이블릿 이론의 함수 분해 이론으로부터, 식(10)의 팽창과 이동 변수가 적절하게 선택되고 가중치가 학습된다면 임의의 함수 $f(x) \in L^2(R)$ 이 웨이블릿 신경회로망으로 근사됨을 알 수 있다.

2. 다중 분해 분석

<그림 2>와 같은 함수를 단일 분해도에서 표현하고자 한다면, 급한 경사 구간의 간격에 의해 결정되는 분해도(h)를 사용하여 표현하여야 하며, 그만큼 많은 데이터 포인트들이 필요하다($h=1$ 의 경우). 만약 다중 분해도를 사용하여 이 함수를 표현한다면, 정확한 표현에 필요한 데이터 포인트들을 줄일 수 있다.

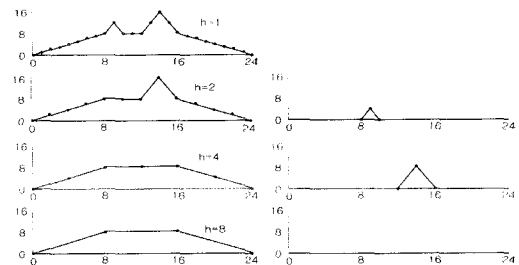


그림 2. 함수의 다중 분해도 분석
Fig. 2. Multiresolution Analysis of a Function.

III. 스케일링-웨이블릿 복합 신경회로망 설계

1. 스케일링-웨이블릿 복합 신경회로망

일반적으로 구성된 웨이블릿 신경회로망의 구조에서 다중 분해 분석을 통해 스케일링 함수를 추가로 구성하여 스케일링 함수와 웨이블릿 함수로 구성된 복합 구조를 이끌어낸다.

이 복합 구조는 다음 그림과 같다.

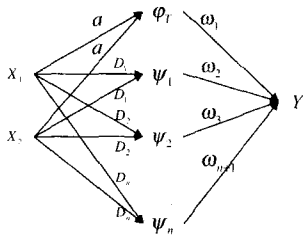


그림 3. 스케일링-웨이블릿 복합 신경회로망의 구조
Fig. 3. Structure of Complex Scaling-Wavelet Neural Network.

위 그림에서 D_1, D_2, \dots, D_n 은 각각 1, 2, ..., n 번째 웨이블릿의 신축 계수이며, ϕ_T 는 이동 계수를 T 로 하는 스케일 함수이며 자체적으로 스케일 변수, a 를 포함하고 있으며, ψ_n 는 웨이블릿 함수를 표현한다. 작은 입력층과 출력층 사이의 가중치는 $\omega_1, \omega_2, \dots, \omega_{n+1}$ 로 표현되며 여기서 n 은 웨이블릿만으로 이루어진 은닉층의 개수이다.

제안한 구조를 결정함에 있어서 가장 중요한 문제는 스케일링 함수를 선택하고 변수를 결정하는 일이다. 결정된 스케일링 함수의 스케일 변수와 이동 변수는 대상 함수와의 단순 비교를 통하여 대강 근사를 이룰 수 있는 최소한의 값을 부여한다. 스케일링 함수를 결정하면 전체 구조를 완성하기 위해 웨이블릿 함수의 변수를 결정해야 한다. 이는 본 논문에서 제안한 최적화 알고리즘인 유전 알고리즘을 이용하여 결정하고자 한다.

<그림 3>에서 제안한 구조의 전체 출력식을 표현하면 다음과 같다.

$$Y = \sum_{T=1}^n \phi_T(X_i(t)) W_T \sum_{n=1}^n \sum_{m=1}^n \psi_n(X_i(t)) W_{n+1} \quad (11)$$

여기서, i 는 입력의 개수, T 는 스케일 변수값, n 는 웨이블릿으로 구성된 은닉층 개수이며, W_n 는 은닉층과 출력층 사이의 가중치이다. 스케일 변수는 앞에서 설명한 것처럼 근사 대상 함수를 효과적으로 근사할 수 있는 최소값을 선택하며, 웨이블릿 함수는 신축과 이동 변수를 파라미터로 갖는다. 신축과 이동 변수는 제안한 것처럼 유전 알고리즘을 이용한 최적화 기법을 통해 출력층에서 에러(error)를 최소화할 수 있는 적절한 값을 선택하게 되며, 선택된 신축과 이동 변수의 값을 쌍으로 하는 웨이블릿 함수가 결정된다. 가중치는 일반적으로 사용되는 역전파 알고리즘을 통하여 갱신된다.

IV. 유전 알고리즘을 이용한 스케일링-웨이블릿 복합 신경회로망의 최적 구조 설계

분해도는 스케일링 함수의 스케일 조절 변수, 웨이블릿의 팽창 변수에 해당하고 데이터 포인트들은 스케일링 함수의 최초 이동 변수에 의한 기저함수가 되는 동시에, 이동 변수에 의하여 이동되는 웨이블릿 기저함수에 해당되기 때문에 효율적인 함수 근사의 문제는 결국 낮은 수준 분해도에 따른 스케일링 함수의 최초 스케일 조절 변수와 이동 변수에 좌우되며 또한, 웨이블릿 이론의 팽창과 이동의 결정에 좌우된다고 할 수 있다. 따라서 본 연구에서는 대상 함수를 가장 낮은 수준으로 분해하였을 경우에 대상 함수를 적절히 표현할 수 있는 스케일링 함수를 결정하고 유전 알고리즘을 사용하여 근사 대상 함수에 대한 식(11)의 웨이블릿 신경회로망의 팽창과 이동의 최적 값을 찾는다.

이제 좀더 자세히 제안한 최적구조 설계 알고리즘의 처리 과정에 대하여 알아본다.

[단계 1] 스케일링 함수 변수 결정

스케일링 함수의 초기 변수들을 결정한다.

[단계 2] 각 개체의 초기화

유전 알고리즘의 초기 개체군을 이루는 각 개체의 초기화는 다음 그림과 같이 웨이블릿 함수 결정에 필

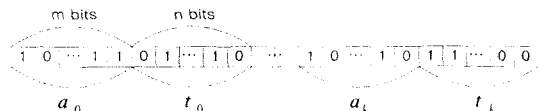


그림 4. 각 개체의 초기화
Fig. 4. Initialization of a Chromosome.

요한 팽창과 이동의 정보를 가질 수 있도록 코딩한다.

팽창 변수(a_i)와 이동 변수(t_i)에 각각 m, n 비트(bit)를 할당한다.

[단계 3] 망 고정 및 적합도 평가

각 개체의 이진 정보를 적절한 팽창(a_i), 이동(t_i)으로 변환시킨 다음 각 개체마다의 적합도 값을 계산한다. 스케일링 변수와 팽창과 이동으로 변환된 적합도 값으로 스케일링-웨이블릿 복합 신경회로망의 입력층과 은닉층 사이의 가중치 및 각 뉴런의 임계값을 고정시킨 후 은닉층과 출력층 사이의 가중치 값을 역전파 학습 알고리즘을 이용하여 구한다.

[단계 4] 새로운 집단 생성

본 논문에서 개체 선택에 사용된 방법은 룰렛 선택법과 엘리트 보존 선택법이며, 교배와 돌연변이는 각각 단순 교배와 다점 돌연변이가 사용되었다.

[단계 5] 새로운 최고의 개체나 원하는 적합도 값이 얻어질 때까지 상기 단계를 반복한다.

[단계 6] 최종적으로 얻어진 개체의 팽창과 이동 값과 스케일링 변수를 이용하여 스케일링-웨이블릿 복합 신경회로망을 고정시켜 최종 구조를 결정한 후 역전파 학습 알고리즘을 이용하여 학습을 수행한다.

V. 임의의 비선형 함수에 대한 근사화 모의 실험 결과

1차원 함수와 2차원 함수에 대하여 근사화 실험을 수행하였으며 비교 대상으로는 기존의 다층 전방향 신경 회로망과 단일 팽창과 균일한 이동을 사용하는 직교 기저 스케일링 함수를 사용한 웨이블릿 신경회로망을 이용하였다.

1. 1차원 비선형 함수에 대한 근사화 실험

스케일링-웨이블릿 복합 신경회로망에 사용된 스케일링 함수는 싱크 함수로 구성된 싱크 스케일링(Sinc Scaling) 함수인 $\phi(t) = \frac{\sin(\pi t)}{\pi t}$ 를 사용했고, 웨이블릿 함수는 싱크 스케일링 함수에서 유도된 싱크 웨이블릿 함수 $\psi(t) = \frac{2\sin(\pi t) - \sin(2\pi t)}{\pi t}$ 이다. 근사 대상 함수

$f(x)$ 는 구간 $[-10, -2)$ 에서 $-2.186x - 12.864$, $[-2, 0)$ 에서 $4.246x$, $[0, 10]$ 에서 $10e^{-0.05x - 0.5 \cdot \sin[(0.03x + 0.7)x]}$ 와 같이 구간별로 정의된 함수이다.

유전 알고리즘에 사용된 제어 파라미터들은 개체군

표 1. 결정된 팽창과 이동

Table. 1. Decided Dilations and Translations.

뉴런	1	2	3	4	5	6	7
팽창	2^{-2}	2^{-2}	2^{-1}	2^0	2^{-2}	2^{-1}	2^{-1}
이동	0.0	-4.0	0.0	7.0	3.0	2.0	3.0

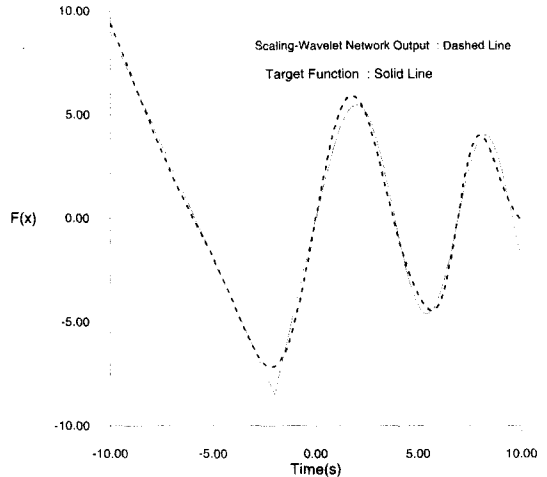


그림 5. 본 논문에서 제안한 스케일링-웨이블릿 복합 신경회로망의 근사 결과

Fig. 5. Approximation of Complex Scaling-Wavelet Neural Network.

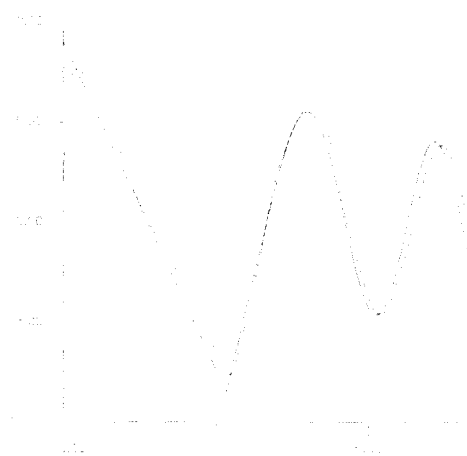


그림 6. 다차원 전방향 신경회로망의 근사 결과

Fig. 6. Approximation of Multilayered Feedforward Neural Network.

크기 60, 교배율 100%, 돌연변이율 30%와 같으며 스케일링으로 구성된 한 개의 뉴런과 <표 1>과 같이 구성된 웨이블릿 7개, 총 8개의 뉴런을 사용하였다.

역전파 학습의 학습률 η 는 0.05를 사용하였고, 총

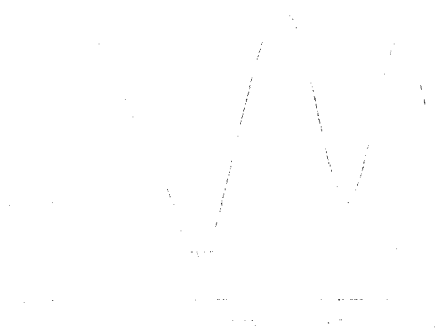


그림 7. 직교기저 웨이블릿 신경회로망의 근사 결과
Fig. 7. Approximation of Orthogonal Wavelet Neural Network.

2000번의 반복 학습을 수행하였다.

직교 기저 웨이블릿 신경 회로망의 웨이블릿 함수로는 직교 기저 스케일링 함수로 널리 알려진 식(12)와 같은 싱크 함수를 사용하였다.

$$\varphi(x) = \frac{\sin x}{x} \quad (12)$$

여기서, 단일 팽창은 2^{-1} 을 사용하였고, 근사 대상 함수의 영역이 완전히 포함될 수 있도록 총 11개의 뉴런을 사용하였다. <그림 5~7>은 각 신경회로망의 근사 결과를 나타낸 것이다. 결과에서 확인할 수 있듯이 유전 알고리즘에 의하여 구성된 스케일링-웨이블릿 복합 신경회로망이 적은 수의 뉴런으로도 더 좋은 근사 결과를 나타냄을 알 수 있다.

VI. 결 론

다중 분해 분석을 통해 낮은 수준 분해도의 분석을 통한 대강 근사와 높은 분해도의 분석을 통해 미세 근사를 시도하기 위해 본 논문에서는 스케일링-웨이블릿 복합 신경회로망의 팽창과 이동을 최적으로 구성하고

학습시키는 알고리즘을 제안하였다. 이 알고리즘은 유전 알고리즘을 이용하여 주어진 문제 해결에 적합한 팽창과 이동 값을 결정한 후에, 다시 역전파 학습 알고리즘을 이용하여 은닉층과 출력층 사이의 가중치를 학습하는 것에 그 특징이 있으며, 이와 같이 구성된 스케일링-웨이블릿 복합 신경회로망이 임의의 비선형 함수에 대한 근사화 실험 결과에서 기존의 다차원 전방향 신경회로망이나 직교 기저 웨이블릿 신경회로망보다 더 적은 뉴런수로도 더 좋은 근사 결과를 나타냄을 확인하였다. 또한 스케일링 함수를 이용하는 경우 웨이블릿을 대체하는 효과를 볼 수 있는 데 수학적 증명을 통한 스케일링 함수와 웨이블릿과의 연관 관계를 해석해 보는 연구가 필요하다.

참 고 문 헌

- [1] James A. Freeman and David M. Skapura, *Neural Networks Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing, 1991.
- [2] V. T. Sunil Elanayar and Y. C. Shin, "Radial Basis Function neural network for approximation and estimation of nonlinear dynamic systems," *IEEE Trans. Neural Networks*, vol. 4, pp. 594~603, 1994.
- [3] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 889~898, Nov. 1992.
- [4] Jaideva C. Goswami and Andrew K. Chan, *Fundamentals of Wavelets*, Wiley Interscience, 1999.
- [5] 서재용, 김용택, 조현찬, 전홍태, "시간-주파수 분석을 이용한 모듈라 웨이블릿 신경망의 최적 구조 설계," *대한전자공학회 논문지*, 제38권, SC편, 제2호, pp. 12~19, 2001

저 자 소 개

金成柱(正會員) 第37卷 SC編 第5號 참조

金成炫(正會員)

현재 : 동원대학 전자과 교수

徐在龍(正會員) 第37卷 SC編 第4號 참조

全洪兌(正會員)

현재 : 중앙대학교 전기전자공학부 교수

趙鉉讚(正會員)

현재 : 한국기술교육대학교 정보기술공학부 교수