한국 퍼지 및 지능시스템학회 논문지 2000, Vol. 10, No. 4 (2000)338~342

유전 알고리즘을 이용한 웨이블릿 신경회로망의 최적 구조 설계

Optimal Structure of Wavelet Neural Network Systems Using Genetic Algorithm

이창민 · 서재용 · 전홍태

Chang-Min Lee, Jae-Yong Seo and Hong-Tae Jeon

중앙대학교 전기전자공학부

ABSTRACT

In order to approximate a nonlinear function, wavelet neural networks combining wavelet theory and neural networks have been proposed as an alternative to conventional multi-layered neural networks. Wavelet neural networks provide better approximating performance than conventional neural networks. In this paper, an effective method to construct an optimal wavelet neural network is proposed using genetic algorithm. Genetic Algorithm is used to determine dilations and translations of wavelet basis functions of wavelet neural networks. Then, with these determined dilations and translations, wavelet neural networks are further trained by back propagation learning algorithm. The effectiveness of the final network is verified through the approximation result of a nonlinear function and comparison with conventional neural networks.

1. 서 론

최근에 신경회로망(Neural Network)은 다소 복잡한 형태의 함수를 학습할 수 있는 능력 때문에 비선형 시스템의 제어에 인기있는 도구가 되었다. 그 중에서도 역전파 학습 알고리즘(Back Propagation Learning)을 이용한 다층 전방향(Multi-layered Feedforward) 신경회로망은 실제 응용에 있어서 자주 사용되는 신경회로망의 형태이다. 그러나, 다층화된 구조와 역전파 알고리즘의 성질 때문에 학습과정은 종종 바람직스럽지 못한 지역 최소치에 빠지거나 수렴 속도가 너무 느린 단점이 있다. 이런 다층 전방향 신경회로망의 단점을 극복하기 위해 제안된 RBF(Radial Basis Function) 회로망은 더욱 간단한 구조(1개의 은닉층 사용)이기 때문에 학습속도가 다층 전방향 신경회로망보다 훨씬 빠르다[1].

함수 근사의 측면에서 볼 때, RBF 회로망은 지역적 특성을 갖는 기저함수(basis function)를 사용하여 함수를 근사화하는 형태이다. 이런 기저함수의 지역성으로 인하여 RBF 회로망은 지역적 변화와 불연속을 갖는 함수를 학습하는데 더욱 더 적합하다. 이러한 장점에도 불구하고 RBF 신경회로망은 기저함수들이 일반적으로 직교하지 않기 때문에 필요 이상의 기저함수들을 사용하게 되고 이것은 효율적이고 고유한 망구성의 방해 요인이 된다. 이러한 문제들을 극복하기 위해 기존의 신경회로망과 웨이블릿 이론을 결합한 웨이블릿 신경회로망(Wavelet Neural Network, WNN)이 제안되었다[2,3]. 웨이블릿 신경회로망은 RBF 회로망의 기저함수들을 웨이블릿 이론(wavelet theory)의 웨이블릿 함수(waveletfunction)로 대체하여 직교성을 갖는 기저함수를 갖도록 한다. 이와같이 구성된 웨이블릿 신경회로망은 RBF 신경회로망의 대부분의 장점을 유지하고, 알려지지 않은 함수에 대하여 효율적이고 고유한 망구성을 제공한다. 그러나 웨이블릿 함수들은 강한 직교 조건을 만족해야하며 이러한 직교 조건은 망의 유연성을 저해하는 요소로 작용하고, 또한 시간-스케일의 지역화에는 낮은 성능을 나타낸다[2,4].

이에 본 논문에서는 직교성을 갖는 웨이블릿 함수 대신에 망에 유연성을 제공하고 직교 웨이블릿 함수와 비슷한 알고리즘을 제공하는 웨이블릿 프레임(Frame) 함수로 웨이블릿 신경회로망을 구성한 뒤, 망의 각 파라미터들을 적절히 구성하고 학습시키는 최적 구조 설계 알고리즘을 제안한다. 제안한 웨이블릿 신경회로망의 최적 구조 설계 알고리즘은 다음과 같은 두 과정으로 구성되어 있다. 첫번째 과정은 최적화 알고리즘으로 널리 알려진 유전 알고리즘을 사용하여 웨이블릿 신경회로망의 기저함수들을 최적으로 구성한다. 두번째 과정은 문제 해결에 적합한 최적의 기저함수를 이용하여 웨이블릿 신경회로망을 구성하고 기존의 역전파 학습으로 은닉층과 출력층 사이의 가중치를 조절한다. 제안한 웨이블릿 신경회로망의 최적 구조 설계 알고리즘을 함수 근사화 문제에 적용하여

본 연구는 과학기술부의 뇌과학연구 프로그램의 지원 하에 수행되었습니다.

그 결과를 기존의 다층 전방향 신경회로망과 직교 웨이블릿 신경회로망의 결과와 비교하여 우수성을 검증한다.

본 논문은 다음과 같이 구성되어 있다. 2절, 3절에서 웨이블릿 이론과 웨이블릿 신경회로망, 유전 알고리즘에 관하여 알아본 후, 4절에서는 유전 알고리즘을 이용한 웨이블릿 신경회로망의 구성 방법에 대하여 알아본다. 5절에서 임의의 비선형 함수에 대한 근사 실험 결과를 기술한다.

2. 웨이블릿 이론과 웨이블릿 신경회로망

웨이블릿 이론에 의하면, 다음과 같이 $L^2(R)$ 공간을 직교 분해할 수 있는 웨이블릿 함수 $\psi(x)$가 존재한다 [4-6].

$$L^2(R) = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots \tag{1}$$

$$\psi_{m,n}(x) = 2^{m/2}\psi(2^m x - n) \tag{2}$$

여기서 W_m은 $[2^{m/2}\psi(2^m x - n)]_{n=-\infty}^{n=+\infty}$으로 표시할 수 있는 부분공간이며, m, n은 정수이다. 웨이블릿 $\psi(x)$는 스케일링(scaling) 함수라 알려진 $\varphi(x)$로부터 구할 수 있으며 스케일링 함수의 팽창항(dilation)과 이동항(translation)은 $L^2(R)$의 다중 분해 분석(multiresoulution analysis, MRA)을 가능케하고 다음과 같은 관계를 성립케한다[4-6].

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L_2 \tag{3}$$

여기서, Vm은 $[2^{m/2}\varphi(2^m x - n)]_{n=-\infty}^{n=+\infty}$으로 표시할 수 있는 부분공간이다. 또한, V_m은 W_m과 다음과 같은 직교 여공간의 관계에 있다.

$$V_{m+1} = V_m \oplus W_m \tag{4}$$

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

$$f(x) = \sum_n \langle f, \varphi_{m_0,n} \rangle \varphi_{m_0,n}(x) + \sum_{m \ge m_0, n} \langle f, \psi_{m,n} \rangle \psi_{m,n}(x) \tag{5}$$

또는 $f(x) = \sum_{m,n} \langle f, \psi_{m,n} \rangle \psi_{m,n}(x)$ (6)

여기서, $<\cdot,\cdot>$는 내적을 표시하고, m_0는 가장 낮은 분해능(resolution)을 표시하는 임의의 정수이다. 단일 차원에서의 이와 같은 결과는 다음과 같은 텐서(tensor) 곱 형태로 다차원으로 확장할 수 있다[2,3,7].

$$\psi_d(\mathbf{x}) = \prod_{j=1}^{d} \psi(x_j) \tag{7}$$

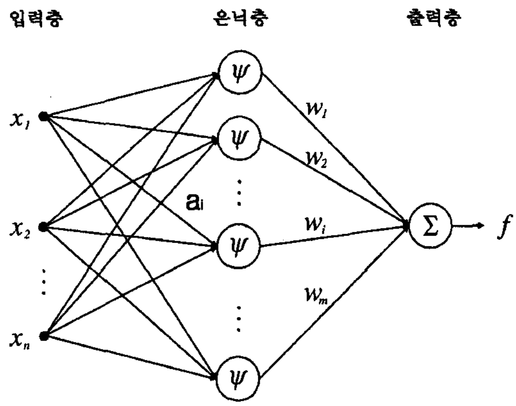


그림 1. 웨이블릿 신경회로망
Fig. 1. Wavelet Neural Network

상기의 논의로부터 다음과 같이 웨이블릿 신경회로망을 구성할 수 있다.

$$f(\mathbf{x}) = \sum_i w_i \psi_d(\mathbf{a}_i \mathbf{x} - \mathbf{t}_i) \tag{8}$$

여기서, $\mathbf{x} \in \mathbf{R}^d$는 각 은닉층 뉴런의 입력 벡터이고, $\mathbf{a}_i \in \mathbf{R}^d$, $\mathbf{t}_i \in \mathbf{R}^d$는 각각 i번째 은닉층 뉴런의 팽창과 이동 벡터이며, w_i는 가중치를 나타낸다. 한편, 직교 웨이블릿 함수는 앞서 서술한 바와 같이 실제 응용에 있어서 부적절한 면이 있으므로, 웨이블릿 프레임 함수를 사용한다. 웨이블릿 프레임 함수는 직교 웨이블릿 함수에 비하여, 잉여성분(redundancy factor)이 존재하는 것이 사실이나, 기저함수의 선택이 직교 웨이블릿 함수에 비하여 용이하며, 잉여성분 또한 경우에 따라서는 에러에 대한 강건성을 제공해 줄 수 있고, 함수 구성에 있어서 직교 웨이블릿 함수와 같은 알고리즘을 제공하기 때문에, 실제 응용에 있어서 많이 사용된다. 프레임 함수의 정의는 다음과 같다[6].

정의 1: Hilbert 공간 $H(L^2(R^d))$과 벡터 $\{h_k\}_{k=-\infty}^{\infty} \subset H$의 수열이 주어졌을 때, 다음 식을 만족하는 상수 $A > 0$고 $B < 0$이 존재하면 $\{h_k\}_{k=-\infty}^{\infty}$를 프레임(frame)이라 한다.

$$A\|f\|^2 \le \sum_k |\langle f, h_k \rangle|^2 \le B\|f\|^2 \tag{9}$$

여기서, $f \in H$이고, A와 B는 프레임 한계(frame bound)이다. 만약 $A = B = 1$이면 프레임 함수 집합은 직교 기저 함수 집합이 된다.

3. 유전 알고리즘

유전 알고리즘은 자연생태계의 진화 매카니즘인 '적

자 생존'의 원리를 모방하여, 최적화를 목적으로 만들어진 알고리즘이다. 유전 알고리즘에서는 임의의 세대(generation)를 형성하는 개체들의 집합, 즉 개체군(population) 중에서 환경에 대한 적합도(fitness)가 높은 개체가 높은 확률로 살아남아 재생할 수 있게 되며, 이 때 교배(crossover) 및 돌연변이(mutation)로서 다음 세대의 새로운 개체군을 형성하게 된다[9,10]. 본 논문에서 사용된 유전알고리즘의 구체적인 처리과정은 다음 절에서 설명한다.

4. 유전 알고리즘을 이용한 웨이블릿 신경회로망의 설계

일반적으로 웨이블릿 신경회로망의 구조는 입력층과 은닉층의 가중치에 해당하는 팽창과 뉴런의 임계값에 해당되는 이동에 의하여 결정되며 효율적인 함수 근사의 문제는 결국 웨이블릿 이론의 팽창과 이동의 결정에 좌우된다 할 수 있다. 따라서 본 연구에서는 유전 알고리즘을 사용하여 근사 대상함수에 대한 식(8)의 웨이블릿 신경회로망의 팽창과 이동의 최적값을 찾는다. 이를 위한 웨이블릿 신경회로망의 최적구조 설계 알고리즘의 흐름도는 다음과 같다.

[단계 1] 각 개체의 초기화

유전 알고리즘의 초기 개체군을 이루는 각 개체의 초기화는 다음 그림과 같이 개체마다 웨이블릿 신경회로망에 필요한 모든 팽창과 이동의 정보를 가질 수 있도록 코딩한다.

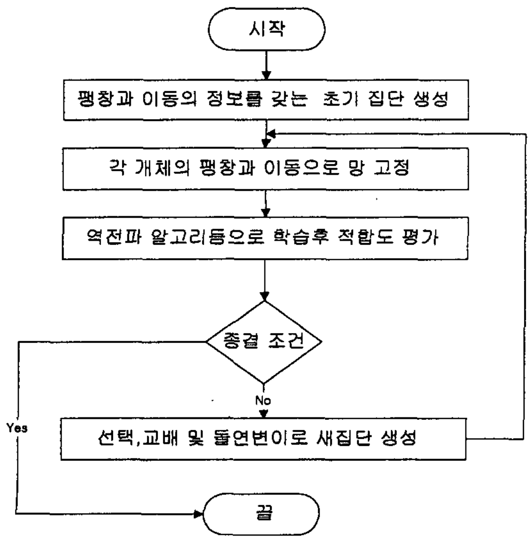


그림 2. 제안한 최적구조 설계 알고리즘의 흐름도
Fig. 2. Flow Chart of the proposed Algorithm

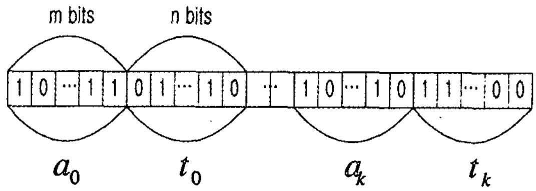


그림 3. 각개체의 초기화
Fig. 3. Initialization of a Chromosome

[단계 2] 망 고정 및 적합도 평가

각 개체의 이진 정보를 적절한 팽창(a_i), 이동(t_i)으로 변환시킨 다음 웨이블릿 신경회로망에 대입하여 역전파 학습을 수행한 후, 각 개체마다의 적합도 값을 계산한다. 여기서, 변환될 수 있는 팽창의 최소값은 단일 웨이블릿 함수를 사용하여도 근사 대상 함수의 영역을 충분히 포함할 수 있도록 설정하며, 최대값은 근사 대상 함수에 따라 적절히 선택하여 준다. 마찬가지로, 변환될 수 있는 이동의 최대값을 t_{max}라 하고, 최소값을 t_{min}이라 하면, 변환될 수 있는 이동(t_i)의 최대값과 최소값은 대상 함수의 영역[x_{min}, x_{max}]과 팽창(a_i)에 의하여 다음과 같이 결정된다.

$$t_{max} = x_{max} \times a_i,\ t_{min} = x_{min} \times a_i \tag{10}$$

위와 같은 억제 조건(constraint condition)은 변환된 이동에 따라 웨이블릿 함수가 이동하여 근사 대상 함수의 영역을 벗어나는 것을 방지해준다. 각 개체의 적합도 값은 변환된 팽창과 이동으로 웨이블릿 신경회로망을 고정시킨 후 은닉층과 출력층 사이의 가중치 값을 역전파 학습 알고리즘을 이용하여 구한다. 학습은 출력층 뉴런이 선형합 형태이고 입력과 은닉층 사이의 가중치는 학습할 필요가 없기 때문에 기존의 신경망의 학습보다는 훨씬 간단하지만 너무 많은 반복횟수는 많은 시간이 소요되므로 각 개체의 적합도 차이를 알 수 있는 범위에서 학습을 중단한다. 적합도는 다음 식에 의하여 구한다.

$$fitness = \frac{1}{\sqrt{\sum_{i=1}^{n} (f(x_i) - g(x_i))^2 / n + 1}} \tag{11}$$

여기서, $f(x_i)$와 $g(x_i)$는 각각 i번째 입력의 근사 대상함수와 웨이블릿 신경회로망의 출력이다.

[단계 3] 새로운 집단 생성

개체군에서 높은 적합도 값을 갖는 개체에 높은 선택 확률을 주고, 선택된 개체 사이에 교배와 돌연변이의 진화과정을 통하여 새로운 개체군을 얻는다. 본 논

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

[단계 4] 새로운 최고의 개체나 원하는 적합도 값이 얻어질 때까지 상기 단계를 반복한다.

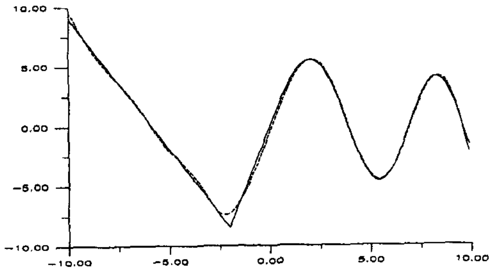
[단계 5] 최종적으로 얻어진 개체의 팽창과 이동으로 웨이블릿 신경회로망을 고정시켜 최적구조를 결정한 후 은닉층과 출력층 사이의 가중치 값을 역전파 학습 알고리즘을 이용하여 충분히 학습을 수행한 후 모의실험을 종료한다.

5. 모의 실험 결과

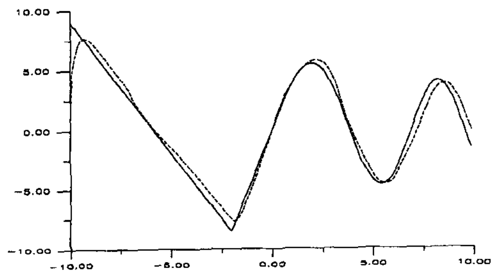
1) 1차원 함수: 웨이블릿 신경회로망에 사용된 웨이블릿 프레임 함수는 $y(x)=xe^{-\frac{1}{2}x^{2}}$이고 총 8개의 뉴런이 사용되었으며, 근사 대상 함수 $f(x)$는 구간 [-10, 10]에서 다음과 같이 구간별로 정의된 함수이다.

$$f(x)=\begin{cases} -2.186x-12.864 & [-10\leq x<-2] \\ 4.246x & [-2\leq x<0] \\ 10e^{-0.05x-0.5}\cdot\sin[(0.03x+0.7)x] & [0\leq x\leq 10] \end{cases} \qquad (12)$$

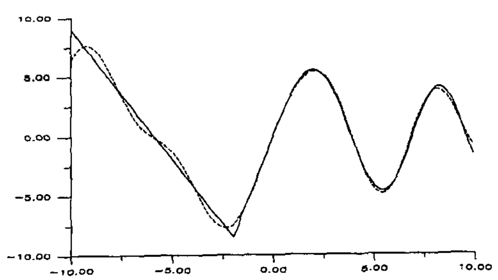
비교를 위하여 시그모이드 함수에 기반한 다층 전방향 신경회로망과 직교 기저 스케일링 함수를 사용한(팽창항과 이동항이 고정된) 웨이블릿 신경회로망의 결과도 함께 보였다. 비교대상인 다층 전방향 신경회로망에는 총 10개의 뉴런을 사용했고 역전파 학습의 학습률 η는 0.05를 사용하였으며, 총 2000번의 반복 학습을 수행하였다. 직교 기저 웨이블릿 신경회로망에서는 단일 팽창으로 2^{-1}을 사용하였고, 근사 대상 함



최적구조 설계 알고리즘을 적용한 웨이블릿 신경회로망의 근사 결과



다층 전방향 신경회로망의 근사 결과



직교기저 웨이블릿 신경회로망의 근사결과

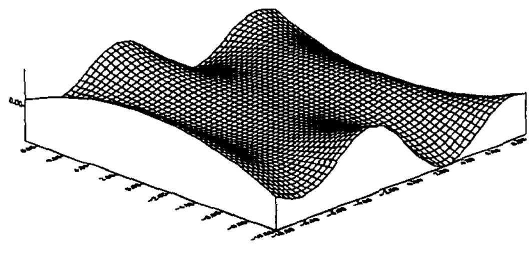
그림 4. 1차원함수의 근사 결과

Fig. 4. Approximation result of 1D function

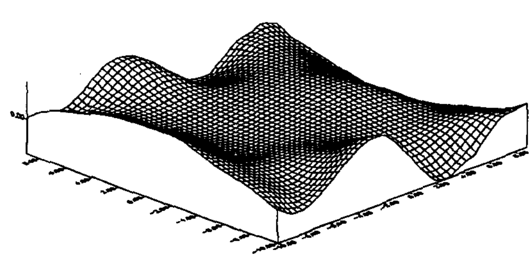
표 1. 최적구조 설계 알고리즘의 WNN과 직교기저 WNN과의 비교

Table 1. Comparison of WNN by the proposed Algorithm & Orthogonal WNN

	뉴런	RMSE
최적구조 설계 알고리즘의 웨이블릿 신경회로망	64	3.452
직교기저 웨이블릿 신경회로망	121	3.246



2차원 근사화 대상 함수



최적구조 설계 알고리즘을 적용한 웨이블릿 신경회로망의 근사 결과

그림 5. 2차원 함수의 근사 결과

Fig. 5. Approximation result of 2D function

수의 영역이 완전히 포함될 수 있도록 총 11개의 뉴런을 사용하였다. 결과에서 확인할 수 있듯이 유전 알고리즘에 의하여 구성된 웨이블릿 신경회로망이 적은 수의 뉴런으로도 더 좋은 근사 결과를 나타냄을 알 수 있다.

2) 2차원 함수의 근사 : 웨이블릿 프레임 함수로 1차 웨이블릿 프레임 함수의 텐서 곱 형태인 $\psi(x) = x_1 x_2 e^{-\frac{1}{2}(x_1^2 + x_2^2)}$이 사용되었으며, 근사대상 함수로 영역 $[-10, 10] \times [-10, 10]$사이에 있는 $f(x) = (x_1^2 - x_2^2)$가 사용되었다. 표 1에서 직교기저 웨이블릿 신경회로망과의 근사 결과를 비교하였다. 결과에서 확인할 수 있듯이, 최적구조 설계 알고리즘을 적용한 웨이블릿 신경 회로망이 절반 정도의 뉴런으로도 직교 기저 웨이블릿 신경회로망과 유사한 결과를 얻었음을 알 수 있다.

6. 결 론

웨이블릿 신경회로망에서 적합한 팽창과 이동 값의 선택은 더욱 효율적인 함수 근사를 위하여 매우 중요하다. 이에 본 논문에서는 웨이블릿 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 웨이블릿 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. 그러나, 일반적으로 웨이블릿 신경회로망은 입력 차원이 높아질수록 많은 수의 뉴런이 필요하며 이는 곧 유전 알고리즘이 탐색해야할 해공간이 늘어나게 되어 최적해 탐색에 많은 시간이 필요함과 동시에 최적해가 아닌 공간으로 수렴할 가능성 또한 늘어난다 할 수 있다. 따라서, 향후에는 이와 같은 문제점을 극복할 수 있도록 보다 발전된 형태의 유전 알고리즘이나 다른 탐색 알고리즘을 적용하여 해공간 탐색의 효율성 및 정확성을 높이는 방향으로의 연구가 필요하다.

참고문헌

[1] 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.

[2] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Networks,* Vol. 3, pp. 889-898, Nov. 1992.

[3] J. Zhang, G. G. Walter, Y. Miao and W. N. W. Lee, "Wavelet neural networks for function learning," *IEEE Trans. Signal Processing,* Vol. 43, pp. 1485-1497, June. 1995.

[4] Jaideva C. Goswami and Andrew K. Chan, *Fundamentals of Wavelets*, Wiley Interscience, 1999.

[5] Martin Vetterli and Jelena Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, 1995.

[6] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms.* Upper Saddle River, NT : Prentice-Hall, 1998.

[7] T. Kugarajah and Q. Zhang, "Multidimensional wavelet frames," *IEEE Trans. Neural Networks*, Vol. 72, pp. 1552-1556, Nov. 1995.

[8] B. Delyon, A. Juditsky, and A. Benveniste, "Accuracy analysis for wavelet approximations" *IEEE Trans. Neural Network*, Vol. 6, pp. 332-348, Mar. 1995.

[9] David E. Goldberg, Genetic Algorithms in Search, *Optimization, and Machine Learning,* Addison-Wesley Publishing, 1989.

[10] L. Davis, *Handbook of Genetic Algorithm*, Van Nostrand Reinhold, 1991

이 창 민 (Chang-Min Lee)

1998년 2월 : 중앙대학교 전자공학과 졸업(학사)
2000년 8월 : 중앙대학교 대학원 전자공학과 졸업 예정(석사)
주관심분야 : 퍼지, 웨이블릿 신경망, 유전 알고리즘 등임

서 재 용 (Jae-Yong Seo) 정회원

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

전 홍 태 (Hong-Tae Jeon) 정회원

현재: 중앙대학교 전기전자공학부 교수