

## 모듈화된 웨이블릿 신경망의 적응 구조 설계

# Adaptive Structure of Modular Wavelet Neural Network

서재용\*, 김성주\*\*, 조현찬\*, 전홍태\*\*

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

\*한국기술교육대학교 정보기술공학부

\*\*중앙대학교 전자전기공학부

### 요 약

본 논문에서는 F-투영과 기하학적인 성장 기준을 이용하여 모듈화된 웨이블릿 신경망(MWNN)의 적응구조를 설계하기 위한 성장과 소거 알고리즘을 제안한다. 기하학적인 성장 알고리즘은 지역 오차를 고려한 추정 에러 기준과 이미 존재하는 다른 웨이블릿 함수와 근사적인 직교성을 보장하도록 웨이블릿 함수를 배열할 수 있는 각도 기준으로 구성되어 있다. 이러한 기준은 망의 설계자가 설계자의 의도대로 MWNN을 구성할 수 있는 방법론을 제시한다. 제안한 성장 알고리즘은 모듈과 모듈의 크기를 성장시킨다. 또한 소거 알고리즘은 MWNN의 모듈에 사용되는 웨이블릿 신경망의 지역화 특성에 기인한 문제점을 극복하기 위해 구성된 MWNN에서 불필요한 모듈의 노드나 모듈을 제거한다. 제안한 MWNN의 적응 구조 설계 알고리즘을 1, 2차원 대상 함수에 적용하여 성능을 검증하였다.

### Abstract

In this paper, we propose an growing and pruning algorithm to design the adaptive structure of modular wavelet neural network(MWNN) with F-projection and geometric growing criterion. Geometric growing criterion consists of estimated error criterion considering local error and angle criterion which attempts to assign wavelet function that is nearly orthogonal to all other existing wavelet functions. These criteria provide a methodology that a network designer can constructs wavelet neural network according to one's intention. The proposed growing algorithm grows the module and the size of modules. Also, the pruning algorithm eliminates unnecessary node of module or module from constructed MWNN to overcome the problem due to localized characteristic of wavelet neural network which is used to modules of MWNN. We apply the proposed constructing algorithm of the adaptive structure of MWNN to approximation problems of 1-D function and 2-D function, and evaluate the effectiveness of the proposed algorithm.

**Key Words** : Modular Wavelet Neural Network, Growing and Pruning Algorithm, Adaptive Structure

## 1. 서 론

신경망에 대한 모델 선택과 망의 성장 및 소거 알고리즘에 대한 연구가 활발하게 진행되고 있다[1]. 또한 모듈화된 설계방식의 장점이 인식되면서 Jordan et al.이 제안한 모듈화된 신경망(Modular Neural Network; MNN)의 성장 및 소거 알고리즘에 대한 연구결과가 보고되고 있다[2][3]. 제안한 방식들의 대부분은 소거 알고리즘을 적용하기 위해 초기에 구성하는 MNN의 적합한 크기를 결정하는 문제와 다분히 시행착오적인 방법을 사용하였다.

따라서 본 논문에서는 시행착오적인 방법보다는 대상 문제에 적응적인 모듈화된 웨이블릿 신경망(MWNN)의 성장과 소거 알고리즘을 제안한다. 제안한 알고리즘은 F-투영 방법과 성장기준을 바탕으로 전역오차가 감소하도록 엑스퍼트 네트워크(EN)의 팽창과 이동계수를 결정

하여 노드를 추가한다[4]. 전역오차가 원하는 문턱치에 도달하여 성장 알고리즘이 종료되면 전역오차의 감소에 기여하지 못하는 각 EN의 노드를 제거하는 소거 알고리즘을 수행하여 최적의 MWNN의 구조를 얻는다.

MWNN의 게이팅 네트워크(GN)은 비선형성이 강한 다층신경망으로 구성하고, 학습 알고리즘은 기울기 학습 방법과 Expectation-Maximization(EM) 알고리즘을 결합하여 사용한다. EM 알고리즘은 수렴 속도가 우수하다는 장점이 있다[5][6].

본 논문에서 제안한 MWNN의 성장 및 소거 알고리즘을 단일 차원 및 다차원 함수의 근사화에 적용하여 수행능력을 검증하였다.

## 2. MWNN의 학습 방법

EM 알고리즘의 반복(iteration)은 추정단계(E-step)와 극대화 단계(M-step)인 두 단계로 구성된다. M-단계는 E-단계의 각 반복 과정에서 재 정의된 유사 함수를 극대화하는 과정이다. EM 알고리즘은 먼저 주어진 데이터

접수일자 : 2001년 12월 1일

완료일자 : 2001년 12월 24일

본 연구는 과학기술부의 뇌과학연구 프로그램의 지원 하에 수행되었습니다.

나 현재 모델에서 완전한 데이터 유사 함수의 기댓값을 정의한다. 이것이 E-단계이다.

$$C_F(\theta, \theta^i) = E[l_c(\theta, D) | X] \quad (1)$$

여기서  $\theta^i$ 는  $i$ 번째 반복에서 매개변수 값,  $E[\cdot]$ 는  $\theta^i$ 에 대한 기댓값,  $X$ 는 불완전 데이터,  $l_c(\theta, D)$ 는 완전 데이터 유사 함수,  $D$ 는 완전 데이터이다.

M-단계에서는 결정론적인 함수  $Q$ 를 고려하여 새로운 매개변수  $\theta^{i+1}$ 을 추정하기 위해  $\theta$ 와 관련된 함수를 극대화한다.

$$\theta^{i+1} = \arg \max_{\theta} Q(\theta, \theta^i) \quad (2)$$

다시 E-단계에서 완전한 데이터 유사 함수를 개선하기 위한 추종이 이루어지고 이 과정을 반복하게 된다. 이러한 일반적인 EM 알고리즘을 Xu et al.가 MN에 적용하여 학습 알고리즘을 개발하였다.

### 2.1 EN의 가중치 갱신 규칙

MWNN의 구조에서  $t$ 번째 학습 데이터에 대한 MWNN의 출력은 다음과 같다[7].

$$y_k^i(x^t, w_k^i) = \sum_{m=1}^M w_{km}^i \psi_{k, mn}^i(x^t) \quad (3)$$

다음은  $t$ 번째 학습 데이터가 주어질 때 비용함수에 대한  $k$ 번째 모듈에서  $j$ 번째 가중치 벡터  $w_k^i = [w_{k1}^i \ w_{k2}^i \ \dots \ w_{kM}^i]^T$ 의 변화량은 체인룰을 적용하면 다음과 같이 표현할 수 있다.

$$\frac{\partial C_F}{\partial w_k^i} = \frac{\partial C_F}{\partial y_k^i(x^t, w_k^i)} \frac{\partial y_k^i(x^t, w_k^i)}{\partial w_k^i} \quad (4)$$

결국 M-단계에서  $k$ 번째 모듈인 WNN의 가중치 벡터는 다음 식과 같이 미세조정 값에 의해서 매 반복마다 갱신된다.

$$w_k^i(i+1) = w_k^i(i) + \eta \sum_{m=1}^M h_k^i \frac{e_k^i}{\sum_k} \cdot \psi_{k, mn}^i(x^t) \quad (5)$$

여기서  $i$ 는 반복을 나타낸다.

### 2.2 GN의 가중치 갱신 규칙

$t$ 번째 학습 데이터가 주어질 때 비용함수에 대한 GN의 출력층 가중치  $a_k$ 의 변화량은 체인룰을 적용하여 다음과 같은 식을 얻을 수 있다.

$$\frac{\partial C_F}{\partial a_k} = \sum_{i=1}^N \sum_{k=1}^K h_k^i (1 - g_k(x^t, a_k)) \frac{\partial u_k}{\partial a_k} \quad (6)$$

여기서  $g_k(x^t, a_k)$ 는  $t$ 번째 학습 데이터에 대한 GN의  $k$ 번째 사전확률을 의미한다. 확률적인 기울기 상승법에 의해 매 반복마다 다음과 같은 식에 의해 갱신된다.

$$a_k(i+1) = a_k(i) + \eta \sum_{i=1}^N \sum_{k=1}^K h_k^i (1 - g_k(x^t, a_k)) f(net_k^i) v_m \quad (7)$$

위와 유사한 방법으로 유사함수에 대한  $m$ 번째 입력

층과  $k$ 번째 은닉층의 가중치 갱신규칙은 다음과 같다.

$$a_k^m(i+1) = a_k^m(i) + \eta \sum_{i=1}^N \sum_{k=1}^K \{h_k^i (1 - g_k(x^t, a_k^m)) f(net_k^i) a_k^m\} v'_m x_i^t \quad (8)$$

위의 가중치 갱신규칙은 MWNN의 확률적인 기울기 상승 학습[7]과 유사한 형태이지만 모든 학습 데이터에 대해서 사후확률과 오차를 고려한다는 것이 상이하다. 즉 E-단계에서 사후확률에 의해 지역오차가 고려되고, M-단계에서 가중치가 극대화된다.

## 3. MWNN의 성장과 소거 알고리즘

### 3.1 MWNN의 성장 알고리즘

MWNN의 성장 알고리즘은 모듈과 모듈을 구성하는 노드를 성장시키며 직관적인 방법을 사용한다. 추가되는 EN(WNN)은 입력공간을 포괄하고 전역 오차를 감소시키게 된다. 또한 기하학적인 성장 기준을 사용하기 때문에 추가되는 EN은 준 직교성을 보장하는 프레임으로 구성되며 최적의 노드 구성이 가능하게 된다.

#### • 초기화 과정

초기화 과정은 WNN의 성장 알고리즘과 동일하다. MWNN의 성장알고리즘은 초기에 하나의 모듈에서 시작한다. 따라서 초기화 과정에서는 EN 모듈의 초기 크기를 결정하고 관련 매개변수를 설정하는 과정이다.

- [단계 1] 초기 팽창 계수 결정
- [단계 2] 노드 사이의 거리 값 초기화
- [단계 3] 초기 노드의 이동계수 결정
- [단계 4] 초기 가중치 결정

#### • 적응 과정

적응과정에서는 초기화 과정을 통해 단일 모듈을 구성하고, 대상문제에 적용하여 학습을 수행한 다음 다중 모듈을 추가하는 과정이다. WNN의 성장 알고리즘에서 제안한 방법[4]을 사용하여 전역 오차를 감소시키도록 새로운 모듈을 추가한다. 추가되는 모듈을 구성하는 WNN의 매개변수는 발생하는 지역 오차와 입력시간을 반영하여 결정한다. 전역 오차가 원하는 문턱치에 도달할 때까지 모듈의 성장을 반복한다. 수행 목표치가 원하는 문턱치( $\epsilon_{MAE}$ )에 도달하면 성장 알고리즘을 종료한다.

- [단계 1] 새로운 모듈 추가(WNN의 매개변수 결정)
  - 새롭게 추가되는 모듈의 매개변수 결정
  - GN의 추가된 노드의 가중치 결정
- [단계 2] 새롭게 추가된 모듈 학습
- [단계 3] 전체 모듈 학습

### 3.2 MWNN의 소거 알고리즘

현재 일부 연구자들에 MNN과 HMNN의 전지 알고리즘이 개발되어 함수근사화 문제와 패턴인식에 적용하여 좋은 수행결과를 얻고 있다. Ramamurti et al.[8]은 적당하게 수행능력이 우수하도록 초기 MNN을 구성하고 학습 데이터를 이용하여 단순화시키는 방법을 사용하였다.

모든 학습 데이터를 이용하여 각 모듈의 사후확률의 합을 구하여 전지의 기준으로 사용하였다([8]-식 (9) 참조). Waterhouse et al.[9]은 HMNN에 전지 알고리즘을 적용하였다. 학습 데이터를 이용하여 CART 트리구조의 모든 노드에 대한 전지 문턱치  $J_i^{(t)}$ 를 구하고 노드의 활성정도를 판단하여 매 단계마다 적용적으로 노드를 삭제하는 방식을 제안했다. 또한 Whitworth et al.[10]은 Waterhouse가 제안한 구조를 이용하여 서브트리의 루트 GN의 크로스 엔트로피를 구하여 연속적으로 하나의 모듈이 존재할 때까지 서브트리를 제거하면서 오차를 저장하는 방법을 사용하였다. 전지된 구조 중에서 오차가 가장 작은 구조를 선택하여 사용한다. 위에서 제안한 방식들은 전체 모듈을 제거하는데 전지 알고리즘을 사용하였다. 또한 초기 MNN의 구성에도 적합한 구조를 찾는 데 시행착오 방법을 사용하였다. 따라서 초기 구성된 MNN의 크기에 따라 최적의 MNN 구조를 찾는 데 실패할 가능성도 존재한다. 따라서 본 논문에서는 이러한 문제점을 해결하고, MWNN의 구조에 적합한 전지 알고리즘을 제안한다. 제안한 알고리즘은 수행능력이 부족한 모듈을 제거하는 방식을 사용하지 않고, 각 모듈의 노드를 제거하도록 설계하였다. 웨이블릿 신경망의 지역화 특성을 근거로 GN에 의해 입력공간의 분할된 상태를 고려하여 지역 오차 감소에 영향을 미치지 못하는 각 모듈의 노드를 제거한다.

제안한 MWNN의 성장 알고리즘은 새로운 모듈을 추가할 때 전체 입력공간을 고려하기 때문에 GN에 의해 부분적으로 입력공간이 분할되어 특정 입력공간에서만 각 GN이 동작한다. 그림 1은 MWNN에서 게이팅의 역할과 대상 함수에 대한 모듈의 출력을 나타낸 것이다. GN은 완전하게 학습이 종료되면 입력공간을 소프트하게 분할한다. 설명의 단순화를 위해서 그림 1과 같은 분할된 영역의 대상 함수를 고려하였다. 그림과 같이 3영역으로 정확하게 구분되는 대상 함수가 주어지고 MWNN이 3개의 모듈로 구성될 경우 GN은 정확하게 3개의 입력공간을 소프트하게 분할하여 특정 영역에서 각각의 모듈이 동작하도록 스위칭하게 된다.

그림 1에서 첫 번째 그림 (a)는 GN에 의해서 입력공간을 3개의 모듈이 분할한 결과를 이해하기 쉽게 나타낸 것이다. 여기서 타원의 중심과 크기는 입력공간에서 분할되는 부공간에 존재하는 입력의 평균과 분산으로 생각할 수 있다. 그림 1의 (c)는 (b)와 같은 대상 함수가 주어지고 학습이 완료된 후의 영역분할을 수행한 GN의 출력을 나타낸 것이다. 그림 (d)는 엑스퍼트의 출력을 나타낸 것이다. (d)에서 확인할 수 있듯이 Module I은 GN에 의해서 Region I에서 동작하고((d) 참조), Region II와 Region III에서의 출력은 0에 가깝다. 따라서 그림 (c)의 ㉠영역에서만 Module I이 동작하는 것과 동일하다. 또한 Module II는 ㉡영역에서 Module III은 ㉢영역에서 동작하게 된다. 다음 표 1은 각 모듈의 동작 영역과 비동작 영역을 정리한 것이다.

표 1. 모듈의 동작영역과 비동작 영역

	동작 영역	비동작 영역
Module I	㉠	㉡+㉢
Module II	㉡	㉠
Module III	㉢	㉠+㉡

새로운 모듈을 추가할 때 전체 입력 공간을 고려하기 때문에 각각의 모듈의 출력이 0인 영역의 노드는 지역 오차를 감소시키는데 기여하지 못함을 알 수 있다. 따라서 불필요한 노드를 제거해도 전체 수행능력에는 큰 영향을 미치지 못한다. 이것은 모듈로 사용하는 WNN의 지역화된 특성(웨이블릿 함수의 지역화 특성)에 기인한다. 위와 같은 원리를 바탕으로 전지 알고리즘은 각 모듈의 최적구조를 설계하기 위해 불필요한 모듈의 노드를 제거함으로써 MWNN의 최적 구조를 설계할 수 있는 방법을 제공할 수 있다.

전지 알고리즘은 성장 알고리즘이 종료된 후에 저장된 MWNN의 관련 매개변수를 로드하여 초기 전지 알고리즘을 적용하기 위한 구조로 사용한다.

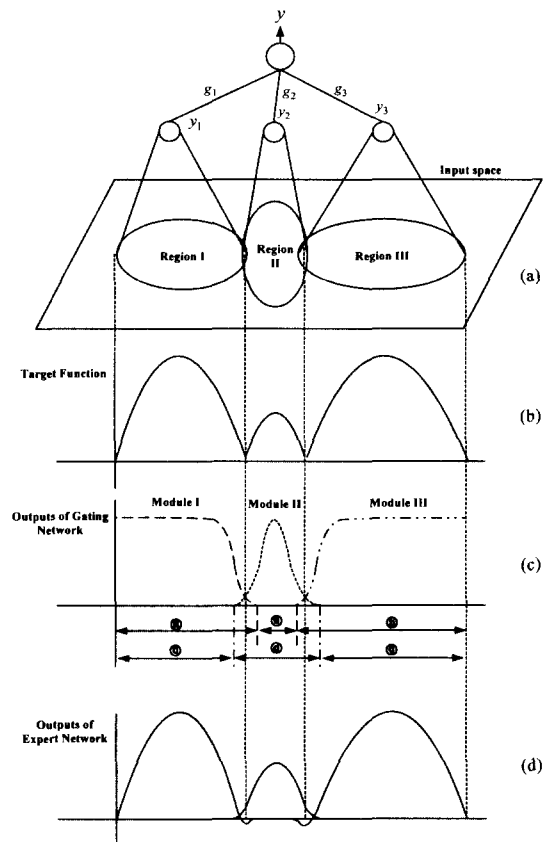


그림 1. GN에 의한 입력공간 분할

다음은 전지 알고리즘을 단계별로 상세하게 설명한다.

[단계 1] 모듈의 불필요한 노드 제거

GN의 출력을 참조하여 실제 지역오차를 감소시키는데 작용하지 않는 모듈의 노드를 제거하는 단계이다. 각 모듈의 노드로 사용되는 웨이블릿 함수의 이동계수와 팽창계수에 해당되는 입력공간을 고려하여 GN의 출력을 검사한다. 이 때 GN의 출력이 문턱치( $\epsilon_p$ ) 이하인 경우 노드를 제거하고 제거한 노드와 연결된 가중치를 0으로 초기화하여 새로운 구조의 모듈을 생성한다. 위와 같은 과정을 다른 모듈에 반복적으로 적용하여 불필요한 노드를 제거함으로써 모든 모듈의 최적구조를 찾는다. 실제 적용에서는 모듈의 노드를 제거하는 기준이 되는  $\epsilon_p$ 의

값을 작은 값으로 설정한다. 만약 모듈의 전체 노드가 제거된다면 모듈 자체를 제거시킨다. 이 과정에서는 GN의 해당 출력층 노드를 제거하고 제거된 노드와 관련된 가중치를 0으로 초기화하는 과정이 추가된다.

[단계 2] 학습단계

[단계 1]에서 각 모듈의 불필요한 노드를 제거하여 최적구조로 설계된 MWNN의 재학습을 수행하는 단계이다.  $\epsilon_p$ 의 값을 0으로 설정하면 실제적으로 학습과정이 필요하지 않지만 0이 아닌 작은 값으로 설정되면 제거된 노드가 지역 오차 감소에 어느 정도 기여하고 있기 때문에 재학습 과정이 필요하다. 학습단계에서는 GN의 매개변수를 조정할 필요가 없기 때문에 고정시키고 모듈의 가중치를 조정하는 학습단계를 수행한다. 오차의 변화량이 E-6 이하일 경우 학습단계를 종료한다.

4. 모의실험 및 검토

Mexicanhat 웨이블릿 함수에 대한 Daubechies의 계산결과를 참조하여  $a=2$ (최대 고주파수 표현이 가능한 최소 정수)로 정의하고,  $b \in (0.25, 1.875)$ 에 속하는  $b=1.0493$ 을 선택하였다.

4.1 모의실험 1

모의실험 1에서의 근사화 대상함수는 다음 식과 같은 연속함수인 Hermite 함수(식 9)이고, 시간축 상에서 [-4, 4]를 고려하였으며 학습과 테스트 데이터 201개를 사용하였다.

$$f(x) = 1.1(1 - x + 2x^2)e^{-\frac{x^2}{2}} \quad (9)$$

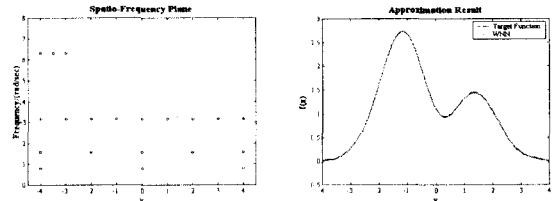
이 함수는 선형과 비선형이 혼합되어 있는 함수로 MWNN으로 함수 근사화기를 설계할 경우 GN이 근사적으로 2개의 영역으로 입력공간을 분할할 수 있다는 가능성을 확인하고 적용한 것이다.

그림 2는 MAE를 0.1로 설정하고 성장 알고리즘을 적용한 결과를 보인 것이다. 제시한 근사화 수행결과에서 확인할 수 있듯이 MWNN의 성장 알고리즘이 성공적으로 적용되었음을 확인할 수 있다. 그러나 MWNN의 성장 알고리즘은 전체 입력공간을 고려하여 MAE를 근거로 모듈의 노드를 추가하기 때문에 추가되는 노드의 수가 증가하였다. 즉 추가되는 모듈에 의해 일부 입력공간에서는 지역 오차를 감소시킬 수 있지만 GN에 의해 영역이 분할되어 특정 영역에서는 해당 모듈이 동작하지 못하기 때문에 지역 오차를 감소하는데 기여하지 못하게 된다. 따라서 불필요한 잉여 노드를 추가되어 WNN의 성장 알고리즘에 비해 추가되는 노드의 개수가 증가되었다. 성장과정에서 4개의 모듈이 사용되었다.

그림 3은 MWNN의 성장 알고리즘이 적용되는 성장패턴을 나타낸 것이다. 가로축은 추가되는 노드의 개수를 나타내고, 세로 축은 새로운 노드가 추가되기 전의 RMSE를 나타낸 것이다. 초기에 각 모듈이 추가되면서 RMSE가 급격하게 감소하고, 후반부에서는 미세조정 단계로 동작하여 RMSE의 감소량이 작다.

그림 4는 MWNN의 성장 알고리즘의 단점을 극복하기 위해 성장 알고리즘이 종료하고 저장한 MWNN의 매개

변수를 이용하여 본 논문에서 제안한 전지 알고리즘을 적용한 결과를 보인 것이다.



(a) 노드의 시간-주파수 분포 (b) 함수의 근사화 결과  
그림 2. 성장 알고리즘의 적용 결과(MAE: 0.1)

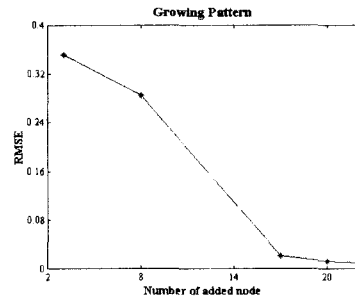
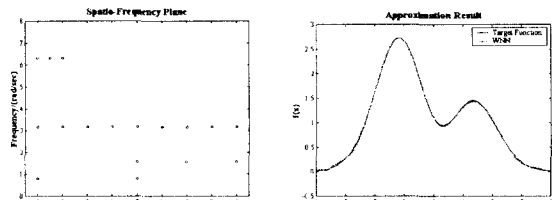


그림 3. MWNN의 성장 패턴(MAE: 0.1)

모듈의 노드를 제거하는 기준이 되는  $\epsilon_p=0.001$ ,  $\epsilon_v=0.01$ 로 설정하였다. GN에 의해 입력공간이 분할되고 각 모듈을 구성하는 EN은 해당 영역에 한정되어 학습을 수행하기 때문에 성장 알고리즘이 종료된 후에 EN의 출력은 지역화되는 특성을 나타낸다. 그림 2의 (a)에서 확인할 수 있듯이 첫 번째 모듈과 두 번째 모듈은 각각 입력공간의 왼쪽과 오른쪽으로 분할되어 지역화되었기 때문에 전지 알고리즘을 적용할 경우, 각 모듈이 분할된 해당 영역에서 출력을 내지 못하는 노드가 제거되었다.

그림 4의 (b)는 전지 알고리즘을 적용하여 해당 모듈의 노드를 제거한 후 재학습을 수행한 결과를 나타낸 것이다. 그림 2의 (b)와 비교해 볼 때 특정 모듈의 노드가 삭제된 후에도 수행결과가 크게 악화되지 않았음을 확인할 수 있다. 이것은 모듈의 구성하는 WNN의 지역화 특성에 기인한 것이다.



(a) 노드의 시간-주파수 분포 (b) 함수의 근사화 결과  
그림 4. 소거 알고리즘의 적용 결과(MAE: 0.1)

표 2는 MWNN의 성장과 전지 알고리즘을 적용한 수치적인 결과를 나타낸 것이다. 전지 알고리즘을 적용한 후에 전체 노드의 개수가 감소되었음을 확인할 수 있다. RMSE는 성장과 전지 알고리즘이 종료된 후 테스트 데

이터를 이용하여 수행한 결과를 나타낸 것이다. 성장과 전지 알고리즘의 수행결과를 비교하면 전지 알고리즘을 수행한 후의 RMSE가 증가되었음을 확인 할 수 있다. 이것은 삭제된 모듈의 노드가 팽창계수의 폭 만큼 주변 노드에 영향을 미치기 때문이다.

표 2. 성장과 소거 알고리즘 적용 결과

	전체 노드의 개수	RMSE
성장 알고리즘 적용	20	0.00982
전지 알고리즘 적용	17	0.01024

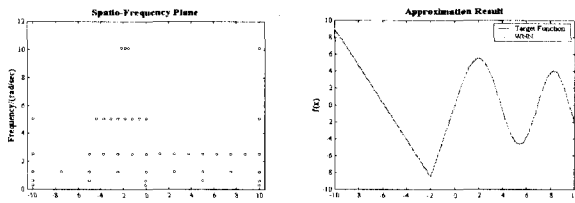
4.2 모의실험 2

모의실험 2에서는 식 (10)과 같은 부분적으로 연속적인 근사화 함수를 사용하였다. 근사화 대상함수는 시간축 상에서 [-10, 10]을 고려하였고, 학습과 테스트 데이터는 201개를 사용하였다.

$$f(x) = \begin{cases} -2.186x - 12.864 & -10 \leq x < -2 \\ 4.246x & -2 \leq x < 0 \\ 10e^{-0.05x-0.5} \sin((0.03x+7)x) & 0 \leq x \leq 10 \end{cases} \quad (10)$$

일반적으로 1-D 함수의 근사화 문제에 자주 적용되는 부분적으로 연속적인 함수의 근사화 문제는 모듈화식 설계 방식의 장점을 설명하기 위해 많이 사용되고 있다. 본 모의실험에서는 MWNN의 모듈화식 설계 방식이 이러한 함수의 입력공간을 적절하게 분할하여 근사화 수행능력의 향상을 달성할 수 있는지 여부와 제안한 전지 및 성장 알고리즘의 성능평가를 위한 것이다. 모듈의 노드를 제거하는 기준과 과도적응을 판단하는 기준이 되는 문턱치는 각각  $\epsilon_p = 0.05$ ,  $\epsilon_e = 0.12$ 로 설정하였으며, 입력공간이 좀 더 복잡하기 때문에 GN의 은닉층은 4개를 사용하였다.

그림 5는 MAE를 0.2로 설정하고 MWNN의 성장 알고리즘을 수행한 후의 결과를 보인 것이다. 전체 5개의 모듈이 사용되었으며, 모의실험 1과 유사하게 성장 알고리즘이 성공적으로 적용되었음을 확인할 수 있다. 그림 6은 성장 알고리즘의 성장 패턴을 나타낸 것이다. 성장 패턴의 유형이 모의실험 1과 유사하고, 함수의 진폭이 큰 부분에 미세조정을 위해 추가 모듈이 사용되었다.



(a) 노드의 시간-주파수 분포 (b) 함수의 근사화 결과  
그림 5. 성장 알고리즘의 적용 결과(MAE: 0.2)

그림 7은 전지 알고리즘을 적용한 결과를 보인 것이다. 마지막 추가된 모듈을 제외하고 나머지 4개 모듈의 노드가 삭제되었음을 확인할 수 있다. 각각의 모듈이 지역 오차를 보상하도록 부분적으로 추가되었기 때문에 전지 알고리즘에 의해 각 모듈에서 많은 노드가 제거되는 않았다. 따라서 모듈화된 설계 방식을 적용할 경우, 부분

적으로 연속적인 함수의 근사화 문제에 대해서 최적구조 설계가 가능함을 확인할 수 있었다.

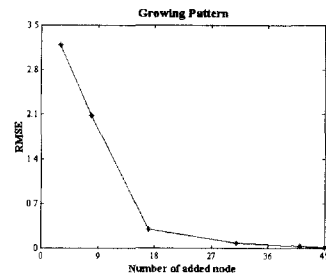


그림 6. MWNN의 성장 패턴(MAE: 0.2)

표 3은 MWNN의 성장과 전지 알고리즘을 적용한 수치적인 결과를 나타낸 것이다. MAE를 0.2로 설정하여 모의실험을 수행한 결과에서 추가되는 모듈의 노드가 MAE 값을 기준으로 입력공간에서 지역 오차가 크게 발생하는 영역에만 추가되어 지역화가 부분적으로 달성되기 때문에 모듈에서 삭제되는 노드의 수가 크게 증가하지 않았다. 따라서 MAE를 이전의 모의실험 보다 작은 값(0.1)으로 설정하여 MWNN의 전지 알고리즘의 수행능력을 검증하고 그 결과를 표 3에 나타내었다. 수행결과에서 확인 할 수 있듯이 MAE를 작은 값으로 선정하여 추가되는 노드의 개수가 증가했고, RMSE가 감소되었다. 따라서 지역화가 완료된 후에 각 모듈에서 삭제되는 노드의 수가 증가하였음을 확인할 수 있다.

표 3. 성장과 소거 알고리즘 적용 결과

MAE : 0.2	전체 노드의 개수	RMSE
성장 알고리즘 적용	45	0.01147
전지 알고리즘 적용	33	0.01893
MAE : 0.1	전체 노드의 개수	RMSE
성장 알고리즘 적용	62	0.008731
전지 알고리즘 적용	41	0.012731

4.3 모의실험 3

모의실험 3에서 2차원 근사화 대상함수는 식 (11)과 같은 Harmonic 함수이고, 시간축 상에서 [0, 1]×[0, 1]을 고려하였으며, 학습과 테스트 데이터는 225개를 사용하였다.

$$f(t_1, t_2) = 42.659[0.1 + \hat{t}_1(0.05 + \hat{t}_1 - 10\hat{t}_1\hat{t}_2 + 5\hat{t}_2^2)] \quad (11)$$

여기서,  $\hat{t}_1 = t_1 - 0.5$ 이고,  $\hat{t}_2 = t_2 - 0.5$ 이다.

표 4의 수치결과(MAE : 0.2)와 같이 정밀한 지역 오차를 보여줘야 모듈의 노드가 추가되지만 MWNN의 모듈화 특성 때문에 추가되는 노드의 개수가 WNN의 성장 알고리즘 보다 감소하였다. 따라서 모듈화 설계의 장점을 확인할 수 있었다. 또한 전지 알고리즘도 성공적으로 적용되었다. MAE를 0.2와 0.4로 설정한 모의실험 결과를 비교하면 전지 알고리즘에 의해 제거되는 노드의 개수는 유사하다. 이것은 MWNN의 성장 단계에서 모듈의 지역화가 달성되고 되어 지역 오차가 급격하게 감소하였기 때문이다.

표 4. 성장과 소거 알고리즘 적용 결과

MAE : 0.4	전체 노드의 개수	RMSE
성장 알고리즘 적용	194	0.028240
전지 알고리즘 적용	139	0.041236
MAE : 0.2	전체 노드의 개수	RMSE
성장 알고리즘 적용	260	0.001938
전지 알고리즘 적용	206	0.002713

### 5. 결론

본 논문에서는 MWNN의 성장과 소거 알고리즘을 개발하였다. 제안한 알고리즘의 목표는 MWNN의 초기 매개변수 결정과 대상문제에 적합한 모듈의 개수 및 모듈의 노드 개수를 결정하여 MWNN의 최적구조를 찾는 것이다. MWNN의 성장 알고리즘은 지역오차를 감소시키도록 모듈의 크기를 결정하여 자동으로 모듈을 추가하게 된다. 또한 성장 알고리즘이 종료된 후에 모듈의 잉여노드를 제거하여 MWNN의 최적구조 설계를 완성한다.

제안한 MWNN의 성장과 소거 알고리즘을 함수근사화 문제에 적용하여 최적구조 설계 알고리즘의 우수성과 근사화 능력을 검증하였다. 또한 제안한 알고리즘은 단일차원의 함수근사화 문제에 성공적으로 적용될 수 있음을 검증하였다. 그리고 모의실험 3을 통해 시간-주파수 분석법[7]을 이용한 최적구조 설계 알고리즘의 문제점인 다차원 확장 문제를 해결할 수 있음을 확인하였다. 이러한 결과를 통해서 모듈화식 설계가 갖는 장점을 적용하고, 적합한 망의 크기를 결정할 수 있는 가능성을 확인하였다.

### 참고 문헌

[1] R. Reed, "Pruning Algorithm - A Survey," *IEEE Trans. on Neural Networks*, vol. 6, pp. 610-622, 1995.

[2] J. Ghosh and K. Tumer, "Structural Adaptation and Generalization in Supervised Feedforward Networks," *Journal of Artificial Neural Networks*, vol. 1, no. 4, pp. 431-458, 1994.

[3] J. Fristch, M. Finke, and A. Waibel, "Adaptively Growing Hierarchical Mixtures of Experts," *Advances in Neural Information Processing Systems-9*, M. I. Jordan, M. C. Mozer, and T. Petsche, Eds. Cambridge, MA: MIT Press, pp. 459-465, 1997.

[4] 서재용, 김용택, 조현찬, 김용민, 전홍태, "F-투영법을 이용한 웨이블릿 신경망의 성장 알고리즘," *대한전자공학회 2001 하계종합학술대회 논문집*, 제24권 제1호, vol. 3, pp. 165-168, 2001.

[5] S. R. Waterhouse and A. J. Robinson, "Constructive Algorithms for Hierarchical Mixture of Experts," *Advances in Neural Information Processing Systems-8*, D. Touretzky, M. Mozer, and M. Hasselmo, Eds. Cambridge, MA:MIT Press, pp. 584-590, 1996.

[6] 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.

[7] 서재용, 김용택, 조현찬, 전홍태, "시간-주파수 분석을 이용한 모듈라 웨이블릿 신경망의 최적 구조 설계," *대한전자공학회 논문지*, 제38권, SC편, 제2호, pp. 12-19, 2001.

[8] R. A. Jacobs 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.

[9] Masumi Ishikawa, "Learning of modular structured networks," *Artificial Intelligence*, vol. 75, pp. 51-62, 1995.

[10] 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.

### 저자 소개

#### 서재용(徐在龍)정회원

1969년 7월 21일생  
 1996년 2월 중앙대학교 전자공학과 졸업(학사)  
 1998년 2월 중앙대학교 대학원 전자공학과 졸업(석사)  
 2001년 9월 중앙대학교 대학원 전자공학과 졸업(박사)  
 2000년 8월 ~ 현재 : 한국기술교육대학교 정보기술공학부  
 객원교수

주관심분야 : 퍼지, 유전알고리즘, 신경망, FNN, 웨이블릿  
 신경망 등임

#### 김성주(金成柱)정회원

정회원, 1971년 12월 12일생  
 1999년 2월 중앙대학교 전자공학과 졸업(학사)  
 2001년 2월 중앙대학교 대학원 전자공학과 졸업(석사)  
 2001년 3월 ~ 현재 : 중앙대학교 대학원 전자전기공학부  
 박사과정 재학중

주관심분야 : 신경망, 로보틱스, 퍼지, 웨이블릿 신경망 등임

#### 조현찬(趙鉉讚)정회원

현재 : 한국기술교육대학교 정보기술공학부 전자공학과 교수

#### 전홍태(全洪兌)정회원

현재 : 중앙대학교 전기전자공학부 교수