

동적 신경망의 층의 분열과 합성에 의한 비선형 시스템 제어

論 文

48A-4-10

Control of Nonlinear System by Multiplication and Combining Layer on Dynamic Neural Networks

朴省昱* · 李載寬** · 徐輔煥***

(Seong-Wook Park · Jae-Kwan Lee · Bo-Hyeok Seo)

Abstract - We propose an algorithm for obtaining the optimal node number of hidden units in dynamic neural networks. The dynamic neural networks comprise of dynamic neural units and neural processor consisting of two dynamic neural units; one functioning as an excitatory neuron and the other as an inhibitory neuron. Starting out with basic network structure to solve the problem of control, we find optimal neural structure by multiplication and combining dynamic neural unit. Numerical examples are presented for nonlinear systems. Those case studies showed that the proposed method is useful in practical sense.

Key Words : Evaluation structure, Multiplication and combining, Learning algorithm

1. 서 론

'92년에 생물학적 신경을 모형화한 상태변수 모형을 가진 동적 신경기가 개발되었는데 기존의 케 환 신경기에 비해 학습회수를 줄일 수 있고 회상능력이 우수한 특성을 가진다[1]. 이 동적 신경기는 시냅스(synapse) 입력이 바로 출력에 영향을 주지 않고 어느 정도 시간이 지난 후에 영향을 주는 중앙신경 시스템의 반향회로(reverberating circuit)와 소마전류(soma current)를 변환시키는 축삭(axon)으로 구성되어 있다. 동적 신경기로 구성된 신경망은 온라인으로 동적 시스템의 식별과 제어가 가능하지만 중간층의 노드 증가에 따라 비선형 함수의 이득과 학습변수값이 불안정하여 출력이 발산한다. 저자는 학습처리시간이 빠르고 안정하게 적응변수를 학습할 수 있는 동적 신경망을 개발하였다[2]. 동적 신경망을 모아서 하나의 전송요소로 하여 이 전송요소의 출력에 따라 활성화 전송군(excitatory neurons)과 금지화 전송군(inhibitory neurons)을 가진 동적 신경망으로 비선형 다변수 시스템을 식별하고 제어하였다[3]. 하지만 여러 실제응용에 있어서, 어떤 구조의 중간층을 사용해야 하는가와 몇 개의 중간층 노드를 사용할 것인지를 결정해야 한다. 두 개의 중간층을 가진 구조는 원하는 문제를 해결하는 능력이 있으므로 특별히 중간층 노드를 어떻게 선택해야 하는가는 상당히 중요하다. 다층형 신경망에서 중간층의 수가 많을수록 문제해결에 필요한 많은 결정함수(discriminant function)를 가지고 있다. 실제응용에 있어, 여러 개의 중간층을 사용하는 대신에 중간층의 노드의 수를 가지고 문제를 해결한다. 신경망의 구조에 관한 연구는 확률적인 접근방법[4,5]

을 사용하여 연구되고 있지만 적용분야에 따라 일관성이 없고 많은 훈련용 패턴을 요구하므로 적용하는 경우에 많은 제한이 있으므로 중간층의 특성에 관한 연구를 활발히 하고 있다[6,7]. 여기서 확률적인 접근방법이란 시스템을 입력과 출력의 조건부 확률로 표현하고 전방향 신경망을 입력과 연결계수를 조건부로 하는 출력으로 표현하여 시스템을 가장 잘 근사화 할 수 있는 신경망 구조와 연결계수를 찾는 것으로 최적 모형과 변수를 동시에 구하는 과정을 말한다. 하지만 케 환 신경망의 구조에는 적용할 수 없고 비용함수에 대한 최적의 분산행렬을 알 수 없기 때문에 샘플 데이터에 의존해야 한다[4]. 다른 방법으로 출력층으로 전달되는 중간층의 합이 가우시안(Gaussian) 분포를 가진다는 가정에서 신경망 구조를 평가하는 조건을 얻을 수 있지만 많은 샘플 데이터와 큰 신경망 구조에만 적용 가능하다[5]. 중간층의 역할과 중간층과 입력층의 연결계수의 최대 상관계수 그리고 선형 출력오차에 대한 중간층과 출력층간의 연결계수를 조절하였지만 일반적인 신경 회로망에 적용하기가 어렵다[6]. 중간층과 출력오차 사이의 분산이 최대가 되도록 중간층 노드를 추가하는 방법 그리고 비용함수를 사용하여 노드에 대해 최적의 비용함수를 가지도록 신경망 구조를 찾는 방법이 연구중이다[7].

본 논문에서는 실시간 적용 가능한 동적 신경모형을 이용하여 비선형 시스템을 제어하는 동적 신경망 제어를 구성하고, 학습도중에 학습한계를 판단하는 해석적인 방법과 층의 노드 분열과 합성하는 기법을 연구하여 원하는 출력을 얻는 적절한 노드를 가진 동적 신경망 제어를 설계한다.

2. 사용된 동적 신경 모형

시냅스, 동적 인자 프로세서, 그리고 소마이득(soma gain)으로 구성된 동적 신경모형을 그림 1에 나타내었다. 동적 인자 프로세서는 두 개의 지연요소와 전향 가중치와 케 환 가중치로 구성된다. $x(k)$, $v_1(k)$ 는 동적 인자 프로세서의 입력과 출력을 나타낸다. 사용된 동적 신경모형은 시냅스 기능과 소

* 正 會 員 : 龜尾1大 電氣科 助教授 · 工博

** 正 會 員 : 東北大 電氣 · 通信工學科 助授 · 工博

*** 正 會 員 : 慶北大 電子 · 電氣工學科 教授 · 工博

接受日字 : 1998년 9월 23일

最終完了 : 1999년 3월 30일

마 기능을 생물학적 관점과 같은 기능이 되도록 역치 (threshold)를 비선형 함수의 변수에 포함시켜 $v_1(k)$ 가 $\theta(k)$ 보다 큰 경우에 비선형 변환된다. 그림 1에서 $I_j(k)$ ($j=1, \dots, n$)는 다른 신경으로부터의 입력이고 $x(k)$ 는 학습된 시냅스 결합에 입력신호를 곱한 값의 합으로 임의의 신경이 가지는 외부 입력의 총합이다. 이 외부입력 신호는 반향회로 즉 신경 시스템의 동적 인자 프로세서를 통하여 축색으로 전달된다. 전달된 신호는 생물학적으로 소마전위(soma electro-chemical potential)에 해당되고 축색은 이 전위를 연속적인 비선형 전류로 변환한다. 이 전류는 함수의 기울기, 함수의 최대 기울기를 조절하는 역치, 그리고 소마전위를 변수로 하는 활성화 함수로 나타낼 수 있다. 그림 1의 선택 부분의 적용계수가 학습된다. 동적 인자 프로세서는 Rao의 동적 신경모형[1]의 2차 구조를 달리 표현한 구조로 리아푸노프(Lyapunov) 안정도가 보장되는 조건하에서 적용계수를 학습한다[2]. 동적 신경모형은 신경내부에 동적 인자 프로세서를 포함한 구조를 가지므로 한 개의 신경은 내부 신경의 전달함수의 동적 특성을 반영한다. 동적 인자 프로세서를 상태공간 표현식으로 나타내면 식(1-2)와 같다. 식(1-2)를 z변환하여 정리하면 식(3-4)와 같다.

$$s(k+1) = As(k) + Bx(k) \tag{1}$$

$$v_1(k) = Cs(k) + Dx(k) \tag{2}$$

단, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, $C = [c_1 \ c_2]$, $D = d_0$

$$\frac{S(z)}{X(z)} = (zI - A)^{-1} B X(z) = \frac{1}{(z - a_{11})(z - a_{22}) - a_{12}a_{21}} \begin{bmatrix} b_1 z - a_{22}b_1 + a_{12}b_2 \\ b_2 z - a_{11}b_2 + a_{21}b_1 \end{bmatrix} \tag{3}$$

$$V_1(z) = CS(z) + DX(z) = C(zI - A)^{-1} BX(z) + DX(z) \tag{4}$$

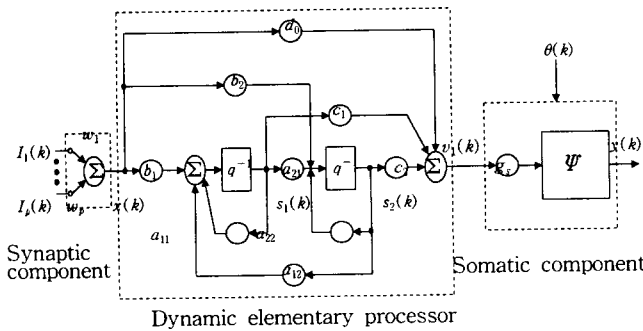


그림 1 동적 인자 프로세서를 가진 동적 신경기
Fig. 1 Dynamic neural unit with dynamic elementary processor

식(3)의 우변항 분모분자예를 z^{-2} 곱하여 정리하여 이산표현식으로 나타내면 식(5)와 같다

$$\frac{v_1(k)}{x(k)} = \frac{d_0 + (b_1 c_1 + b_2 c_2 - d_0 K_1)q^{-1} + (K_3 + d_0 K_2)q^{-2}}{1 - K_1 q^{-1} + K_2 q^{-2}} \tag{5}$$

단, $K_1 = a_{11} + a_{22}$, $K_2 = a_{11}a_{22} - a_{12}a_{21}$
 $K_3 = a_{12}b_2c_1 + a_{21}b_1c_2 - a_{22}b_1c_1 - a_{22}b_2c_2$

3. 동적 신경망의 층의 분열과 합성

신경망의 구조를 평가하기 위해서 중간층의 수와 중간층의 노드 수를 해석적으로 결정할 필요가 있다. 중간층의 노드 수에 따라 출력오차를 줄일 수 있으므로 노드의 수를 찾는 방법에 대해 연구한다. 즉 신경망 구조가 정해지면 층의 노드 수를 적절히 선택할 수 있어야 한다. 본 논문에서는 중간층의 노드 분열과 합성 알고리즘을 개발하여 동적 신경기에 적용하여 비선형 시스템을 제어하였다.

3.1 학습한계와 중간층 노드의 분열

신경망의 중간층 노드 수를 적절히 선택하기 위해서 신경망의 학습한계와 분열조건을 정의할 필요가 있다. 신경망의 학습한계라는 것은 중간층의 노드 수가 부족하여 학습이 효율적으로 진행되지 않는 상태로 정의한다[8]. 학습이 진행되는 동안 연결계수의 변화가 없는 경우로 시간에 따른 연결계수의 변화율로 학습한계로 판별하였다. 본 논문에서는 중간층 노드간의 상관계수와 오차의 실효치로 학습한계(learning limit)를 판단하였다. 신경망이 학습한계로 판명되면 그 한계 상태를 벗어나기 위해 중간층의 노드를 추가해야 한다. 추가 노드 수를 층과 노드를 이용한 그래프 구조로 증가시키는 방법[7]도 있지만 평가함수를 사용해야 하고 최적 해 접근방법을 사용하므로 많은 데이터와 시간이 요구되고 전방향 신경망에서만 사용 가능하다. 확률적인 개념을 이용한 해 접근 방법[4]이 있으나 신경망 구조에 대한 정보를 평가할 때, 비용함수에 대한 최적의 분산행렬을 알 수 없으므로 추가 노드에 의한 신경망 구조 변경시 추가 노드에 의한 변동으로 인해 학습능력이 처음의 신경망 구조보다 떨어진다[7]. 진화이론을 사용할 경우 신경망 구조에 대한 평가를 할 수 있으나 출력층으로 전달되는 중간층의 출력 합이 가우시안 분포를 가져야 하고 400개 이상의 입력 패턴을 분류할 만큼 입력층과 중간층의 노드가 충분해야하므로 학습하는데 많은 시간이 소요된다[5]. 결국 위의 방법들은 실시간 연산 처리가 가능한 동적 신경망에 적용하여 중간층의 노드를 분열하기에 적합하지 못하다. 그래서 아래와 같이 지수적으로 추가 노드 수를 결정하였다[9].

$$N_{new} = 2N_{old} \tag{6}$$

여기서 N_{old} , N_{new} 는 추가 전과 후의 중간층 노드의 총수를 나타낸다. 추가되는 중간층 노드의 수는 추가전의 노드 수와 동일하게 추가되므로 지수적으로 분열한다. 본 논문에서는 임의의 3층 이상의 동적 신경망에서 중간층의 노드 수를 합성하고 분열하였다.

3.2 중간층 노드의 합성

중간층 노드의 분열후 최적의 노드의 수를 가지기 위해서 불필요한 노드를 제거하는 합성조건을 알아야 한다. 신경망의 출력 자승오차와 상관관계를 이용해서 합성시기를 결정한다. 상관계수의 값이 잠시 일정한 값을 가지고 있다가 학습능력이 바뀌면서 상관계수의 값이 변화되므로 상관계수 변화

량으로 중간층 노드의 합성시기를 결정한다. 중간층 노드의 역할이 학습한계 상태가 되면 상관계수의 시간적 변화가 일정하다. 중간층 노드의 상관관계를 정량적으로 알기 위해 두 개의 중간층 노드 출력계열 $\{V_{ip} | p=1,2,\dots\}$, $\{V_{jp} | p=1,2,\dots\}$ 의 상관계수를 이용한다.

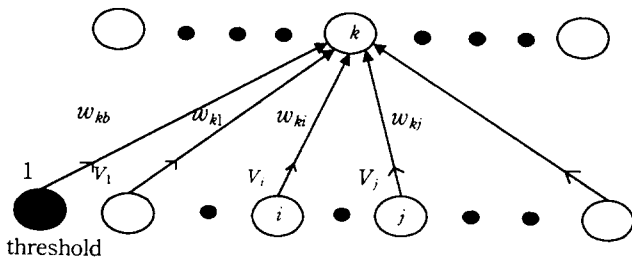
$$r_{ij} = \frac{E(V_{ip}V_{jp}) - E(V_{ip}) \cdot E(V_{jp})}{\sqrt{E(V_{ip}^2) - E(V_{ip})^2} \cdot \sqrt{E(V_{jp}^2) - E(V_{jp})^2}}, \quad -1 \leq r_{ij} \leq 1 \quad (7)$$

여기서 $E(V_{ip})$ 와 $E(V_{jp})$ 는 p 개의 데이터에 대한 중간층 노드 i 및 j 의 출력 기대치이다.

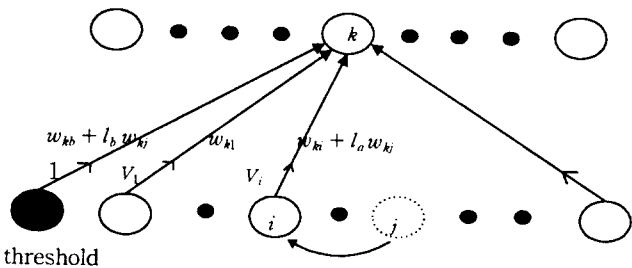
신경망의 학습상태가 학습한계 극한값보다 작으면 그림 2(a)에서 동일 중간층내의 노드 i, j 의 상관계수 r_{ij} 와 분산 s_i^2, s_j^2 가 식(8)을 동시에 만족할 때, 중간층 노드를 하나로 합성한다. θ_1 에 따라 합성을 통해 신경망의 노드를 줄일 수 있는 상수이고 θ_2 는 학습시 학습률을 변화시킬 수 있다.

$$|r_{ij}| \geq \theta_1, \quad s_i^2 > s_j^2 \geq \theta_2 \quad (8)$$

여기서, $s_i = \sqrt{E(V_{ip}^2) - E(V_{ip})^2}$, $s_j = \sqrt{E(V_{jp}^2) - E(V_{jp})^2}$



(a) 중간층 노드 i, j 와 문턱 노드
(a) Hidden units i, j and threshold unit



(b) 중간층 노드 i 와 j 의 합성
(b) Combination of hidden unit i and j

그림 2 과잉 중간층 노드의 합성

Fig. 2 Combination of extra hidden units

동일 중간층내의 노드 i, j 의 출력계열 $\{V_{ip}, V_{jp}\}$ 은 다변량해석(multi-variable analysis)에 의해 선형식으로 근사화할 수 있다[10]. 두 개의 노드가 가지는 값들을 이용하여 노드간의 관계를 선형 회귀식으로 표현하면 식(9)와 같다.

$$V_{jp} = l_a V_{ip} + l_b \quad (9)$$

여기서 l_a 는 직선의 기울기 계수이고 l_b 는 직선의 절편이다. l_a 와 l_b 는 아래와 같다.

$$l_a = \frac{E(V_{ip}V_{jp}) - E(V_{ip}) \cdot E(V_{jp})}{E(V_{ip}^2) - E(V_{ip})^2} \quad (10)$$

$$l_b = E(V_{jp}) - l_a E(V_{ip}) \quad (11)$$

식(10-11)을 사용하여 상위에 있는 중간층의 노드 k 에서의 입력을 구하면 식(12)가 된다.

$$\begin{aligned} w_{ki} V_i + w_{kj} V_j + w_{kb} \cdot 1 + \sum_{i \neq j} w_{ki} V_i \\ = w_{ki} V_i + w_{kj} (l_a V_i + l_b) + w_{kb} \cdot 1 + \sum_{i \neq j} w_{ki} V_i \\ = (w_{ki} + l_a w_{kj}) V_i + (w_{kb} + l_b w_{kj}) \cdot 1 + \sum_{i \neq j} w_{ki} V_i \end{aligned} \quad (12)$$

그림 2(b)에 나타난 것같이 중간층 노드 i 와 문턱신호(threshold signal)에 대한 연결계수는 식(13-14)로 갱신되고

$$w_{ki} = w_{ki} + l_a w_{kj} \quad (13)$$

$$w_{kb} = w_{kb} + l_b w_{kj} \quad (14)$$

중간층 j 노드를 소거한다. 합성에 있어서 출력계열의 분산이 작은 노드를 분산이 큰 노드에 합성시킨다.

4. 사용된 동적 신경 모형

비선형 함수의 식별능력과 동적 신경계어의 성능을 조사하기 위해서 사례연구하였고 결과를 비교 분석하였다. 중간층 노드의 분열과 합성을 통해 비선형 시스템을 식별하고 제어하므로써 동적 신경망의 유용성을 검증하였다.

사례연구 1에서 동적 신경기를 사용하여 임의의 비선형 함수를 식별하여 비선형 함수의 출력을 근사화하는 문제를 다룬다. 비선형 함수 $f[x(k)]$ 가 식(15)와 같은 경우에 식별오차가 작게 수렴하는 가를 조사한다.

$$f[x(k)] = 0.2x^3(k) + 0.3\sin(2\pi \cdot x(k-1)) + 0.1\sin(5\pi \cdot x(k-2)) \quad (15)$$

여기서 $x(k)$ 는 크기가 $[-1, 1]$ 인 정현파 입력으로 식(16)으로 표현되고 k 는 이산 시간으로 매 시각 샘플링을 나타낸다.

$$x(k) = \sin\left(\frac{2\pi}{250} k\right), \quad 0 < k \leq 1000 \quad (16)$$

출력오차로부터 학습 알고리즘[2]을 사용하여 동적 신경기의 학습변수를 조절하여 비선형 함수의 출력을 근사화한다.

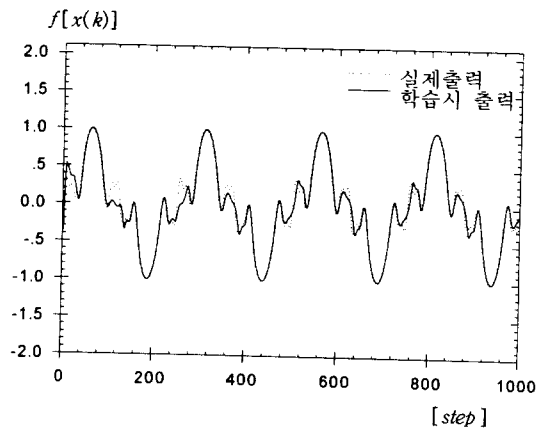


그림 3 동적 신경기를 이용한 비선형 함수의 식별
Fig. 3 Nonlinear function and its approximation by dynamic neural unit

학습시 사용한 적응변수 초기치, 학습율, 그리고 활성화 함

수와 문턱신호는 참고문헌[2]에 정의된 변수값에 따른다. 사례연구 1에서는 사용한 동적 신경모형의 함수 근사화 능력을 알아보기 위해 층이 한 개인 동적 신경기를 사용하여 임의의 비선형 함수를 식별하였다. 학습한 결과를 그림 3에 나타내었다. 학습하면서 비선형 함수를 식별하는 기법을 사용하여 임의의 비선형 함수를 원하는 만큼 정확히 근사화함을 알았다.

다른 케환 신경망을 사용한 결과와 비교 검토하기 위해 다층 케환 신경망과 혼돈(chaos) 신경망으로 반복 학습한 결과를 그림 4와 그림 5에 나타내었다. 학습한 결과로부터 오차의 실효치를 표 1에 나타내었다. 사용한 신경모형을 사용한 경우 오차의 실효치가 가장 작음을 알 수 있었다. 케환 신경망과 혼돈 신경망은 입력층이 3개, 중간층이 30개, 출력층이 1개인 3층 신경망을 사용하여 입·출력쌍을 2000번 학습하였다. 학습된 총 학습 변수는 중간층과 출력층의 상호 간섭항(interacting weight parameter)과 자신의 케환항을 합쳐 혼돈 신경망 식별기에서 1051개이며 케환 신경망에서는 1050개이다. 학습하는데 시간이 많이 소요되었고 혼돈 신경망을 사용하여 학습한 경우에 오차의 실효치가 케환 신경망보다 전반적으로 작은 값으로 수렴함을 알 수 있었고 온 라인으로 학습을 하는 동적 신경기보다 오차가 큰 것을 알 수 있었다.

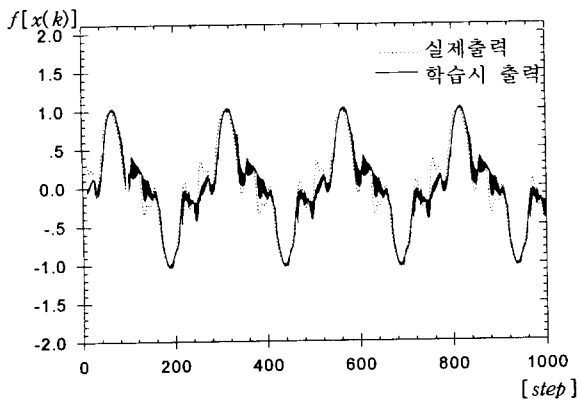


그림 4 3층 케환 신경망을 이용한 비선형 함수의 식별
Fig. 4 Nonlinear function and its approximation by three layer recurrent neural networks

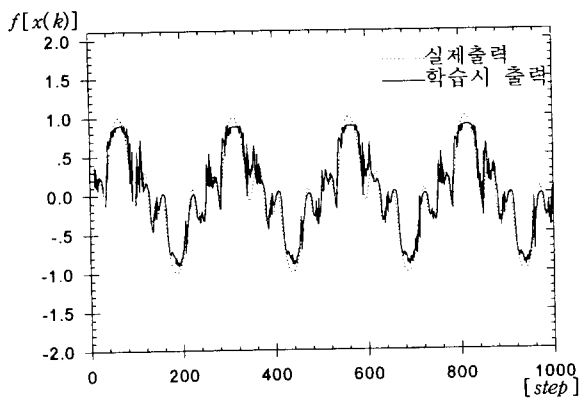


그림 5 3층 혼돈 신경망을 이용한 비선형 함수의 식별
Fig. 5 Nonlinear function and its approximation by three layer chaotic neural networks

표 1 사례연구 1에 대한 오차의 실효치

Table 1 Root mean square error for case study 1

오차	오차의 실효치		
	동적 신경기	케환 신경망	혼돈 신경망
사례연구 1			
비선형 함수 근사화	0.087566	0.1637/2000	0.1621/2000

사례연구 2는 시스템에 비선형성이 있고 플랜트의 동적 특성이 변화하는 경우에 대해 동적 신경기의 제어 성능을 조사한다. 모형 비선형 시스템은 식(17)과 같은 차분방정식으로 표현된다.

$$y(k) = \sum_{i=1}^n \alpha_i y(k-i) + \sum_{j=1}^m \beta_j u(k-j) + f_{nonlin} \quad (17)$$

여기서 $y(k)$ 와 $u(k)$ 는 이산시간 플랜트의 출력과 입력을 나타내며 바운드(bound)되어있다. f_{nonlin} 는 비선형 함수로 식(18)과 같고 n 과 m 은 출력과 입력의 최대지연을 나타낸다.

$$f_{nonlin} = f[y(k-1), \dots, y(k-n), u(k), \dots, u(k-m)] \quad (18)$$

비선형 함수는 시간의 함수로 식(19)와 같고 이산시간 k 가 150이후부터 식(20)와 같이 비선형 함수의 형태가 변한다.

$$f_{nonlin} = 0.4 \left(e^{-(y(k-1)^2 + y(k-2)^2)} + \sqrt{u(k)^2 + u(k-1)^2 + u(k-2)^2} \right) \quad (19)$$

$$f_{nonlin} = \frac{0.4 - 0.4 \cos(7\pi(y(k-1)^2 + y(k-2)^2)) + e^{-u(k)}}{4 + u(k-1)^2 + u(k-2)^2} \quad (20)$$

플랜트의 전달함수는 시간에 따라 변하고 식(21)과 같다.

$$\begin{aligned} \beta_{ff} &= [1.2, 1.0, 0.8] \quad \alpha_{fb} = [1.3, 0.9, 0.7] \quad , 0 < k \leq 200 \\ \beta_{ff} &= [1.2, 1.0, 1.4] \quad \alpha_{fb} = [1.3, 0.9, 1.3] \quad , 200 < k \leq 350 \\ \beta_{ff} &= [1.2, 1.0, 1.4, 0.2] \quad \alpha_{fb} = [1.5, 0.9, 1.3, 0.5] \quad , 350 < k \leq 500 \end{aligned} \quad (21)$$

여기서 β_{ff} 는 플랜트의 2차 전달함수의 분자를 나타내는 계수이고 α_{fb} 는 분모의 계수를 나타내고 k 가 350이후에 플랜트의 차수가 2차에서 3차로 변한다. 동적 신경제어기는 비선형 시스템의 출력이 기준 모형의 출력을 추종한다. 기준 모형의 입·출력 관계식은 식(22-23)과 같다.

$$r(k) = 0.3 \left[\sin\left(\frac{\pi}{2400}d\right) + \sin\left(\frac{\pi}{628}d\right) + \sin^2\left(\frac{\pi}{418}d\right) + \cos^2\left(\frac{\pi}{314}d\right) + \sin\left(\frac{\pi}{250}d\right) + \sin\left(\frac{\pi}{200}d\right) \right] \quad (22)$$

$$y_d(k) = 0.4y_d(k-1) + 0.6r(k) \quad (23)$$

단, $d=100k$

비선형 시스템의 출력은 원하는 기준 모형 출력과 함께 그림 6과 같은 학습과 제어 알고리즘에 의해 동적 신경기의 학습 변수를 조절한다. 왜선 부분은 비선형 시스템과 이산시간 플랜트와 플랜트의 입출력으로 구성된 비선형 함수를 나타낸다. 동적 신경기는 그림 6에서 동적 인자 프로세서와 시그모이드함수로 나타내었다. 표 2는 학습시 사용한 적응변수의 초기치, 학습율, 그리고 활성화함수와 문턱신호를 나타낸다.

입력이 복잡한 경우에 신경의 상태가 수렴하는 가를 조사하기 위해 샘플링의 주기를 이산시간 변수의 100배로 하였다. 비선형 함수의 비선형성과 플랜트의 차수가 변화되는 모

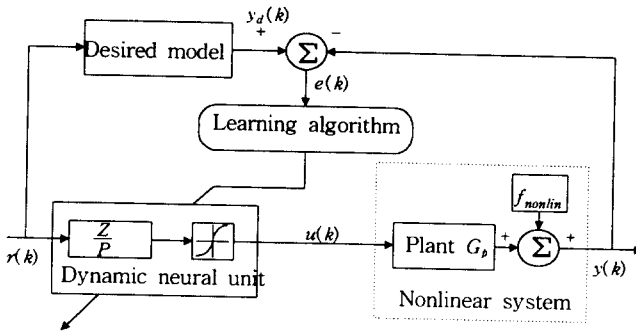


그림 6 동적 신경기를 이용한 비선형 시스템 제어구조
 Fig. 6 Control scheme for an unknown nonlinear system using dynamic neural unit

표 2 사례연구 2에서 사용한 변수와 초기치
 Table 2 Parameters and initial values used in case study2

Method	Proposed dynamic neural unit
Case study 2	
Nonlinear system control (1-1-1)	$\theta = 0.2, \eta_x = 0.6, \eta_z = 0.4, \eta_y = 0.4$ $g_s = 0.6$ $z[0] = 0.2, z[1] = 0.2, z[2] = 0$ $\mu[0] = 1.0, \mu[1] = 0.2, \mu[2] = 0$ $g_1 = 0.8, g_2 = \sqrt{\mu[2] - g_1^2}$ $g_3 = 0.3$ $g_4 = (0.5z[1] - z[0]g_1)/g_3$
Nonlinear system control (2-2-1)	$g_{1e} = 0.6, \theta_e = 0.2, \eta_{e1} = 0.6, \eta_{e2} = 0.1, \eta_{e3} = 0.1, \eta_{e4} = 0.1$ $g_n = 0.6, \theta_n = 0.2, \eta_{n1} = 0.6, \eta_{n2} = 0.4, \eta_{n3} = 0.4, \eta_{n4} = 0.1$ $W_{e1} = W_{e2} = 0.2, W_{e3} = W_{e4} = 0.5, V_1 = V_2 = 0.5, \eta_w = 0.2, \eta_v = 0$ $z[0] = 0.2, z[1] = 0.2, z[2] = 0$ $\mu[0] = 1.0, \mu[1] = 0.2, \mu[2] = 0$ $g_1 = \mu[1]/2, g_2 = \sqrt{\mu[2] - g_1^2}$ $g_{d1} = 0.6, g_{d2} = 0.6$ $g_4 = (0.5z[1] - z[0]g_1)/g_3$
Nonlinear system control (2-2-1-1)	$g_{1e} = 0.8, g_n = 0.8, g_{2e} = 0.8, g_{2n} = 0.8$ $\theta_e = .1, \eta_{e1} = .6, \eta_{e2} = .1, \eta_{e3} = .1, \eta_{e4} = 0, g_{d1} = .6$ $\theta_n = .2, \eta_{n1} = .6, \eta_{n2} = .2, \eta_{n3} = .2, \eta_{n4} = 0, g_{d2} = .6$ $\theta_{2e} = .1, \eta_{2e1} = .6, \eta_{2e2} = .1, \eta_{2e3} = .1, \eta_{2e4} = 0, g_{2d1} = .6$ $\theta_{2n} = .2, \eta_{2n1} = .6, \eta_{2n2} = .1, \eta_{2n3} = .1$ $W_{e1} = W_{e2} = .5, W_{e3} = W_{e4} = 1, W_{2e1} = W_{2e2} = .5, V_{11} = V_{12} = .5$ $W_{n1} = W_{n2} = 1, W_{2n1} = W_{2n2} = 0, \eta_w = .2, \eta_v = .3, \eta_v = .4$
Nonlinear system control (2-2-2-1)	$g_{1e} = 0.8, g_n = 0.8, g_{2e} = 0.8, g_{2n} = 0.8$ $\theta_e = .1, \eta_{e1} = .6, \eta_{e2} = .1, \eta_{e3} = .1, \eta_{e4} = 0, \eta_{2e1} = 0$ $\theta_n = .2, \eta_{n1} = .6, \eta_{n2} = .2, \eta_{n3} = .2, \eta_{n4} = 0, g_{d1} = 0$ $\theta_{2e} = .1, \eta_{2e1} = .6, \eta_{2e2} = .1, \eta_{2e3} = .1$ $\theta_{2n} = .2, \eta_{2n1} = .6, \eta_{2n2} = .1, \eta_{2n3} = .1$ $W_{e1} = W_{e2} = .5, W_{e3} = W_{e4} = 1, W_{2e1} = W_{2e2} = .5$ $V_{11} = V_{12} = .5, g_{2d1} = .6, g_{2d2} = 0.6, g_{d1} = .6$ $W_{2e1} = W_{2e2} = 1, \eta_{w1} = .3, \eta_{w2} = .4, \eta_{2e3} = 0$ $W_{n1} = W_{n2} = 1, W_{2n1} = W_{2n2} = 0, \eta_w = .2, g_n = .6$

형 제어 시스템을 가지고 사례연구한 결과를 그림 8 - 그림 11에 나타내었다.

동적 신경기는 한 개의 입력과 한 개의 출력으로 구성되는 전송요소를 가지고 있고 이 전송 요소를 병렬로 연결하면 동적 신경군이 된다[3]. 동적 신경기와 동적 신경군(neural processor)은 동적 신경망의 기본 요소로 연결방식에 따라 신경망의 구조가 결정된다. 예로 입력노드가 2개, 첫 번째 중간층 노드가 2개, 두 번째 중간층 노드가 1개, 그리고 출력층 노드가 1개인 동적 신경망은 1개 동적 신경군(중간층 1)과 1

개의 동적 신경기(중간층 2)로 구성된다. 입력층은 입력패턴 즉 입력 자료정보만 포함하며 이 정보는 중간층의 입력으로 사용되며 출력층은 중간층의 선형 합으로 출력정보를 가진다. 동적 신경망의 기본 요소로 학습한 결과를 그림 9와 그림 10에 나타내었다. 동적 신경기는 우수한 학습능력을 가지고 있고 실시간 처리가 가능하다. 반면에 동적 신경군은 출력정보를 입력으로 사용할 수 있고 입력이 복수개인 MIMO 시스템에 주로 사용된다[3].

사례연구 2에서는 신경망 구조가 (1-1-1)에서 학습한 후 입력층과 중간층의 노드를 각각 한 개씩 증가시켜서 신경기의 직렬구조와 병렬구조시의 학습결과를 비교하였고 다시 중간층의 층을 증가시켜 학습하여 층의 증가시의 학습능력을 살펴보았다. 결과그림으로부터 동적 신경기가 동적 신경군보다 오차가 적고 층의 증가와 노드의 증가시 오차가 커지는 경향이 있음을 알 수 있었다. 입력노드가 두 개이고 첫 번째 중간층 노드가 두 개, 두 번째 중간층 노드가 한 개, 출력층 노드가 한 개인 신경망에서 학습한 결과와 두 번째 중간층의 노드 수를 하나 추가하여 학습한 결과를 그림 10과 그림 11에 나타내었다. 사례연구한 결과, 출력의 심한 변화에 따라 처음부분과 중간부분에 원하는 출력을 추종함에 있어 다소 오차가 있음을 알 수 있었다. 입력수가 많을수록 지연시간이 생기고 층의 수가 많을수록 층의 전체 수렴성이 부분 최소점으로 빠지는 경향이 있음을 알 수 있었다.

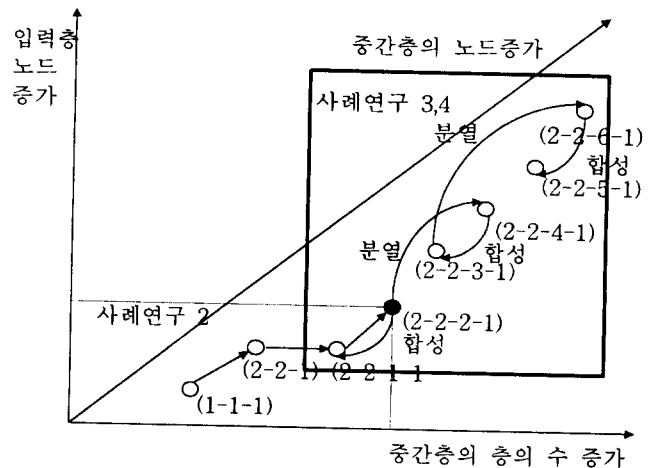


그림 7 중간층의 노드와 중간층의 층의 분열과 합성과정
 Fig. 7 Multiplication and combination process for node and layer of hidden layer

따라서 층의 연결 방법에 따라 출력오차를 줄일 수 있으므로 층의 선택을 해석적으로 결정해야 한다. 즉 신경망 구조가 정해지면 층의 노드 수를 적절히 선택할 수 있어야 한다. 본 논문에서는 층의 분열과 합성 알고리즘을 통해 원하는 결과를 얻을 수 있었다. 그림 7은 사례연구시 층의 변화와 노드의 변화과정을 나타낸다.

사례연구 3에서는 기본 신경망을 2-2-2-1으로 하여 신경망의 구조를 분열과 합성을 통해 중간층의 노드수를 생성하여 노드의 영향을 살펴보고 신경망 구조부터 오차의 실효치를 구해 신경망을 평가하였다. 그 결과를 그림 12에 나타내었고

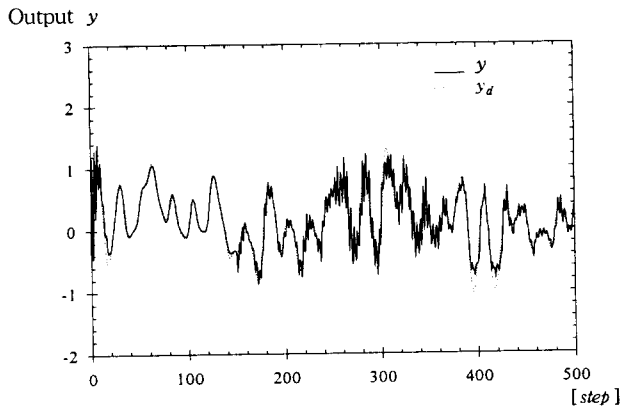


그림 8 동적 신경제어기를 이용한 시스템 파라미터의 변동과 차수변화에 따른 비선형 시스템 제어 (1-1-1)
Fig. 8 Control for nonlinear control system with parameter changes and structural disturbance using dynamic neural controller (1-1-1)

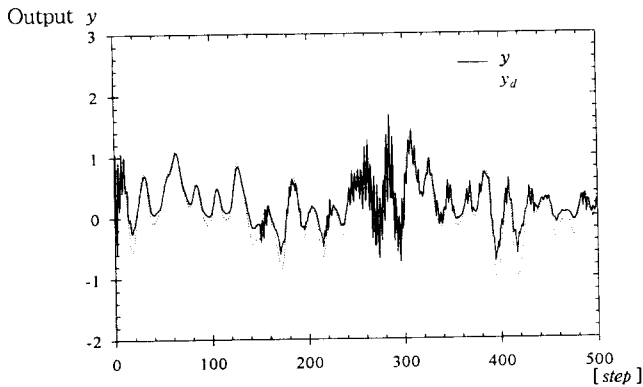


그림 9 동적 신경제어기를 이용한 시스템 파라미터의 변동과 차수 변화에 따른 비선형 시스템 제어 (2-2-1)
Fig. 9 Control for nonlinear control system with parameter changes and structural disturbance using dynamic neural controller (2-2-1)

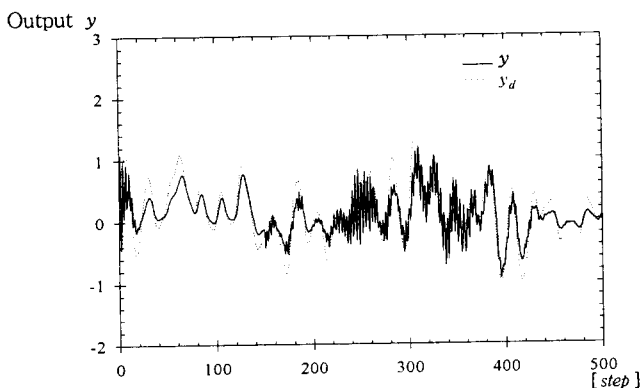


그림 10 동적 신경제어기를 이용한 시스템 파라미터의 변동과 차수변화에 따른 비선형 시스템 제어 (2-2-1-1)
Fig. 10 Control for nonlinear system with parameter changes and structural disturbance using dynamic neural controller (2-2-1-1)

노드생성과정을 표 3에 도시하였다. 분열과 합성을 적용한 경우에 오차의 실효치가 감소됨을 알 수 있었다. 오차의 실효치는 임의로 선택한 기본 신경망(2-2-2-1) 경우에 비해 오차의 실효치가 1.14배 감소됨을 표 3을 통해 알 수 있었다.

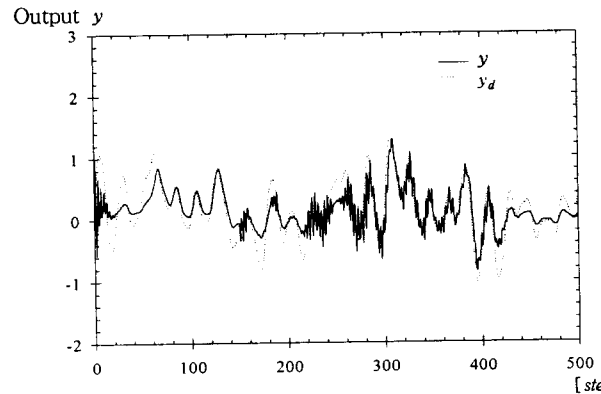


그림 11 동적 신경제어기를 이용한 시스템 파라미터의 변동과 차수변화에 따른 비선형 시스템 제어 (2-2-2-1).
Fig. 11 Control for nonlinear system with parameter changes and structural disturbance using dynamic neural controller (2-2-2-1)

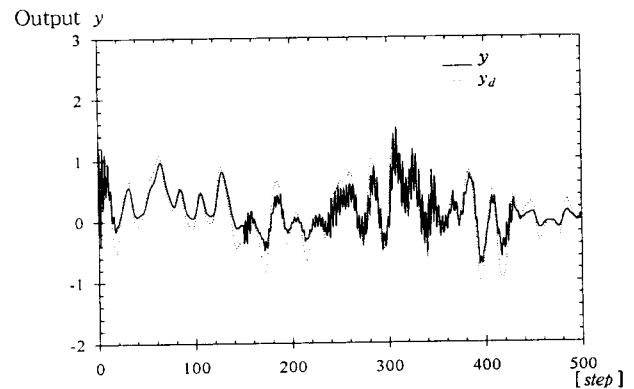


그림 12 분열과 합성에 의한 시스템 파라미터의 변동과 2차에서 3차로 차수변화에 따른 비선형 시스템 제어 (2-2-3-1)
Fig. 12 Control for nonlinear system with parameter changes and structure change to the third order by multiplication and combination(2-2-3-1)

본 논문에서는 중간층이 두 개인 4층 신경망을 기본 구조로 하여 학습을 시작하였고 두 번째 중간층의 노드 수를 분열과 합성 알고리즘을 사용하여 최적의 노드 수를 결정하였다. 입력층의 노드가 2개, 중간층의 노드가 각각 2개, 2개이고 출력층 노드가 1개인 동적 신경망에서 출력오차와 두 번째 중간층의 노드간의 상관계수로부터 노드 수를 증가하는 전방향 분열과 기본 구조에서 합성조건을 만족하는 경우에 노드 수를 줄이는 역방향 합성을 동시에 하였다. 첫 번째로 기본 신경망을 학습하여 오차의 실효치가 구해지면 분열조건인 학습한계 설정치, 학습한계 극한치 조건을 만족하고 중간층 노드 간의 상관계수가 설정치 이상이면 분열을 시작한다. 이 과정에서 두 번째 노드의 출력 분산이 첫 번째 노

드보다 적어서 첫 번째 노드로 합성하여 학습을 진행하였다. 학습과정에서 구한 결과를 표 3에 나타내었고 역방향 합성결과를 그림 10에 나타내었다.

표 3 사례연구 3에 대한 노드 생성과정

Table 3 Node creation step for case study 3

정보	분열과 합성 정보				
	분열/합성	오차의 실효치	합성노드 /상관 계수	직선의 기울기와 절편 / 표준편차	학습변수 설정치
동적 신경망 구조					
2-2-1-1	합성	0.2649	(1,2)-(1)	$V_1 = -0.175 V_2 + 0.058$	학습종료 설정치 (0.1728)
2-2-2-1	기본 구조	0.2726	0.884631	$s_1 = 0.1259, s_2 = 0.024$	
2-2-4-1	분열	0.3103	0.369243 0.780787	$s_3 = 0.0064, s_4 = 0.005$	학습한계 극한값 (0.2258)
2-2-3-1	합성	0.2391	(3,4)-(3)	$V_3 = 0.717 V_4 + 0.075$	
2-2-6-1	분열	0.2554	0.843995 0.005731 0.022477	$s_1 = 0.0519, s_2 = 0.013$	상관계수 설정치 (0.75)
2-2-5-1	합성	0.3578	(1,2)-(1)	$V_1 = 0.290 V_2 + 0.077$	

조는 플랜트가 2차에서 1차로 변하는 경우에도 동일한 최적 표 4 사례연구 4에 대한 노드 생성과정

Table 4 Node creation step for case study 4

정보	분열과 합성 정보				
	분열/합성	오차의 실효치	합성노드 /상관 계수	직선의 기울기와 절편 / 표준편차	학습변수 설정치
동적 신경망 구조					
2-2-1-1	합성	0.2729	(1,2)-(1)	$V_1 = 0.238 V_2 - 0.080$	학습종료 설정치 (0.1791)
2-2-2-1	기본 구조	0.2738	0.914280	$s_1 = 0.1511, s_2 = 0.0394$	
2-2-4-1	분열	0.2794	0.329246 0.838878	$s_3 = 0.0082, s_4 = 0.0062$	학습한계 극한값 (0.2247)
2-2-3-1	합성	0.2473	(3,4)-(3)	$V_3 = 0.717 V_4 + 0.075$	
2-2-6-1	분열	0.3796	0.696359 0.021138 0.026609	$s_1 = 0.0571, s_2 = 0.0105$	상관계수 설정치 (0.65)
2-2-5-1	합성	0.5938	(1,2)-(1)	$V_1 = 0.102 V_2 + 0.066$	

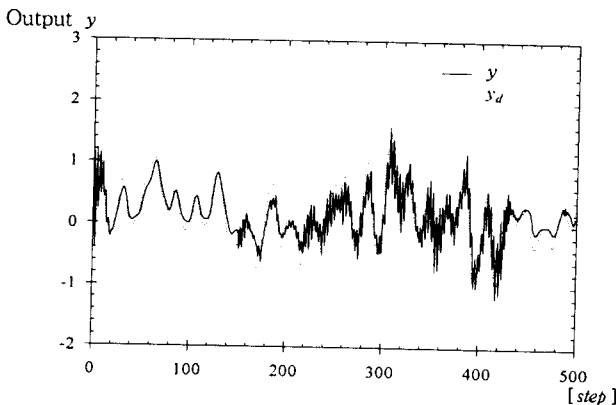


그림 13 분열과 합성에 의한 시스템 파라미터의 변동과 2차에서 1차로 차수변화에 따른 비선형 시스템 제어 (2-2-3-1)

Fig. 13 Control for nonlinear system with parameter changes and structure change to the first order by multiplication and combination (2-2-3-1)

전방향 합성은 먼저 분열후 합성하여 중간층의 노드를 4개로 한 후 오차의 실효치를 구하여 합성조건을 만족하면 합성한다. 이 과정에서 3번째 노드, 4번째 노드가 상관계수의 설정치를 만족하여 3번째 노드로 합성하였다. 합성후 다시 분열하여 노드 수가 6개로 되었고 다시 합성조건을 통해 5개로 되어 학습하였다. 본 논문에서는 오차의 실효치가 학습종료 설정치를 만족하지 못해 분열과 합성이 계속되었고 그 결과를 관측하여 최적의 신경망 구조 2-2-3-1 결정하였다. 이 구

오차 실효치

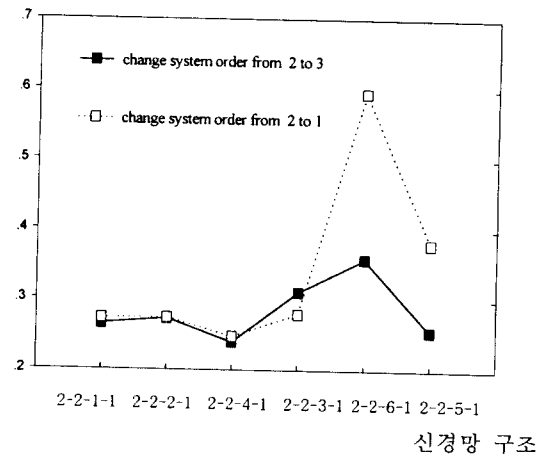


그림 14 사례연구3과 4에 대한 오차 실효치

Fig. 14 Root mean square for case study 3 and 4

구조로 결정되었다. 기본 신경망 구조에서 분열과 합성을 통해 구한 최적 구조 신경망은 상대적으로 다른 신경구조에 비해 오차의 실효치가 적었고 학습종료 설정치와 학습한계 극한값보다 더 큰 값을 가짐을 알 수 있었다. 동적 신경망의 층의 분열과 합성에서 얻은 구조로 활성화 함수의 기울기를 변화시킬 경우 동적 신경망과 신경군 사이의 오차 실효치 0.209067을 얻을 수 있다. 입력이 복수개인 비선형 시스템에 적용하면 보다 최적화 되는 결과를 가질 수 있다. 동적 신경망의 합성과 분열시에 활성화함수의 크기와 기울기 그리고 학습변수는 모든 층의 노드에 동일한 값을 가진다. 분열과

합성과정에서 살아남은 노드는 학습하여 얻은 연결계수와 적응된 값을 유지하고 있고 새로 추가되는 노드는 처음에 결정된 초기 학습계수와 활성화함수의 기울기를 가지고 학습하게 된다. 학습종료 설정치는 동적 신경기에서 얻은 오차의 실효치로 설정하였고 학습한계 극한값은 동적 신경군의 오차 실효치로 하였다. 사전에 학습한 동적 신경기와 신경군의 학습 결과를 본 논문에서는 이용하였기 때문에 학습 중간과정에서의 상관관계 시간 변화정보는 필요가 없었다. 실제로 상관관계의 변화량은 학습과정중 몇 번 있었으나 그 값은 아주 작고 거의 변화가 없었다. 분열과 합성과정에서 상관계수와 오차의 실효치로 학습한계를 판단하였고 두 노드간의 출력 분산으로 합성을 행하였다.

5. 결 론

본 논문에서는 새로운 구조의 동적 신경모형을 사용하여 비선형 이산시간 시스템의 식별과 제어에 적용하여 학습과 제어 알고리즘에 의해 시스템의 출력을 식별하는 제어기를 설계하였다. 비선형 시스템을 식별하는 제한 신경기는 여러 형태가 사용될 수 있으나 학습처리 시간이 빠르면서 안정하게 기준 출력을 식별하며 비선형 시스템을 제어할 수 있는 동적 신경기를 사용하였다.

본 논문에서는 중간층의 분열과 합성에 의해 신경망의 구조를 선택하여 비선형 시스템을 식별하는 제어기를 설계하였다.

제안된 동적 신경모형의 타당성과 중간층 노드 수의 분열과 합성 알고리즘의 유용성을 입증하기 위해 사례연구로 비선형 모형 시스템 적용하였다. 본 논문의 사례연구에서 얻은 결과를 요약하면 다음과 같다.

- (1) 사용된 동적 신경기는 기준 출력을 잘 추종하여 원하는 비선형 함수의 파형을 근사화하였다.
- (2) 사용된 동적 신경망은 비선형에 적용하여 모형 시스템의 출력을 추종하도록 제어하였고 사례연구를 통해 타당성을 입증하였다.
- (3) 하나 이상의 중간층을 가진 동적 신경망에서 중간층 노드의 전방향 분열과 역방향 합성이 가능한 알고리즘을 제시하였다.
- (4) 동적 신경망의 층의 분열과 합성을 통해 실제 시스템의 동적 특성을 정확하게 식별하도록 제어하였다.

본 연구는 한국전력 지원 97년도 단기과제(기초전력공동연구소 과리번호 97-005) 연구비 지원으로 수행되었음을 알리며 관계 당국에 감사드립니다.

참 고 문 헌

[1] M. M. Gupta and D. H. Rao, "Dynamic Neural Units With Applications to the Control of Unknown Nonlinear System." *The Journal of Intelligent and Fuzzy Systems*, Vol. 1, pp.73-92, 1993.

[2] 서보혁, 박성욱, "상태변수 모형을 가진 동적 신경망을 이용한 비선형 시스템의 제어기 설계," 대한전기학회 논문지, 제 45권 12호, pp. 1742-1753, 1996.

[3] 서보혁, 박성욱, "동적 신경망을 이용한 비선형 다변수 시스템의 제어기 설계," 대한전기학회논문지, 제 46권 12호, pp. 1818-1824, 1997.

[4] Murata, Yoshizawa and Amari, "Network Information Criterion-Determining the Number of Hidden Units for an Artificial Neural Network Model," *IEEE Trans, Neural Networks*, Vol. 5, No. 6, pp. 865-871, 1994.

[5] D. B. Fogel, "An Information Criterion for Optimal Neural Network Selection," *IEEE Trans, Neural Networks*, Vol. 2, No. 5, pp. 490-497, 1991.

[6] J. N. Hwang, S. S. You, S. R. Lay and I. C. Jou, "The Cascade-Correlation Learning: A Projection Pursuit Learning Perspective," *IEEE Trans, Neural Networks*, Vol. 7, No. 2, pp. 278-288, 1996.

[7] Axel Doering, Mirosław Galicki and Herbert Witte, "Structure Optimization of Neural Networks with the A*-Algorithm," *IEEE Trans, Neural Networks*, Vol. 8, No. 6, pp.1434-1445, 1997.

[8] T. Ash, "Dynamic Node Creation in Backpropagation Networks," ICS Report 8901, UC San Diego, 1989.

[9] 増田, 藤井, 池谷, 相部, "隠れユニットの合成による階層型ニューラルネットワークのコンパクト構造化," 計測自動制御學論, 28, No. 4, 1992.

[10] K. A. Yeomans, *Introducing statistics*, in *Statistics for the social scientist: volumn one*, pp. 166-207, Penguin Books Ltd, 1980.

저 자 소 개



박 성 욱 (朴省昱)

1964년 11월 16일생. 1987년 경북대 공대 전기공학과 졸업. 1989년 동 대학원 전기공학과 졸업(석사). 1996년 동 대학원 전기공학과 졸업(공학박). 1998.2.~1998.12. 영국 버밍엄 대학 Post-Doc. 연수 1992년 ~ 현

재 구미1대학 전기과 조교수.
Tel : 연구실 (0546) 440-1203
E-mail : psw@kumi.ac.kr



이재관 (李載寬)

1967년 2월 17일생. 1990년 경북대 공대
전기공학과 졸업. 1993년 동 대학원 전기
공학과 졸업(석사). 1998년 일본 동북대학
대학원 공학연구과 전기·통신공학전공 졸
업(박사). 1998 ~현재 일본 동북대학 대학

원 공학연구과 전기·통신공학전공 조수

Tel : 81-22-217-7075

E-mail : lee@abe.ecei.tohoku.ac.jp

서보혁 (徐輔赫)

전기학회논문지 제48A권 제1호 참조

Tel : 053-950-5604

E-mail : bhsuh@bh.kyungpook.ac.kr