

고차 뉴런을 이용한 KOHONEN의 자기 조직화 맵

정 종 수, 하기오라 마사후미
(데이터박스(주), 게이오 대학교(일본))

Supervised Kohonen Feature Map Using Higher Order Neuron

Jong-Soo JUNG, Massfume HAGIWARA
(Databox Co. Ltd., Keio University(Japan))

Abstract - 본 논문은 교사 있는 학습기의 Kohonen Feature Map에 고차 뉴런을 도입, 고차 뉴런을 이용한 Kohonen의 자기 조직화 맵을 제안한다. 일반적인 Kohonen Feature Map의 특징은 입력신호를 받아 출력 면(Kohonen Feature Map) 내의 특정한 위치 주위에 집중하는 메커니즘으로 즉, 국소집중 반응을 구하는 구조이다. 본 논문에서는 종래형의 Kohonen Feature Map의 특징을 보유하며 교사 있는 학습기의 Kohonen Feature Map에 고차 뉴런을 도입하여 국소 집중반응 및 특징 추출이 용이하도록 네트워크 구조를 개선한 것이다. 특히, 일차 뉴런의 문제점인 비선형 분리 문제에 대하여 교사 있는 학습기의 Kohonen Feature Map의 입력층에 고차 뉴런을 도입함으로써 비선형 분리 가능한 형태의 네트워크 구조로 형성하였다. 그러나, 일반적인 고차 뉴런의 문제점을 보완하기 위해 본 논문에서는 오직 2차 뉴런만을 생성하였으며 중복되는 뉴런을 최대한 억제하였다. 본 제안 모델의 특성을 살펴보기 위해 XOR문제와 20개의 Alphabet을 식별하는 패턴인식 시뮬레이션을 했으며, 본 제안 모델의 범화 능력을 알아보기 위하여 Mirror Symmetry를 사용하여 계산기 시뮬레이션을 했다. 그 결과, 본 제안 모델이 종래형의 네트워크 구조보다 뛰어난 인식률을 얻을 수 있었다.

리보다는 비선형 분리 문제가 더 요구된다. 따라서, 이러한 문제를 해결하기 위해 일차 다층형(Single Multi Layer)과 고차 단층형(Multiple Single Layer)의 네트워크가 제안되었다. 고차 뉴럴 네트워크의 경우, 간단한 네트워크 구조를 형성하고 있으며, 고차 뉴런에 의한 비선형 분리가 가능한 특성을 갖고 있다. 이러한 특성을 계층형 네트워크에 적용하여 위치변형, 확대 및 축소에도 적합한 네트워크 구조임을 알 수 있다 [7][8][10][11]. 이러한 특성들을 이용하여 본 논문에서는 고차 뉴런을 Kohonen의 자기 조직화 맵에 적용하였다. 그 결과, 비선형 분리 문제에 있어서 본 제안 모델이 종래의 모델 보다 뛰어난 성능을 발휘하여 국소(局所) 집중반응을 정확하게 추출할 수 있었다. 또한, 이러한 비선형 분리에 의한 복잡한 패턴 분리가 가능하여 여러 분야에 응용이 가능한 사실 또한 알 수 있었다. 그러나, 고차 뉴런을 이용함에 있어 고차 뉴런의 결점을 생각하지 않을 수 없다. 따라서, 본 논문에서는 최적의 뉴런만을 형성하기 위해 2차 뉴런을 도입하였다[11]. 본 모델은 교사 없는 Kohonen Feature Map의 알고리즘을 기초로 하여 교사 있는 학습으로 구축된 Kohonen Feature Map에 고차 뉴런을 적용하여 패턴 분류 문제 등을 중심으로 시뮬레이션을 하였다.

1. 서 론

생물의 신경 조직 및 뇌의 정보처리 방식을 도용하여 인간과 같은 유연성이 있는 정보처리를 실현하고자 하는 시도가 인공적 신경회로망(뉴럴 네트워크)을 이용하여 활발하게 연구되고 있다. 이러한, 뉴럴 네트워크의 특징으로는 병렬 분산처리, 학습능력 및 범화능력이 있으며, 뉴럴 네트워크의 특성에 의해 교사 있는 학습 모델과 교사 없는 학습으로 분리된다. 교사 있는 학습으로서 Back Propagation이 대표적인 모델이며, 여러 분야에서 뛰어난 특성을 발휘하고 있다[1].

한편, 교사 없는 학습으로는 대표적인 모델이 Topological Map를 도입한 Kohonen Feature Map 이 있다[2][3][4]. Kohonen Feature Map는 입력 데이터의 유사 정도를 동적으로 찾아내어 유사한 입력들을 네트워크 상의 가까운 곳에 배치하는 특징이 있다. 즉, 입력 데이터에 대하여 선택적으로 반응하는 네트워크 구조이다. 또한, Kohonen Feature Map는 맵 내의 상호결합이 불필요함과 동시에 미분 연산이 필요 없어 간결한 네트워크 구성이 된다. 따라서, 이러한 특징을 이용하여 Kohonen Feature Map에 교사 신호를 도입한 교사 있는 학습기로 Kohonen Feature Map이 H. Ichiki, M. Hagiwara, 및 M. Nagakagawa 등에 의해 제안되어 패턴인식, 연상기억 등에서 뛰어난 특징이 소개되었다[5][6]. 그러나, 기존의 모델은 일차 뉴런을 갖고 있는 단순구조로 형성되어 있어, 일차 뉴런이 갖고있는 일반적 특징인 선형분리만이 가능한 구조이다. 그러나, 실질적인 패턴인식의 문제에 있어서는 선형 분

2. 본 론

2.1 교사 있는 Kohonen Feature Map의 구조

Kohonen Feature Map는 입력층과 출력층의 2차원 뉴럴 시트(Neural Sheet)의 구조로 되어 있으며, 동작은 입력 벡터에 대하여 2차원의 뉴럴 시트 상에서 거리 관계를 구하여 학습하며 이에 뉴런이 각각 반응하는 특징을 갖고 있는 즉, 토풀로지 맵을 형성하는 네트워크이다[2]. 따라서, 교사 있는 학습 기계로서 Kohonen Feature Map는 입력 벡터와 교사 벡터를 하나의 학습 벡터로 하여 학습하는 방식이다[5][6]. 교사 있는 학습기의 Kohonen Feature Map의 구조는 그림 1과 같으며 교사 없는 뉴럴 네트워크의 대표적인 모델인 Kohonen Feature Map의 특징을 유지, 교사 있는 학습기로 변형하기 위한 입력 벡터 I 는

$$I = (x_1, x_2, \dots, x_n) \tag{1}$$

이다. 여기서, x_n 은 입력 벡터이고, n 은 입력 벡터 수이다. 입력 벡터에 대응하는 교사 벡터 T 는

$$T = (y_1, y_2, \dots, y_m) \tag{2}$$

이고 여기서, y_m 은 교사 벡터이고, m 은 교사 벡터 수이다. 네트워크 상에서 학습을 하기 위해서는 입력 벡터와 교사 벡터를 하나로 결합된 학습 벡터 X_n 은

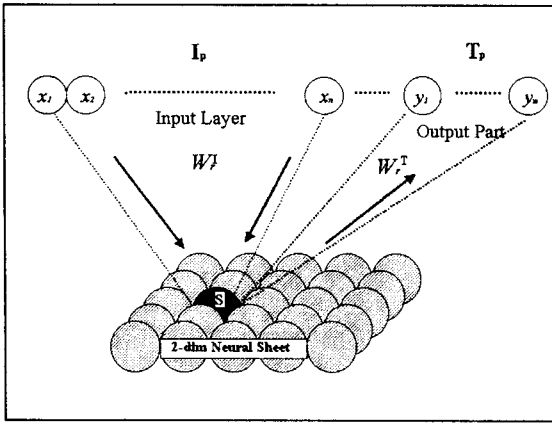


그림 1. 교사 있는 학습기의 Kohonen Feature Map.

$$X_n = \alpha \begin{bmatrix} I \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ T \end{bmatrix} \quad (3)$$

이다. 단, 입력 벡터의 크기에 의해 맵이 형성 되도록 입력 벡터에 대한 가중치인 $\alpha \gg 1$ 로 한다. 그러므로, 입력 벡터의 가중치에 의해 교사 벡터의 연결강도가 투형 되는 형태의 맵이 형성된다. 따라서, 연결강도는 다음과 같이 분리하여 표현된다.

$$W_r = \begin{bmatrix} W_r^I \\ W_r^T \end{bmatrix} \quad (4)$$

여기서, r 은 맵 상에서 뉴런의 위치를 나타내는 2차원 위치벡터이고, W_r^I 는 입력 벡터에 해당하는 연결강도이며, W_r^T 는 교사벡터에 대한 연결강도이다.

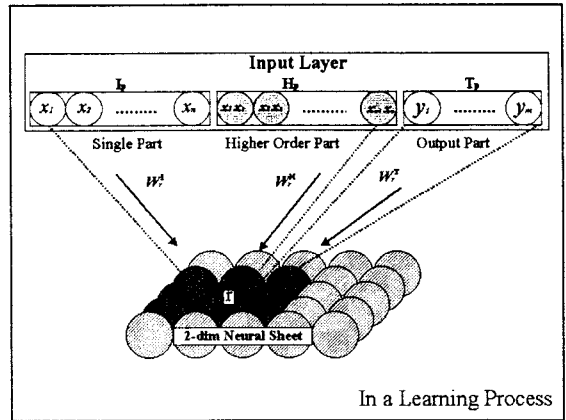
2.2 고차 뉴런을 이용한 KOHONEN 자기 조직화 맵의 네트워크 구조 및 알고리즘: 제안 모델

본 논문에 이용된 고차 뉴런의 특징은 일차 뉴런이 갖고 있는 구조적 결합, 즉 선형분리만이 가능함으로 실질적인 패턴인식의 문제에 있어 비선형 분리 문제에 대하여 많은 한계를 갖고 있다. 따라서, 이러한 문제를 해결하기 위해 간단한 네트워크 구조를 갖고 있는 특징과 고차 뉴런에 의해 비선형 분리가 원활히 이루어지는 특징을 갖고 있는 고차 뉴런을 이용한 Kohonen의 자기 조직화 맵을 제안하다. 본 제안 모델의 구조는 교사 있는 Kohonen Feature Map의 입력층에 고차 뉴런을 도입하는 방식으로 교사 뉴런은 (1)식의 입력 벡터에 대하여 자신의 입력 벡터를 대입하여 복합적 산출 방식으로 고차 뉴런의 입력 벡터 H 는

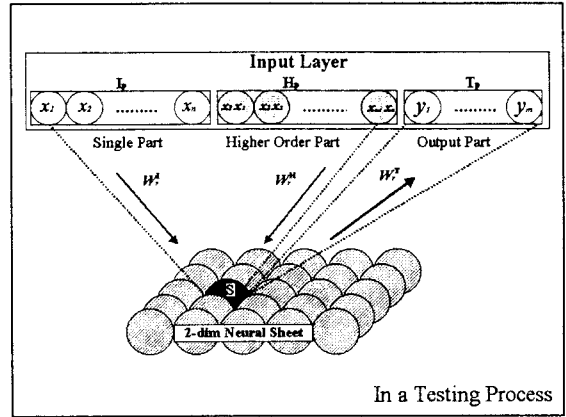
$$H = (x_1 \ x_2 \ x_1 \ x_3 \ \dots \ x_{n-1} \ x_n) \quad (5)$$

이다. 따라서, 고차 뉴런을 포함하고 있는 입력 벡터를 교사 있는 학습기로 학습하기 위해서 고차 뉴런을 포함하고 있는 입력 벡터와 교사 벡터를 하나로 형성한 학습 벡터 X_p 는

$$X_p = \alpha \begin{bmatrix} I \\ H \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T \end{bmatrix} \quad (6)$$



In a Learning Process



In a Testing Process

그림 2. 고차 뉴런을 이용한 Kohonen의 자기 조직화 맵.

이다. 또한, 교사 벡터에 있어 종래의 교사 있는 Kohonen Feature Map과 동일하게 입력 벡터의 가중치를 부가, 입력 벡터에 의해 착용되는 맵이 형성 되도록 한다. 따라서, 고차 뉴런을 이용한 Kohonen의 자기 조직화 맵 상의 연결강도 W_r^P 는 다음과 같다.

$$W_r^P = \begin{bmatrix} W_r^I \\ W_r^H \\ W_r^T \end{bmatrix} \quad (7)$$

여기서, r 은 맵 상에서 뉴런의 위치를 나타내는 2차원 위치 벡터이고, W_r^H 는 고차 뉴런 벡터에 해당하는 연결 강도이다. 본 제안 모델의 구조는 그림 2와 같다.

본 제안 모델의 학습은 입력 벡터 X_p 를 네트워크에 입력하여 출력을

$$y_r = X_p \cdot W_r^P \quad (8)$$

으로 계산하며, 식별기구로서는 2차 뉴런 시트 즉, 맵 상에서 가장 발화하고 있는 뉴런의 출력 y_{\max} 는

$$\begin{aligned} y_{\max} &= \max(\sum I W_n^I + \sum \sum I_{n-1} I_n W_n^H + \sum T W_n^T) \\ &= \max(\sum I W_n^I + \sum H W_n^H + \sum T W_n^T) \\ &= \max[X_p \cdot W_r^P] \end{aligned}$$

표 1. XOR 문제.

입력 벡터	교사 신호
$[-1, -1]$	-1
$[-1, +1]$	+1
$[+1, -1]$	+1
$[+1, +1]$	-1

$$= \max[y_r] \tag{9}$$

이다. 여기서 최대치를 나타내는 뉴런의 위치를 S라고 하며, 최대로 발화된 뉴런의 출력 값을 1로 하고 나머지는 0으로 한다. 그리고, 각 뉴런에 대한 연결강도는

$$W_n^{(new)} = W_n^{(old)} + \epsilon h_n (X_p \cdot W_n^{(old)}) \tag{10}$$

으로 변경되며 여기서, ϵ 는 결합강도의 변환 계수이고 h_n 에 의해 학습하는 연결강도의 범위가 정하여 진다. 여기서, h_n 는

$$h_n = \exp\left(-\frac{\|r - s\|^2}{\sigma^2}\right) \tag{11}$$

이며, σ 는 근접학습의 범위를 결정하는 파라메타 이다. σ 가 크면 클수록 학습 범위가 넓어진다. 통상 σ 는 초기단계에서는 넓은 범위에서 학습을 하며 점차 감소시킨다. 또한, 변경된 연결강도를 절대값화 한다. 상기 알고리즘을 반복함으로써 토폴로지 맵이 형성되며, 형성된 맵의 특징으로 입력 데이터간의 거리 관계가 정확하게 맵상에 이미지 화되어 보존된다. 또한, 빈번한 입력 값은 그렇지 않는 입력 값에 비하여 넓은 범위로 형성된다. 따라서, 고차 뉴런을 사용함으로써 국소 집중반응 및 특징 추출이 용이하여 기존의 일차 뉴런구조의 모델이 갖고 있는 문제점을 해결할 수 있었으며, 여러 응용분야에 적용이 가능하도록 학습 회수의 절감 및 인식률의 향상을 얻을 수 있었다.

2.3. 실험 및 결과분석

2.3.1 XOR 문제.

본 제안 모델의 동작을 조사하기 위해 비교적 간단한 문제이며, 선형분리 불가능한 표 1의 XOR의 문제를 가지고 시뮬레이션을 했다[1]. 고차 뉴런을 이용한 자기조직화 맵은 정규화 했으며, 고차 뉴런을 포함하고 있는 입력벡터와 교사 벡터는 각각 같은 2진수인 -1과 1을 가지고 시뮬레이션 했다. 자기 조직화 맵은 종래형과 비교하기 위해 5X5개의 2차원 시트 맵을 사용했으며, 입력 벡터의 가중치는 5.0, 결합강도의 변환 계수는 0.4, 근접학습 범위의 초기 값을 4.0에서 최종 값 0.4로 하여 시뮬레이션을 하였다. 학습 회수에 의한 학습 인식률

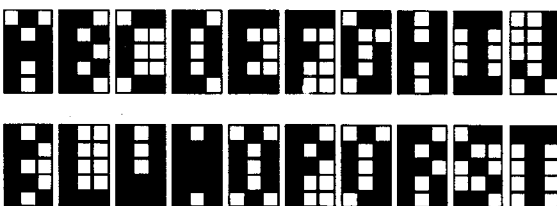


그림 4. 20 Alphabets 문자.

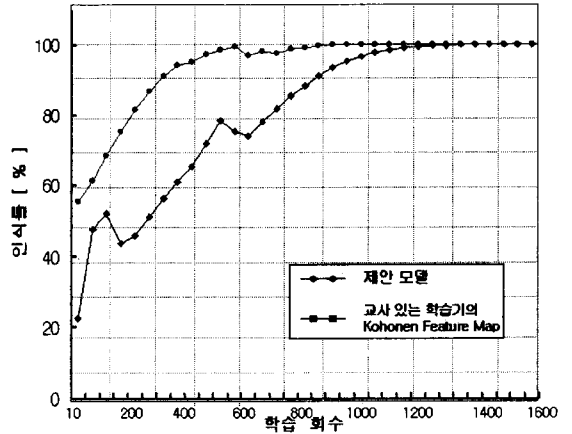


그림 3. XOR의 인식을 결과

은 그림 3과 같다. 이 결과에서 알 수 있듯이 고차 뉴런을 이용함으로써 학습에 따르는 인식에 있어 종래의 학습 회수를 삭감할 수 있었으며, 종래 모델의 학습 인식률보다 높은 인식률을 얻을 수 있었다. 이것은 국소 집중반응 및 특징 추출을 산출함에 있어 고차 뉴런이 일차 뉴런보다 월등한 성능을 발휘하고 있으며, 고차 뉴런을 도입함으로써 비선형 분리 능력이 한층 향상됨을 잘 보여주는 결과이다.

2.3.2 20 Alphabets 식별문제.

본 논문에 있어서, 15Pixel로 이루어진 20개의 Alphabet에 대한 식별을 목적으로 식별문제에 대한 시뮬레이션을 했다[6]. 상기 20개의 Alphabet는 가장 적은 픽셀로 이루어진 패턴으로 패턴 식별문제에 있어서 일반적으로 사용되는 패턴이다. 또한, 고차 뉴런의 생성은 본 논문에 기준한 2차 뉴런을 생성하였으며, (5)식에 따라 중복되는 뉴런의 생성을 억제했다. 특히, 종래형과 비교를 위해 조직화 맵의 크기를 5X5로 했으며 입력 벡터에 대한 가중치를 5.0으로 했다. 결합강도의 변환 계수는 0.6이고, 근접학습 범위의 초기 값을 4.0에서 최종 값 0.5를 가지고 시뮬레이션을 하였다. 입력

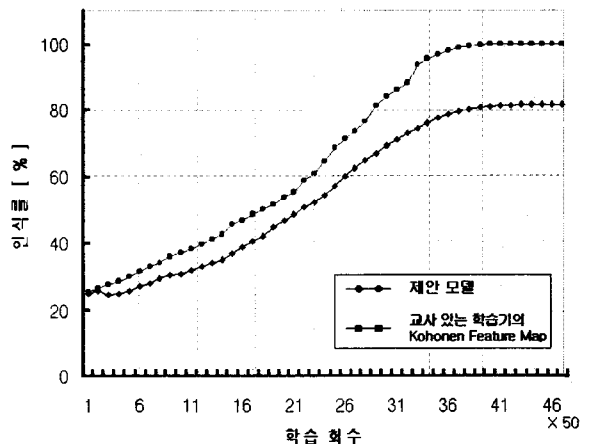


그림 5. 20 Alphabets 식별문제 결과.

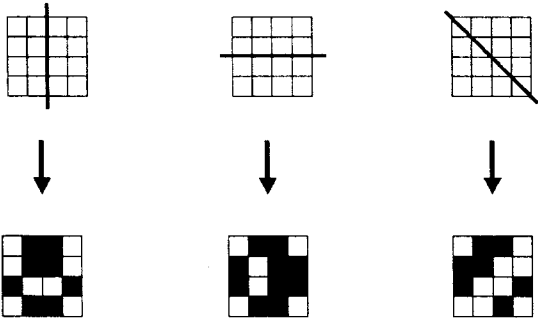


그림 6. 4X4 Mirror Symmetry 문제의 패턴 예.

벡터인 20개의 Alphabet는 그림4와 같다. 그림 5는 2000회에 이르는 학습 회수에 대한 종래형과 제안 모델의 인식률을 나타낸 것이다. 이 결과에서 알 수 있듯이 복잡한 패턴 인식문제에 대하여 고차 뉴런을 이용함이 보다 효율적이라는 것을 알 수 있다. 이것은 고차 뉴런을 사용함으로써 한층 비선형 분리 능력이 향상됨으로 복잡한 패턴 인식문제에 대하여 종래형 보다 높은 인식률을 얻을 수 있었다.

2.3.3 Mirror Symmetry 문제.

본 제안 모델의 범화 능력을 알아보기 위해 그림 6과 같은 4X4 Mirror Symmetry를 사용하여 시물레이션을 하였다[9]. 4X4 Mirror Symmetry 문제는 입력 벡터의 2진수 값의 대칭축을 학습 후, 인식하는 형태로 4X4 Mirror Symmetry 인식 카테고리에 대한 패턴 인식을 하는 문제로 네트워크의 응용력, 즉 범화 능력을 측정하는 좋은 예제이다. 그림 6에 있어 좌측 순으로 좌우대칭, 상하대칭 및 사면대칭에 관한 예제를 나타낸 것이다. 상기 대칭 패턴은 대략 1500개 정도 형성할 수 있으며 그중 일정한 비율의 학습 데이터를 선택하여 학습 후, 각 일정한 미학습 데이터 패턴에 대하여 대칭축을 인식하는 시물레이션을 하였다. 또한, 본 논문에서는 종래형과 제안 모델의 특성을 비교하기 위해 10X10의 2차원의 자기 조직화 맵을 사용하여 시물레이션을 했다. 여기서, 입력 벡터의 가중치는 5.0으로 하였으며, 결합 강도의 변환 계수는 0.4, 근접학습 범위는 초기 값을 4.0에서 최종 값 0.4를 가지고 시물레이션을 했다. 그림 7은 2000회 학습 후, 일정 분포의 미지 테스트 패턴을 네트워크에 대입하여 각각의 패턴에 대한 인식률의 곡선이다. 이 결과에서 알 수 있듯이, 고차 뉴런을 도입함으로 미지 대칭축에 대한 인식률이 종래형 보다 높다는 결과를 알 수 있다. 따라서, 본 제안 모델이 종래형 보다 높은 범화 능력을 갖고 있어 여러 응용 분야에 있어서 종래형 네트워크 구조 보다 본 제안 모델이 한층 적용 가능함을 보여준 결과라고 할 수 있다.

3. 결 론

본 제안 모델인 고차 뉴런을 이용한 Kohonen의 자기 조직화 맵을 교사 있는 학습기로 적용하여도 Kohonen 자기 조직화 맵의 특징인 교사 없는 학습의 기능 즉, 특징 추출의 능력을 유지 할 수 가 있다. 또한, 입력층에 고차 뉴런을 도입하여 초기 단계에서 비선형 분리가 가능하게 되어 복잡한 문제에 대하여 더욱더 대응 가능하도록 개선되었다. 본 제안 모델의 특성을 알아보기 위해서 XOR, 20개의 Alphabet 문제 및 4X4 Mirror Symmetry 문제를 사용하여 계산기 시물레이션을 했다. 그 결과, 종래형 보다 높은 인식률

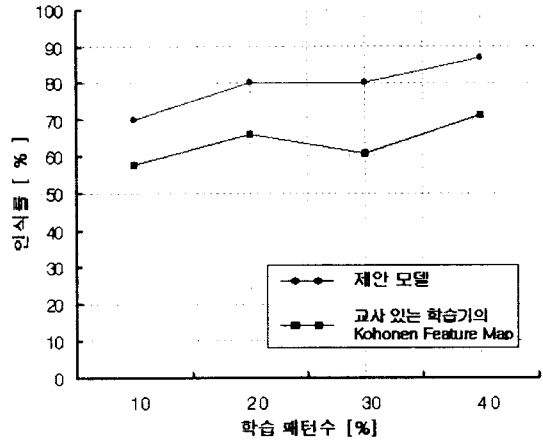


그림 7. 4X4 Mirror Symmetry 문제에 대한 인식률.

을 얻을 수 있었으며, 고차 뉴런에 의한 특징 추출이 한층 용이한 사실을 알 수 있었다. 그러나, 본 제안 모델의 문제점으로 지적되는 고차 뉴런에 의한 방대한 양의 결합강도를 어느 적정 수준에서 억제할 수 있을 것인가 하는 문제이다. 따라서, 자기 문제해결을 위해 자기 조직화 맵 상에서 공헌도가 적은 결합강도에 대하여 도태시키는 결합강도 삭감문제에 대하여 연구가 수행되어야 할 것이다.

(참 고 문 헌)

- [1] Remelhart and McClelland, "Parallel Distributed Processing", MIT, Press, 1986
- [2] T. Kohonen, "Self-Organized Formation of Topographically Correct Feature Map", Biol, Cybern, 43, pp.59-69, 1982
- [3] T. Kohonen, "Self-Organization and Associative Memory: second edition", Springer-Verlag, Press, 1988
- [4] T. Kohonen, "Self-Organizing Semantic Maps", Biol, Cybern, 61, pp.241-254, 1989.
- [5] H. Ichiki, M.Hagiwara and M. Nakagawa, "Self-Organizing Multi-Layer Semantic Maps", IJCNN, Vol.1, pp.357-360, 1991.
- [6] H. Ichiki, M.Hagiwara and M. Nakagawa, "Kohonen Feature Maps as a Supervised Learning Machine", ICNN, pp.1944-1948, 1993
- [7] Y. C. Shin and R. Sridhar, "Network Configurations and Training Speeds of Second-Order Neural Network", WCNN, Vol.1, pp.585-588, 1993
- [8] J. G. Taylor and S. Coomber, "Learning Higher-Order Correlation", Neural Network, Vol.6, pp.423-427, 1993
- [9] T. J. Sejnowski and P. K. Kienker, "Learning Symmetry Groups with Hidden Unit: Beyond the perceptron", Physica, 22D, pp.260-275, 1986
- [10] Lilly Spirkovska and Max B. Reid, "Connectivity Strategies for Higher-Order Neural Network Applied to pattern Recognition", IJCNN, pp.21-26, San Diego, 1990
- [11] 정 중수, 홍성찬, "위치 변환 패턴 인식을 위한 다항식 고차 뉴런 네트워크", 한국정보처리학회 논문지, 제4권, 제12호, pp.3063-3068, 1997