

자기 분열 및 구조화 신경 회로망

○최두일, 박상희
연세대학교 전기공학과

A self creating and organizing neural network

Doo-Il Choi, Sang-Hui Park
Dept. of Electrical Eng., Yonsei Univ.

ABSTRACT

The Self Creating and Organizing (SCO) is a new architecture and one of the unsupervised learning algorithm for the artificial neural network. SCO begins with only one output node which has a sufficiently wide response range, and the response ranges of all the nodes decrease with time. Self Creating and Organizing Neural Network (SCONN) decides automatically whether adapting the weights of existing node or creating a new node. It is compared to the Kohonen's Self Organizing Feature Map (SOFM). The results show that SCONN has lots of advantages over other competitive learning architecture.

I. 서론

일반적으로, 인공 신경 회로망의 학습 법칙인 경쟁 학습 알고리즘은 벡터 양자기를 구성한다고 볼 수 있으며, 회로망은 지속적인 학습을 통해 입력 패턴의 구조를 자동적으로 발견할 수 있다. 현재까지 제안된 많은 경쟁학습 알고리즘 가운데, Kohonen[1]이 제안한 자기 구조화 특징 지도 - Self Organizing Feature Map (SOFM) - 는 입력 패턴을 자동적으로 적응 양자화할 수 있을 뿐 아니라, 출력 노드를 가지런히 구조화 할 수 있다는 점에서 가장 우수한 경쟁 학습 알고리즘 가운데 하나라고 알려져 있다. 그러나, SOFM은 몇가지 단점을 보이고 있음을 실험을 통해 확인할 수 있었다. 첫째로, SOFM은 구석에 위치한 노드가 중심에 위치한 노드보다 더 많이 학습되기 때문에 경계 효과가 발생하게 된다. 특히, 입력패턴이 순환성 분포를 갖을 경우에 경계 효과는 더욱 심각하게 된다. 두번째로, 이는 경쟁 학습 알고리즘의 가장 일반적인 문제이지만, 입력 패턴이 복잡한 분포를 갖는 경우에는 회로망이 매우 불안정해져서 입력 패턴을 균일하게 분할하는데 실패하는 경우가 빈번하게 발생

된다. 이러한 문제점을 해결하기 위하여 Kang[3]은 수정된 SOFM 알고리즘을 제안하였으나 크게 안정성을 향상시키지는 못하고 있다

본 연구에서는, 상술한 바와 같은 문제점을 제거하며, 학습 시간이 빠르고, 최적의 출력 노드의 갯수를 자동적으로 탐색할 수 있는 새로운 구조의 신경 회로망 알고리즘을 제안한다.

II. 자기 분열 및 구조화 신경 회로망 알고리즘

초기 상태에는 뉴런의 갯수가 매우 적으며, 뉴런을 활성화 하는 적절한 자극은 매우 넓은 범위를 갖는 것으로 가정하였다. 즉, 뉴런의 응답 범위는 충분히 넓어서 어떠한 입력에 대해서도 활성화 된다. 시간이 흐름에 따라 모든 뉴런의 응답 범위가 매우 천천히 어떤 값으로 줄어들어 어떤 입력 자극에 대해서 가장 큰 자극을 받지만 활성화되지는 않는 뉴런이 생길 수 있다. 이러한 경우에 가장 큰 자극을 받은 뉴런은 분열하여 자식 뉴런을 만들며 이 때 분열된 자식 뉴런은 부모 뉴런을 닮게 된다고 가정하였다.

본 연구에서 제안한 신경 회로망의 알고리즘은 상술한 바와 같은 가정에 기초하고 있다. 상자 1은 상세한 SCONN 알고리즘을 나타낸다.

[Step 1.] 초기 상태에는 단지 하나의 출력노드가 존재하며, ($j=0$) 이 노드와 입력 노드간의 연결 강도는 작은 값으로 랜덤하게 초기화 된다. 이 때 초기 노드(P cell)의 응답 범위는 어떠한 입력에도 반응하도록 매우 큰 값으로 설정된다.

[Step 2.] 새로운 입력이 순차적으로나 랜덤하게 제시된다.

상자 1. SCONN 알고리즘

Step 1. Initialize Weights of a Primitive Node

Step 2. Present New Input

Step 3. Calculate Distance to All Node(s)

Step 4. Select Winner Node

Step 5. Decide whether Winner Node is Active.
 If winner node is active,
 then goto step 6.
 If winner node is inactive,
 then goto step 7.

Step 6. Organizing an Active Winner Node
 by Adapting Weights goto Step 2

Step 7. Creat a Son Node from an Inactive
 Winner (Mother) Node goto Step 2

[Step 3.] 입력과 각 출력 노드 j 사이의 거리 d_j 를 식 (1)을 이용하여 구한다.

$$d_j = \sum_{i=0}^{N-1} (X_i(t) - W_{i,j}(t))^2 \quad (1)$$

여기서, N은 입력 벡터의 차원, $X_i(t)$ 는 시간 t에서 i번째 입력 노드의 입력, $W_{i,j}(t)$ 는 시간 t에서 i번째 입력 노드와 j번째 출력 노드간의 연결 강도이다.

[Step 4.] 최소 거리를 갖는 단 하나의 출력노드가 경쟁에서 이긴 노드로 선택된다.

[Step 5.] 식 (2)를 이용하여 경쟁에서 이긴 노드가 활성 상태인지 비활성 상태인지를 판별한다.

$$Y_{wj} \begin{cases} \text{is active,} & \text{if } d_{wj} < \theta(t) \\ \text{is inactive,} & \text{otherwise} \end{cases} \quad (2)$$

여기서, Y_{wj} 는 경쟁에서 이긴 노드의 출력, d_{wj} 는 입력과 경쟁에서 이긴 출력 노드간의 연결 강도이고, $\theta(t)$ 는 출력 노드의 응답 범위로서 시간에 따라 지수함수적으로 감소하여 적절한 값으로 수렴한다.

만일 경쟁에서 이긴 노드가 활성 상태이면 신경 회로망은 step 6.에서 구조화되며, 그렇지 않으면 step 7.에서 분열한다.

[Step 6.] 활성 상태의 경쟁에서 이긴 노드의 연결 강도를 식 (3)을 이용하여 학습시킨다.

$$W_{i,wj}(t+1) = W_{i,wj}(t) + \alpha(t) (X_i(t) - W_{i,wj}(t)) \quad (3)$$

여기서, $W_{i,wj}(t)$ 는 입력과 활성 상태의 경쟁에서 이긴 노드간의 연결 강도이고, $\alpha(t)$ 는 학습율로서, 상수이거나 시간에 따라 감소한다.

[Step 7.] 식 (4)와 (5)를 이용하여 부모 노드(비활성 상태의 경쟁에서 이긴 출력 노드)로 부터 자식 노드가 분열한다. 식 (5)에 의하여 자식 노드는 부모 노드를 답게 된다.

$$s_j = s_j + 1 \quad (4)$$

$$W_{i,wj}(t+1) = W_{i,wj}(t) + R(t) (X_i(t) - W_{i,wj}(t)) \quad (5)$$

여기서, $W_{i,wj}$ 는 입력과 분열된 자식 노드간의 연결 강도이고, $R(t)$ 는 다음 요소로서, 0에서 1사이의 값을 갖는다.

이 알고리즘에서 학습을 완료시키는 지표는 3가지가 가능하다. 즉, 학습 시간 (t), 출력 노드의 개수 (s_j) 또는 응답 범위 ($\theta(t)$)가 지표가 될 수 있다.

회로망이 최적의 출력 노드 수를 자동적으로 탐색하기 위해서는 응답 범위 ($\theta(t)$)가 학습을 완료시키는 지표가 되는 것이 바람직하다.

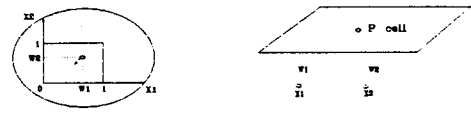
그림 1.은 자기 분열 및 구조화의 개념을 나타내고 있다.

개념 정립을 위하여, 0부터 1사이에서 균일한 분포를 갖는 두개의 상호 독립 입력 $X1$ 과 $X2$ 가 신경 회로망에 입력되며, 초기의 단일 노드의 응답 범위는 1.414보다 크며, 연결 강도는 0부터 1사이의 랜덤한 값으로 초기화되었다고 가정한다. 만일 노드의 응답 범위가 매우 천천히 감소한다면 초기 단일 노드의 연결 강도 벡터는 학습이 반복되면서 입력 패턴의 평균값으로 수렴한다. 그림 1 (a)는 이와 같은 SCONN의 초기 단계를 나타내고 있다. 입력 패턴은 사각형 안에서 균일하게 분포하고 있다. 큰 원은 초기 노드(P cell)의 응답 범위를 나타내고 있으며 매우 천천히 감소하기 때문에 초기 노드는 학습만 될 뿐 분열하지는 않는다.

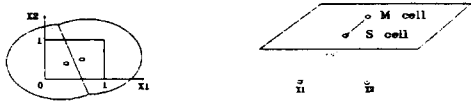
충분한 학습 시간이 흐른 후, 초기 노드의 응답 범위가 서서히 감소함에 따라 어떤 입력에 대해서는 초기 노드가 비활성화될 수 있으며, 이 때 초기 노드는 분열하여 자식 노드를 만든다. 그림 1 (b)는 이와 같은 SCONN의 분열 단계를 나타내고 있다.

태초의 분열 단계 이후에, 모든 노드는 서로 경쟁하여 자기 영역의 입력에 대하여 응답하게 되고 각 노드의 연결 강도는 자기 영역의 입력의 평균값으로 수렴하게 된다. 그림 1 (c)는 이와 같은 SCONN의 구조화 단계를 나타낸다.

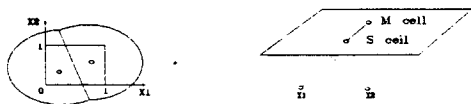
SCONN에서는 분열 단계와 구조화 단계가 계속되면서, 그림 1 (d)와 같이 입력 패턴을 균일하게 양자화하게 된다.



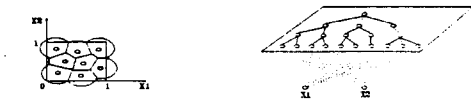
(a) SCONN의 초기 단계



(b) SCONN의 분열 단계

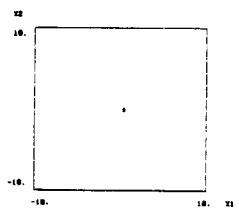


(c) SCONN의 구조화 단계

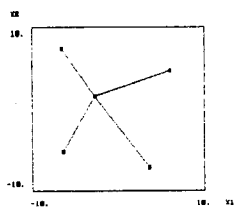


(d) SCONN의 최종 단계

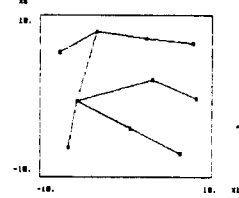
그림 1. SCONN의 개념



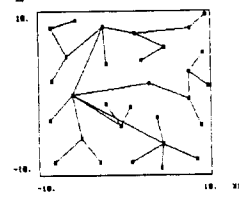
(a) t = 0



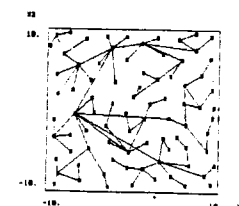
(b) t = 200



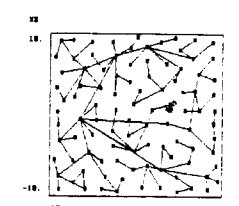
(c) t = 1000



(d) t = 2000



(e) t = 5000



(f) t = 10000

그림 2. 균일한 분포의 입력 패턴에 대한

SCONN의 천이 과정

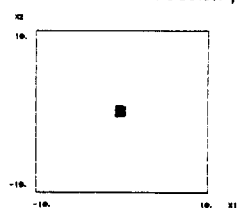
Ⅲ. 결과

적응 벡터 양자기로서 SCONN의 성능을 알아보기 위하여, 입력 패턴의 확률 분포가 1. 균일한 경우, 2. 복잡한 구조를 갖는 경우 및 3. 가우시안 분포를 갖는 경우에 대하여 시뮬레이션하였다. SCONN에 대한 시뮬레이션 결과는 같은 경우의 SOPM에 대한 시뮬레이션 결과와 비교하였다.

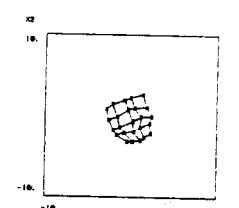
3-1. 균일한 확률 분포의 2차원 입력 벡터

2 차원 입력 벡터 X_1, X_2 가 독립적이고 -10부터 +10 사이에 균일하게 분포된 경우이다. 그림 2.는 이와 같은 입력 패턴에 대하여 SCONN이 천이하는 과정을 보여주고 있다. 여기서, 작은 사각형은 각 노드의 연결 강도를 나타내고 있으며, 부모 노드와 자식 노드의 연결 강도 벡터는 직선으로 연결하였다.

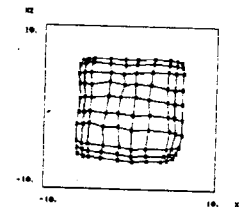
그림 3.은 같은 경우에 대하여 SOPM이 천이하는 과정을 보여주고 있다. 여기서, 작은 사각형은 각 노드의 연결 강도를 나타내고 있으며, 인접한 노드는 직선으로 연결하였다. 이와 같이 입력패턴의 분포가 단순한 경우에는 SOPM도 안정적으로 패턴 분할을 수행할 수 있으나, 경계 효과가 보이는데 이는 학습 횟수를 늘려도 좀처럼 제거되지 않았다. 반면, SCONN은 단시간내에 안정적으로 패턴 분할을 수행하고 있음을 보인다.



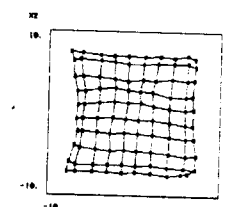
(a) t = 0



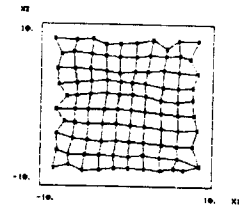
(b) t = 200



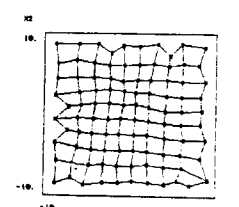
(c) t = 1000



(d) t = 2000



(e) t = 5000



(f) t = 10000

그림 3. 균일한 분포의 입력 패턴에 대한

SOPM의 천이 과정

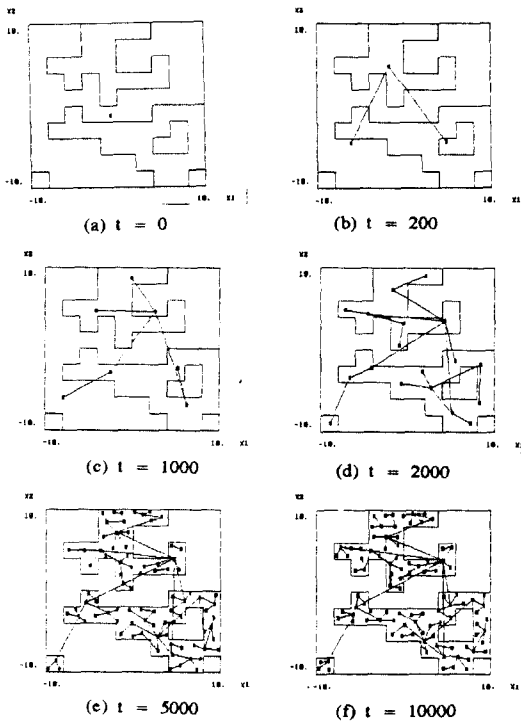


그림 4. 구조화된 입력 패턴에 대한

SCONN의 천이 과정

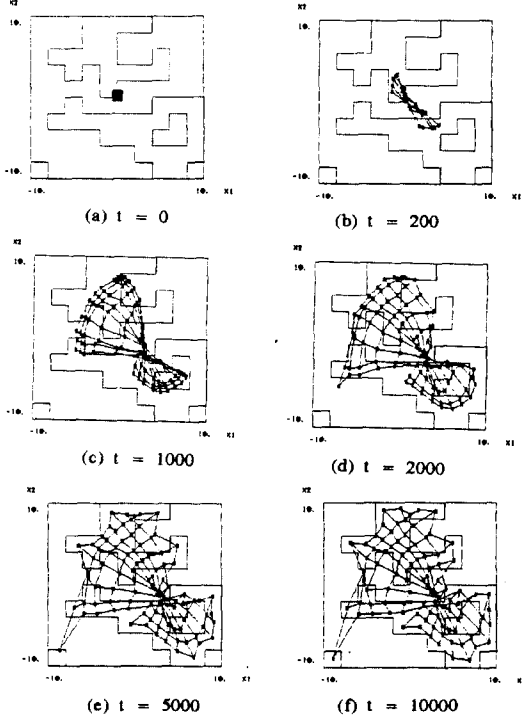


그림 5. 구조화된 입력 패턴에 대한

SOFM의 천이 과정

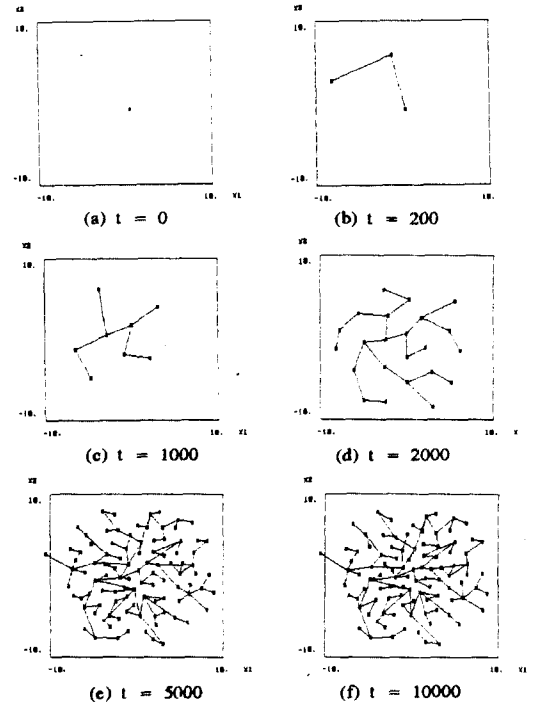


그림 6. 가우시안 분포의 입력 패턴에 대한

SCONN의 천이 과정

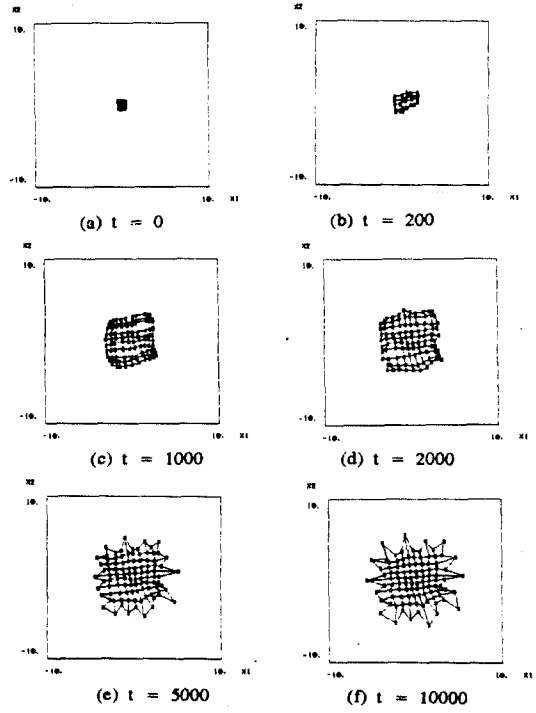


그림 7. 가우시안 분포의 입력 패턴에 대한

SOFM의 천이 과정

3-2. 구조화된 확률 분포의 2차원 입력 벡터

2 차원 입력 벡터가 복잡한 구조를 가지면서 균일하게 분포된 경우이다. 그림 4.는 이와 같은 입력 패턴에 대하여 SCONN이 천이하는 과정을 보여주고 있다. 그림 5.는 같은 경우에 대하여 SOFM이 천이하는 과정을 보여주고 있다. SOFM은 패턴의 구조가 복잡할 수록 쓸모없는 출력 노드들 만들어 경계 효과가 심각해지며 균일한 패턴 분할에 실패한다. 반복된 실험에서도 만족스러운 결과를 보이지는 못했다. 반면, SCONN은 반복된 실험에서, 어느 경우에도 비교적 균일한 패턴 분할을 할 수 있었다.

3-3. 가우시안 분포의 2차원 입력 벡터

2 차원 입력 벡터 X_1, X_2 가 독립적이고, 평균 0, 분산 20/3인 가우시안 분포를 갖는 경우이다. 마찬가지로, 이와 같은 입력 패턴에 대하여 그림 6.은 SCONN이, 그림 7.은 SOFM이 천이하는 과정을 보여주고 있다.

표 1.은 학습이 완료 ($t=10000$) 된 후, 여러가지 입력 환경에 있어서 SCONN과 SOFM에 대한 mean square quantization error를 나타내고 있다. SOFM은 출력 노드들 100개로 고정하여 10000번 학습시켰으며, SCONN은 출력 노드 수가 100이 되면 더 이상 자식 노드를 만들지 않도록 하고 학습 횟수는 10000이 되도록 하였다. mean square quantization error는 식 (6)을 이용하여 구하였다.

$$MSQE = E \{P(X(t)) \cdot (X(t) - W_{win}(t))\}^2$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \|X(t) - W_{win}(t)\| \quad (6)$$

여기서, $X(t)$ 는 제시된 입력 벡터, $W_{win}(t)$ 는 경쟁에서 이긴 노드의 연결 강도 벡터, $\|\cdot\|$ 는 두 벡터의 유클리드 거리를 나타낸다. 실험에서는 $T = 10000$ 으로 하였다.

IV. 결론

본 연구에서, 스스로 분열하고 구조화하는 새로운 형태의 신경 회로망을 제안하였다. 제안된 신경 회로망은 경쟁 학습에 의한 적응 벡터 양자기를 구성하고 있으며, 다음과 같은 특징을 갖고 있다.

1. SCONN은 매우 안정하여 입력 패턴이 복잡한 확률 분포를 갖더라도 입력 패턴을 균일하게 양자화 한다.
2. SCONN은 출력 노드의 최적 갯수를 자동적으로 찾을 수 있다.
3. SCONN은 SOFM에 존재하는 경계효과를 제거하였다.

표 1. Mean Square Quantization Error

	2-D UD		2-D SD	
	SCONN	SOFM	SCONN	SOFM
1	0.780873	0.831283	0.508925	0.594043
2	0.784397	0.810908	0.507568	0.603266
3	0.786804	0.819177	0.502937	0.607996
norm.avg.	1.	1.046467	1.	1.188146

	3-D UD		3-D SD	
	SCONN	SOFM	SCONN	SOFM
1	2.118660	2.308067	1.026361	1.344519
2	2.115548	2.312588	1.032106	1.413899
3	2.116776	2.349412	1.037200	1.384718
norm.avg.	1	1.097478	1.	1.338366

	10-D UD		10-D SD	
	SCONN	SOFM	SCONN	SOFM
1	6.041004	6.216311	9.991150	10.244908
2	6.026380	6.194568	10.012160	10.221279
3	6.032107	6.239135	10.057388	10.223208
norm.avg.	1	1.030416	1	1.020914

※ UD : uniform density SD : structured density
norm. avg. : normalized average

4. SCONN은 필요없는 출력 노드를 갖지 않는다.
5. SCONN은 학습 시간이 매우 빠르다.
6. SCONN은 초기 상태에 의존하지 않는다.

SCONN은 신경 회로망의 경쟁 학습 법칙에 존재하는 대부분의 문제점을 제거하였다. SCONN은 SOFM처럼 출력 노드를 가지런히 배열할 수는 없으나, 출력 노드를 가지 형태로 배열함으로써 인접한 출력 노드는 비슷한 입력패턴에 응답할 수 있으리라고 생각된다.

V. 참고문헌

1. T. Kohonen, *Self Organization and Associative Memory*. 2nd edition, Springer-Verlag, 1988, ch. 5, pp.119-157.
2. T. Kohonen, G. Barna and R. Chrisley, "Statistical Pattern Recognition with Neural Network: Benchmarking Studies", *IEEE Proc. of ICNN*, Vol. 1, pp. 61-68, 1988.
3. J. Kangas, T. Kohonen, J. Laaksonen, O. Simula and O. Ventä, "Variants of Self-Organizing Maps" *IEEE Proc. of IJCNN*, Vol. 2, pp.517-522, 1989.