

자기 분열 및 구조화 신경 회로망

A Self Creating and Organizing Neural Network

崔 斗 一* · 朴 相 暉**
(Doo-Il Choi · Sang-Hui Park)

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.

Key Words : Competitive learning(경쟁학습), Neural network(신경회로망), Self-Organizing(자기구조화), Adaptive Vector Quantization(적응벡터양자화), Neural Vector Quantization(신경벡터양자화)

1. 서 론

고전적인 벡터 양자화(VQ : Vector Quantization) 시스템은 평균 자승 오차(MSQE : mean square quantization error) 또는 엔트로피 성능 지표(entropic performance measure)를 최소화하

도록 설계된다. 입력 패턴 벡터의 확률 정보를 알 수 있다면, 최적 코드북 벡터(optimal codebook vector)를 만들 수 있다.[1] 그러나 일반적으로 이러한 확률 정보를 정확히 얻기는 매우 어렵기 때문에 확률 밀도 함수 $f(x)$ 를 자동적으로 추정하는 적응 벡터 양자화(AVQ : Adaptive Vector Quantization) 기법이 흔히 이용된다.

경쟁학습(competitive learning)을 이용하는 대부분의 인공 신경 회로망은 신경 AVQ시스템으로 볼 수 있다. 경쟁 학습에서 연결 강도 벡터

*正 會 員 : 延世大 大學院 電氣工學科 博士課程
**正 會 員 : 延世大 工大 電氣工學科 教授 · 工博
接受日字 : 1991年 9月 12日
1次修正 : 1992年 2月 6日

(weight vector)는 MSQE를 최소화하는 미지의 도식 \bar{x} 으로 수렴하게 되고, 도식 \bar{x} 근처에서 브로니안 운동(Brownian motion)을 하면서 점근적으로 $E[w_i]=\bar{x}_i$ 를 만족하게 된다. [2] 그러나 일반적으로 경쟁학습은 초기의 랜덤 연결 강도에 크게 의존하기 때문에 안정적으로 패턴 분할을 하지 못하는 경우가 빈번하게 발생된다. 이와 같은 문제점을 제거하기 위하여 많은 방법이 제안되었으며, 현재까지 제안된 많은 경쟁학습 알고리즘 가운데, Kohonen[3]이 제안한 자기 구조화 특징 지도(SOFM : Self Organizing Feature Map)는 입력 패턴을 자동적으로 적용 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 경계효과(boundary effect)가 발생하게 된다. 특히, 입력패턴이 순환성 분포(circular distribution)를 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 발생된다. 세번째로, 확률 밀도가 0인 영역에 존재하는 연결 강도 벡터가 주위의 0이 아닌 입력의 영향을 받아서 입력 벡터 영역으로 수렴하지 못하는 경우가 종종 발생된다. 이러한 현상은 이웃(neighbor)이 감소함에 따라 없어지지만 잔여 효과는 그대로 남게 된다. [4] 이러한 문제점을 제거하기 위하여 Kangas등 [4]은 소위 Minimal spanning tree(MST) topology 를 이용한 local neighborhood개념을 도입하여 수정된 SOFM 알고리즘을 제안하였으나 연결 강도 벡터가 입력 벡터 영역으로 수렴하지 못하는 현상은 여전히 남아 있으며, 경우에 따라서는 SOFM보다 패턴 clustering능력이 떨어지기도 하였다.

본 연구에서는, 상술한 바와 같은 문제점을 제거하여 어떠한 입력 패턴에 대해서도 안정적으로 패턴 clustering을 수행할 수 있으며, 학습 시간이 빠르고, 최적의 출력 노드의 갯수를 자동적으로 탐색할 수 있는 새로운 구조의 신경 회로망 알고리즘을 제안한다.

2. 자기 분열 및 구조화 신경 회로망 알고리즘

초기 상태에는 뉴런(neuron)의 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 줄어들었다면 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. 상자1은 상세한 자기 분열 및 구조화 신경 회로망(SCONN) 알고리즘을 나타낸다.

상자 1 SCONN 알고리즘
Box 1 Algorithm for SCONN

- | |
|--|
| <p>Step 1. Initialize Weights of a Primitive Node</p> <p>Step 2. Present New Input</p> <p>Step 3. Calculate Distance to All Node(s)</p> <p>Step 4. Select Winner Node</p> <p>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.</p> <p>Step 6. Organizing an Active Winner Node by Adapting Weights
goto Step 2</p> <p>Step 7. Creat a Son Node from an Inactive Winner (Mother) Node
goto Step 2</p> |
|--|

[Step 1] 초기 상태에는 단지 하나의 출력노드가 존재하며, ($j=0$)이 노드와 입력 노드간의 연결 강도는 일반적인 신경 회로망에서와 같이 작은 값으로 랜덤하게 초기화 된다.

이 때 초기 노드(P cell : Primitive cell)의 응답 범위 $\theta(0)$ 는 매우 큰 값으로 설정되어 어떠한 입력에 대해서도 초기 노드가 활성화 상태가 되도록 한다. 즉, $\theta(0)=\theta_{max}$ 으로 설정 한다.

[Step 2] 새로운 입력이 순차적으로 랜덤하게 제시된다.

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

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

여기서, N 은 입력 벡터의 차원, $X_i(t)$ 는 시간 t 에서 i 번째 입력 노드의 입력, $W_{ij}(t)$ 는 시간 t

에서 i 번째 입력 노드와 j 번째 출력 노드간의 연결 강도이다.

[Step 4] 최소 거리를 갖는 단 하나의 출력노드가 경쟁에서 이긴 노드로 선택된다. 학습의 초기에는 출력 노드가 단 하나밖에 없기 때문에 출력 노드간의 경쟁없이 초기 노드가 항상 이긴 노드로 선택되지만 학습이 진행되면서 출력 노드들이 분열하게 되어 노드수가 증가하게 되고, 이 때는 모든 출력 노드들이 서로 경쟁하게 된다.

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

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

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. 시뮬레이션에서는 식(3)과 같은 $\theta(t)$ 를 이용하였다.

$$\theta(t) = (\theta_{\max} - \theta_{\min}) \times \exp(-0.001 \times t) + \theta_{\min} \quad (3)$$

여기서, θ_{\max} 와 θ_{\min} 은 상수이다.

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

[Step 6] 활성 상태의 경쟁에서 이긴 노드의 연결 강도를 식(4)를 이용하여 학습시킨다.

$$W_{i,w_j}(t+1) = W_{i,w_j}(t) + \alpha(t)(X_i(t) - W_{i,w_j}(t)) \quad (4)$$

여기서, $W_{i,w_j}(t)$ 는 입력과 활성 상태의 경쟁에서 이긴 노드간의 연결 강도이고, $\alpha(t)$ 는 학습율로서, 상수이거나 시간에 따라 감소한다. 시뮬레이션에서는 $\alpha(t) = 0.085$ 로 고정하였다.

[Step 7] 식(5)와 (6)을 이용하여 부모 노드(비활성 상태의 경쟁에서 이긴 출력 노드)로부터 자식 노드가 분열한다. 식(6)에 의하여 자식 노드는 부모 노드를 닮게 된다.

$$s_j = s_j + 1 \quad (5)$$

$$W_{i,s_j}(t+1) = W_{i,w_j}(t) + R(t)(X_i(t) - W_{i,w_j}(t)) \quad (6)$$

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

이 알고리즘에서 학습을 완료시키는 지표는 3가지가 가능하다. 즉, 학습시간(t), 출력 노드의 갯

수(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 되는것이 바람직하다.

그림1은 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 초기화 되었다고 가정한다. 만일 노드의 응답범위가 매우 천천히 감소한다면 초기 단일 노드의 연결 강도 벡터는 학습이 반복되면서 입력 패턴의 평균값으로 수렴한다. 그림1(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

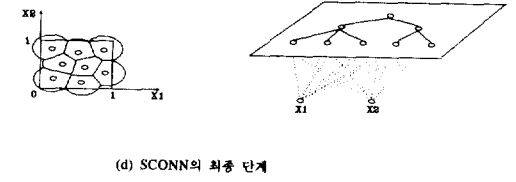
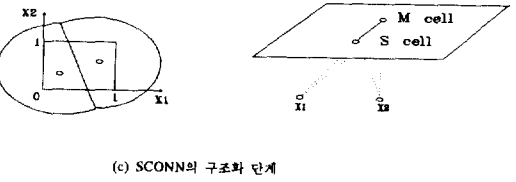
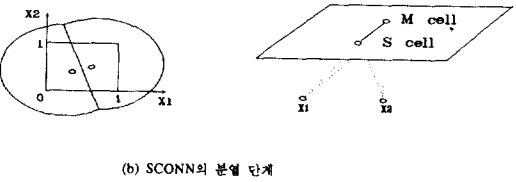
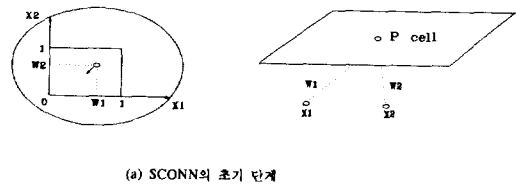


그림 1 SCONN의 개념
Fig. 1 Concept for SCONN

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 이와같은 SCONN의 구조화 단계를 나타낸다.

SCONN에서는 분열 단계와 구조화 단계가 계속

되면서, 그림 1(d)와 같이 입력 패턴을 균일하게 양자화하게 된다.

3. 시뮬레이션 및 결과 고찰

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 SOFM에 대한 시뮬레이션 결과와 비교하였다.

3.1 균일한 확률 분포의 2차원 입력 벡터

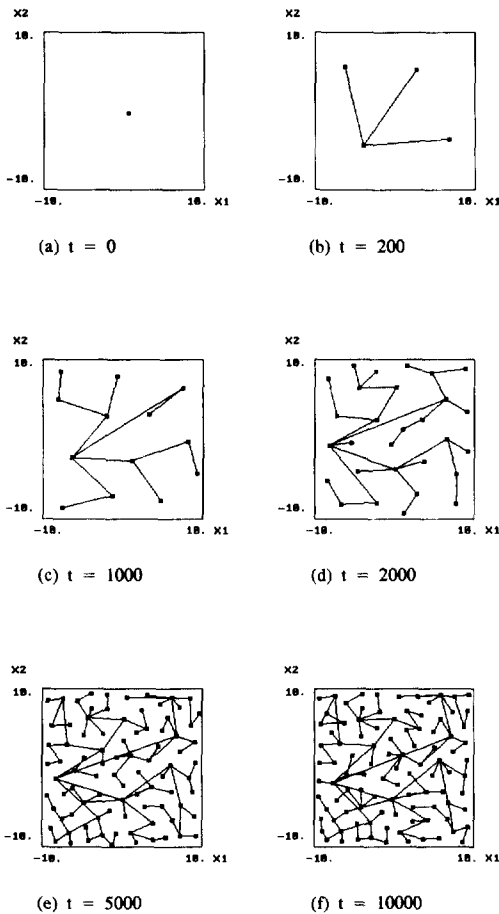


그림 2 균일한 분포의 입력 패턴에 대한 SCONN의 천이 과정

Fig. 2 Transition of SCONN for rectangular uniform input density

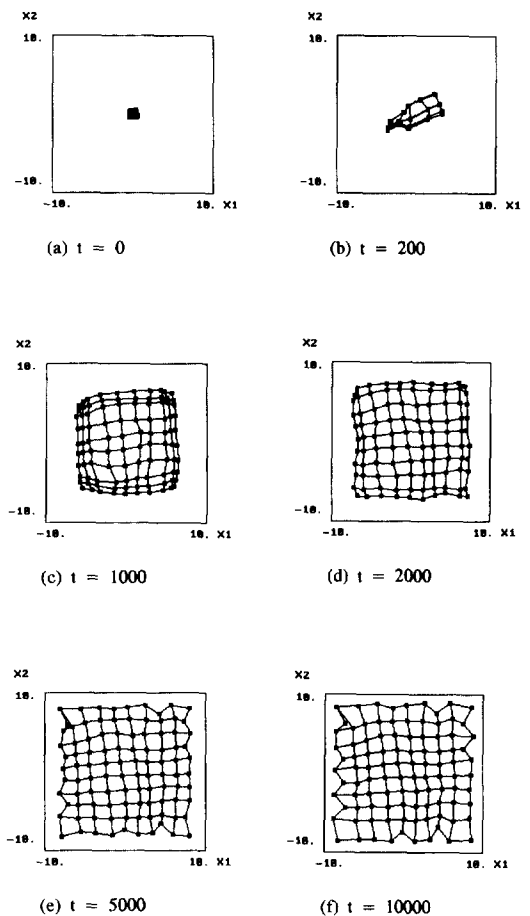


그림 3 균일한 분포의 입력 패턴에 대한 SOFM의 천이 과정

Fig. 3 Transition of SOFM for rectangular uniform input density

2차원 입력 벡터 X_1, X_2 가 독립적이고 -10 부터 $+10$ 사이에서 균일하게 분포된 경우이다. 그림2는 이와 같은 입력 패턴에 대하여 SCONN이 천이하는 과정을 보여주고 있다. 여기서, 작은 사각형은 각 노드의 연결 강도 벡터를 나타내고 있으며, 부모 노드와 자식 노드의 연결 강도 벡터는 직선으로 연결하였다.

그림3은 같은 경우에 대하여 SOFM이 천이하는 과정을 보여주고 있다. 여기서, 작은 사각형은 각 노드의 연결 강도를 나타내고 있으며, 인접한 노드는 직선으로 연결하였다. 이와 같이 입력패턴의 분포가 단순한 경우에는 SOFM도 안정적으로 패

턴 분할을 수행할 수 있으나, 경계 효과가 발생하게 된다. 반면, SCONN은 단시간내에 안정적으로 패턴 분할을 수행하고 있음을 보인다.

3.2 구조화된 확률 분포의 2차원 입력 벡터

2차원 입력 벡터가 복잡한 구조를 가지면서 균일하게 분포된 경우이다. 그림4는 이와 같은 입력 패턴에 대하여 SCONN이 천이하는 과정을 보여주고 있다. 그림5는 같은 경우에 대하여 SOFM이 천이하는 과정을 보여주고 있다. SOFM은 패턴의 구조가 복잡할 수록 쓸모없는 출력 노드를 만들어 경계 효과가 심각해지며 균일한 패턴 분할에 실패

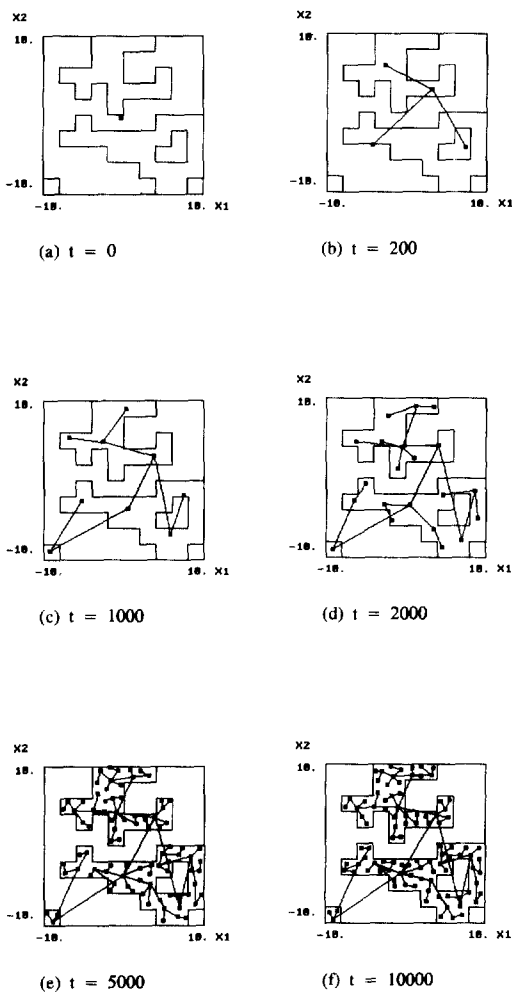


그림 4 구조화된 입력 패턴에 대한 SOFM의 천이 과정
Fig. 4 Transition of SOFM for structured uniform input density

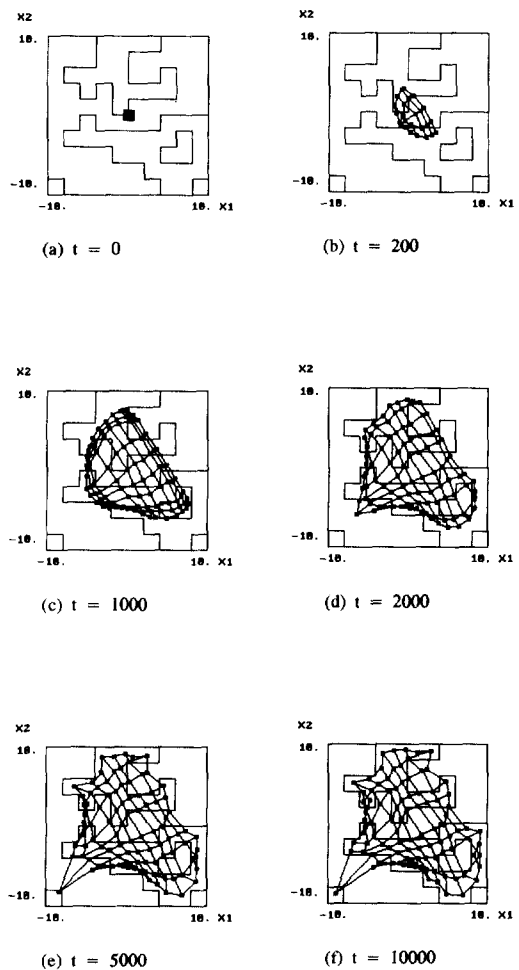


그림 5 구조화된 입력 패턴에 대한 SCONN의 천이 과정
Fig. 5 Transition of SCONN for structured uniform input density

한다. 반복된 실험에서도 만족스러운 결과를 보이지는 못했다. 반면, SCONN은 반복된 실험에서, 어느 경우에도 비교적 균일한 패턴 분할을 할 수 있다.

3.3 가우시안 분포의 2차원 입력 벡터

2차원 입력 벡터 X_1, X_2 가 독립적이고, 평균 0, 분산 $20/3$ 인 가우시안 분포를 갖는 경우이다. 마찬가지로, 이와 같은 입력 패턴에 대하여 그림6은 SCONN이, 그림7은 SOFM이 천이하는 과정을 보여주고 있다.

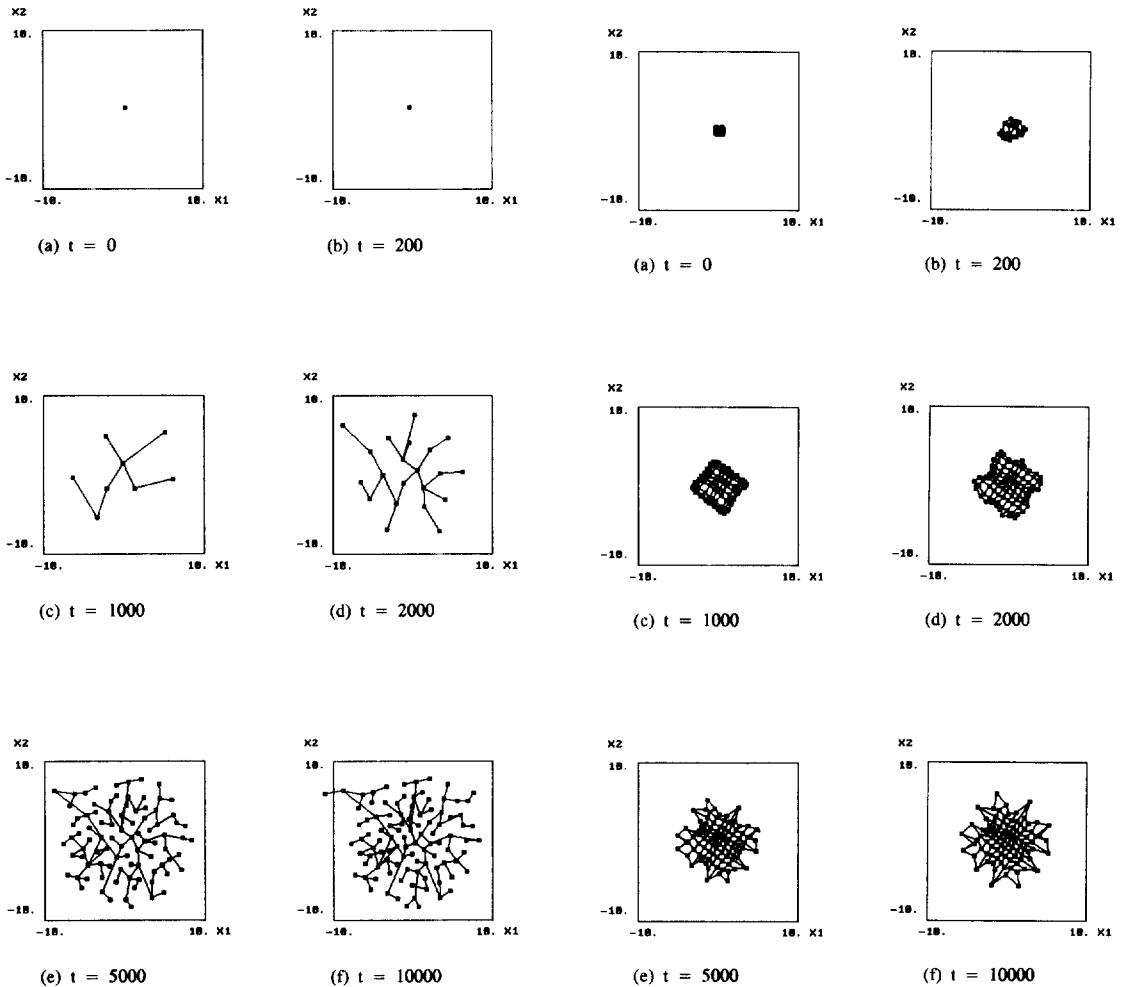


그림 6 가우시안 분포의 입력 패턴에 대한 SCONN의 천이 과정

Fig. 6 Transition of SCONN for gaussian input density

3.4 평균 자승 오차의 비교

표1은 학습이 완료($t=10000$)된 후, 여러가지 입력 환경에 있어서 SCONN과 SOFM에 대한 평균 자승 오차(mean square quantization error)를 나타내고 있다. SOFM은 출력 노드를 100개로 고정하여 10000번 학습시켰으며, SCONN은 출력 노드 수가 100이 되면 더 이상 자식 노드를 만들지 않도록 하고 학습 횟수는 10000이 되도록 하였다.

표1에서 나타난 바와 같이 어느 경우에도 SCONN은 SOFM에 비하여 낮은 MISQE를 보이고 있으며, 특히 SCONN은 입력 패턴이 복잡하게

그림 7 가우시안 분포의 입력 패턴에 대한 SOFM의 천이 과정

Fig. 7 Transition of SOFM for gaussian input density

표 1 평균 자승 오차

Table 1 Mean square quantization error

(a) 2-D 입력인 경우

	2-D uniform density		2-D structured density	
	SCONN	SOFM	SCONN	SOFM
1	0.609762	0.691031	0.259004	0.352887
2	0.615278	0.657571	0.257625	0.363930
3	0.619060	0.671051	0.252945	0.369659
norm.avg.	1	1.095197	1.	1.411789

* 시뮬레이션 조건 (2-D uniform density)

두 입력 x_1, x_2 가 상호 독립이며 $[-10, 10]$ 에서 균일한 경우

ITER=10000, 출력 노드 갯수=100

SCONN : $\alpha(t)=0.085, R(t)=0.85, \theta(t)=8.5\exp(-0.001t)+1.5$

SOFM : $\alpha(t)=0.02, NB=\text{integer}\{10.1\exp(-15t/ITER)\}$

여기서 α 는 학습률, t 는 시간, ITER는 iteration 횟수, R 은 학습 요소, θ 는 응답 영역, NB는 이웃의 범위.

* 시뮬레이션 조건 (2-D structured density)

두 입력 x_1 과 x_2 가 상호 독립이며 $[-10, 10]$ 에서 구조를 갖는 경우

(b) 3-D 입력인 경우

	3-D uniform density		3-D structured density	
	SCONN	SOFM	SCONN	SOFM
1	4.488720	5.327173	1.053416	1.807731
2	4.475543	5.348063	1.065242	1.999110
3	4.480740	5.519736	1.075783	1.917443
norm.avg.	1	1.204535	1.	1.791952

* 시뮬레이션 조건 (3-D uniform density)

세 입력 x_1, x_2 및 x_3 가 상호 독립이며 $[-10, 10]$ 에서 균일한 경우

ITER=10000, 출력 노드 갯수=100

SCONN : $\alpha(t)=0.085, R(t)=0.85, \theta(t)=8.5\exp(-0.001t)+1.5$

SOFM : $\alpha(t)=0.02, NB=\text{integer}\{10.1\exp(-15t/ITER)\}$

* 시뮬레이션 조건 (3-D structured density)

세 입력 x_1 과 x_2 및 x_3 가 상호 독립이며 $[-10, 10]$ 에서 구조를 갖는 경우

(c) 10-D인 입력의 경우

	2-D uniform density		2-D structured density	
	SCONN	SOFM	SCONN	SOFM
1	0.609762	0.691031	0.259004	0.352887
2	0.615278	0.657571	0.257625	0.363930
3	0.619060	0.671051	0.252945	0.369659
norm.avg.	1	1.095197	1.	1.411789

* 시뮬레이션 조건 (10-D uniform density)

열개의 입력 x_1, x_2, \dots, x_{10} 이 상호 독립이며 $[-10, 10]$ 에서 균일한 경우

ITER=10000, 출력 노드 갯수=100

SCONN : $\alpha(t)=0.085, R(t)=0.85, \theta(t)=8.5\exp(-0.001t)+1.5$

SOFM : $\alpha(t)=0.02, NB=\text{integer}\{10.1\exp(-15t/ITER)\}$

* 시뮬레이션 조건 (10-D structured density)

열개의 입력 x_1, x_2, \dots, x_{10} 이 상호 독립이며 $[-10, 10]$ 에서 구조를 갖는 경우

분포된 경우에 매우 강함을 보여서 3차원의 구조화된 입력분포에 대해서는 MSQE를 대략 1/2까지 줄였음을 알 수 있다.

평균 자승 오차는 식(7)을 이용하여 구하였다.

$$MSQE = E\{\|X - W_{win}\|^2\}$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \|X - W_{win}\|^2 \quad (7)$$

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

4. 결론 및 토의

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 개수를 자동적으로 찾을 수 있다.
3. SCONN은 SOFM에 존재하는 경계효과를 제거하였다.
4. SCONN은 필요없는 출력 노드를 갖지 않는다.
5. SCONN은 학습 시간이 매우 빠르다.
6. SCONN은 초기 상태에 의존하지 않는다.

수학적으로 다차원의 정보 공간을 토폴로지를 유지한 채, 2 또는 3차원의 신경장으로 mapping 하는 것은 불가능하다. 그러나 뇌는 다차원의 입력 정보를 2또는 3차원의 신경장 내에서 처리하여 야만 한다. SCONN과 같이 신경 세포를 tree구조로 배열하는 것은 인지 지도를 형성하는 좋은 방법이 될 것으로 생각된다.

한편, SCONN은 학습 시간이 빠르고 패턴 clustering이 매우 안정적이므로 음성이나 영상 신호와 같은 fuzzy한 신호의 인식 및 coding에 매우 유용하리라 생각된다.

참 고 문 헌

- [1] A. Gersho, "Asymptotically Optimal Block Quantization," IEEE Trans. Inform. Theory, Vol. 25, pp. 373~380, July 1979.
- [2] B. Kosko, "Stochastic Competitive Learning," IJCNN-90, Vol. 2, pp. 215~226, June 1990.
- [3] T. Kohonen, Self Organization and Associative Memory. 2nd edition, Springer-Verlag, ch. 5, pp. 119~157, 1988.
- [4] J. Kangas, T. Kohonen, J. Laaksonen, "Variants of Self-Organizing Maps," IEEE Trans. on Neural Network, Vol. 1, pp. 93~99, 1990.
- [5] 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.

저 자 소개



최두일(崔斗一)

1963년 1월 13일생. 1985년 연세대 공대 전기공학과 졸업. 1987년 동 대학원 전기공학과 졸업(석사). 현재 연세대 대학원 전기공학과 박사과정수료.



박상희(朴相暻)

1939년 8월 25일생. 1962년 연세대 공대 전기공학과 졸업. 1964년 동 대학원 졸업. 1971년 동 대학원 졸업(공학). 현재 연세대 공대 전기공학과 교수. 당학회 편집이사(편집위원장).