

복수 패턴 인식을 위한 변형된 네오코그니트론 (Modified Neocognitron for Recognizing Multi-Patterns)

金泰雨*, 崔炳旭*

(Tae-Woo Kim and Byung-Uk Choi)

要約

본 논문에서는 입력영상에 존재하는 복수 패턴을 한번의 패스(pass)로 인식할 수 있는 변형된 네오코그니트론을 제안한다. 변형된 네오코그니트론은 특징을 추출하는 S세포층 및 V세포층과 패턴의 크기, 이동, 회전 등 변형에 덜 민감하게 하는 C세포층으로 구성된 계층적 신경망이다. S세포와 V세포들은 DCC(don't care condition)을 적용하여 영상에 존재하는 모든 패턴에 대한 특징을 추출하는 역할을 하며, S세포와 C세포는 추출된 패턴에 대한 위치정보도 함께 가진다. 위치정보는 추출된 특징들이 서로 간섭을 받지 않고 올바른 특징만이 추출되도록 하는 데 이용된다.

제안한 방법은 피드백이 존재하는 선택적 주의 방법과는 달리 단일패스로 인식이 가능하므로 인식시간이 짧으며, DCC와 위치정보를 이용하므로 서로 접촉되어 있는 복수 패턴의 인식도 가능하다.

Abstract

In this paper, the modified neocognitron, which has capability of recognizing multi-patterns in input image in one pass, is proposed. It is the hierarchical neural network composed of S and V layer which are able to extract features and of C layer with little effect from deformation, changes in size, shifts in position. S and V cells extract the features of all patterns in input image by applying DCC(don't care condition) to those cells. S and C cells also have position informations of extracted patterns. Position information is used in extracting good features without extracted features being interfered one another.

The proposed method is shorter in recognition time than the selective attention method with backward connection, because of recognizing multi-patterns in one passe. The modified neocognitron can recognize attached multi-patterns because of using DCC and position informations.

1. 서론

*正會員, 漢陽大學校 電子通信工學科
(Dept. of Elec. Communication Eng.,
Hanyang Univ.)
接受日字 : 1994年 1月 13日

인공 신경망(Artificial Neural Network)은 인간의 정보처리 메카니즘과 유사하게 작용하도록 모델링하여 병렬처리 능력, 예제학습 능력, 잡음에 강한

특성을 가진다. [1] 이러한 장점 때문에 여러 분야에서 인공신경망의 응용에 관한 연구가 활발히 이루어지고 있다. 특히 문자인식 분야에서 인공 신경망의 응용에 대한 연구가 많이 진행되어 문자인식에 적합한 여러 신경회로망이 발표되었다. 또 한글의 경우에도 신경회로망을 이용한 인식에 관한 연구가 꾸준히 진행되어 왔다. [2,3]

지금까지 제안된 여러가지 신경회로망 중 네오코그니트론(Neocognitron)모델 [5,8]은 패턴인식에 있어서 위치나 크기 변화 등 패턴의 변형에 어느 정도 대응하는 특징이 있으므로 필기체 문자인식에 적합한 신경망 모델로 평가받고 있다. [9] 또한 다른 문자에 비하여 비교적 복잡한 구조를 가지는 한글의 문자인식에 이용하여 자모의 인식과 문자의 인식에 있어서 좋은 결과를 보이고 있다. [10-13]

그러나 기존의 네오코그니트론 모델은 입력영상에서 하나의 패턴을 인식하는 데에는 큰 문제가 없으나 복수개의 패턴이 입력될 경우에는 올바른 인식이 어려워졌다. 이러한 문제점을 고려하여 복수의 패턴을 추출할 수 있는 선택적 주의(Selective Attention)방법이 제안되었다. [14] 그러나, 이 방법은 후향연결(Backward-connection)을 사용하여 여러 번의 패스(pass)에 걸쳐 인식하기 때문에 인식시간이 길어지는 단점을 가진다.

따라서 본 논문에서는 입력영상에 존재하는 복수 패턴을 한번의 패스(pass)로 인식할 수 있는 변형된 네오코그니트론을 제안한다. 변형된 네오코그니트론은 특징을 추출하는 S세포층 및 V세포층과 패턴의 크기, 이동, 회전 등 변형에 덜 민감하게 하는 C세포층으로 구성된 계층적 신경망이다. S세포와 V세포들은 DCC(don't care condition)를 적용하여 영상에 존재하는 모든 패턴에 대한 특징을 추출하는 역할을 한다. DCC는 패턴간의 영향을 제거하여 입력영상에 존재하는 모든 패턴을 동시에 추출할 수 있게 해준다. 그 결과 종속패턴도 동시에 추출되므로 위치정보를 이용하여 제거한다. 위치정보는 첫번째 단에서 입력영상에 존재하는 패턴에 대하여 패턴 단위로 추출되어 S세포와 C세포에 저장되며, 발화된 세포의 위치정보는 다음 단으로 전달된다. 이 정보는 추출된 특징들이 서로 간섭을 받지 않고 올바른 특징만이 추출되도록 하는 데 이용된다.

제안한 방법은 피드백이 존재하는 선택적 주의(Selective Attention) 방법과는 달리 단일패스로 인식이 가능하므로 인식시간이 짧으며, DCC와 위치정보를 이용하므로 많이 붙어 있는 복수 패턴의 인식이 가능하다.

II. 네오코그니트론

1. 네오코그니트론의 구조

네오코그니트론은 잡음이나 크기, 위치변화에 어느 정도 대응할 수 있는 특징을 가지는 계층적 신경망이다. 그 구조는 몇 개의 단(stage)으로 구성되어 있으며, 각 단은 S-층과 C-층으로 구성되어 있다. 각 층(layer)은 세포들의 2차원 배열인 세포평면(cell plane)으로 구성되며, 각 세포평면들은 서로 다른 패턴을 추출하도록 학습된다. 각 세포평면의 세포들은 각각 다른 수용영역(receptive region)에서 정해진 패턴을 추출한다. 세포(cell)는 그 기능에 따라 S, C, V등으로 구분되며 S세포로 이루어진 S층(S layer)과 C세포로 이루어진 C층(C layer)이 교대로 위치한다. S세포는 패턴을 추출하는 기능을 가지며, C세포는 추출된 패턴을 위치변화에 덜 민감하게 하는 기능을 가진다. S세포, V세포와 C세포의 발화값은 각각 식 (1), 식 (2), 식 (3)과 같이 결정된다. [5,9] 여기서 n은 수용영역(receptive region)의 중심, κ 는 S층의 세포평면 번호, k는 C층의 세포평면 번호, r_1 은 선택성 계수, $K_{C_{i-1}}$ 은 k의 범위, A_1 은 수용영역, a_1 은 C세포와 S세포사이의 학습가능한 가중치(weight), b_1 은 V세포와 S세포를 연결하는 학습가능한 가중치, c_1 은 C세포와 V세포를 연결하는 고정된 가중치, d_1 은 S세포와 C세포를 연결하는 고정된 가중치를 나타낸다.

$$u_s(n, \kappa) = r_1 \varphi \left[\frac{1 + \sum_{k=1}^{K_{C_{i-1}}} \sum_{v \in A_1} a_1(v, k, \kappa) u_{C_{i-1}}(n+v, k)}{1 + r_1 b_1(\kappa) u_v(n) / (1 + r_1)} - 1 \right] \quad (1)$$

(여기서 $\varphi[x] = \max(x, 0)$)

$$u_v(n, \kappa) = \left[\sum_{k=1}^{K_{C_{i-1}}} \sum_{v \in A_1} c_1(v, k, \kappa) u_{C_{i-1}}(n+v, k)^2 \right]^{1/2} \quad (2)$$

$$u_c(n, k) = \varphi \left[\sum_{\kappa=1}^{K_S} j_1(\kappa, k) \sum_{v \in D_1} d_1(v) u_s(n+v, \kappa) \right] \quad (3)$$

(여기서 $\varphi[x] = \frac{\varphi[x]}{1 + \varphi[x]}$)

여기서, $j_1(\kappa, k)$ 은 S층과 C층 사이의 연결계수를 나타내는데, S층에서 추출된 패턴들 중 유사한 패턴들은 선택적으로 병합되어 C층으로 전달되도록 하는 역할을 한다.

가중치는 a, b, c, d로 나뉘며, 학습가능한 가중치인 a, b의 학습은 식 (4)와 식 (5)에 의해 결정된

다. 여기서 q_1 은 수정량을 나타낸다.

$$\Delta a_i(v, k, \kappa) = q_1 c_i(v, k, \kappa) u_{i-1}(n + v, k) \quad (4)$$

$$\Delta b_i(\kappa) = q_1 u_{v_i}(n) \quad (5)$$

이때 해당되는 세포평면에 속하는 모든 세포의 가중치 값은 대표세포의 가중치와 똑같이 조정된다. 가중치 c 는 식 (6)에 의하여 결정되며 고정된 값이다.

$$\sum_{k=1}^{K_{S_i-1}} \sum_{v \in A_i} c_i(v, k, \kappa) = 1 \quad \kappa = 1, 2, \dots, K_{S_i} \quad (6)$$

각 S층의 세포들은 정해진 크기의 특징을 추출하며, 첫번째 층의 세포는 작은 국부특징을 추출하고, 추출된 특징들은 다음 층에서 더 큰 특징을 추출하는데 사용된다. S세포층에서 추출된 특징이 다음 S세포층으로 전달될 때 C세포층에서 번짐현상(blurring)이 일어나 위치변화에 덜 민감한 인식이 가능하다.

2. 복수패턴의 추출

네오코그니트론은 입력영상에 하나의 패턴이 존재할 경우에는 인식이 잘 이루어지지만 2개 이상의 패턴인 복수패턴이 존재할 경우에는 어려움이 따른다. 이것은 원형정합(template matching)에 의해 특징을 추출하는 S세포와 V세포의 역할 때문이다.^[4,5]

이러한 원형정합을 통하여 특징을 추출하기 때문에 하나의 수용영역에 하나의 패턴이 존재하지 않으면 올바른 특징을 추출할 수 없다. 즉, 복수의 패턴을 가진 영상이 입력될 경우에 인식층 세포의 발화값이 낮아지며, 극단적인 경우에는 발화값이 현저히 낮아져서 패턴의 인식이 불가능한 경우도 있고, 또 어떤 경우에는 하나의 패턴만이 인식이 된다.

예를 들면, 그림 1의 (a)와 같은 패턴에 대하여 학습을 한 후 그림 1의 (b)와 같은 패턴이 입력되면, 학습한 원형과의 차이가 크게 나타나면 올바르게 특징을 추출할 수 없다. 즉, 두개의 입력패턴 중 학습 패턴이 아닌 패턴이 잡음으로 작용하게 되는데, 그 영향이 작으면 패턴 추출이 가능하지만 영향이 커서 원형과 차이가 크게 나타나면 패턴 추출이 불가능하다. 그 이유는 입력층에 가까운 단(stage)에서는 수용영역이 작으므로 원형에 나타나는 다른 패턴의 영향이 작지만, 최종단에 가까운 단에서는 수용영역이 입력영상의 크기와 비슷하여 한 패턴의 원형에 나타나는 다른 패턴은 잡음으로 나타나기 때문이다. 즉, '1'과 '7'이 원형정합이 이루어질 때, 서로 잡음으로 작용하여 각 인식층 세포의 발화값이 낮아지게 한다.

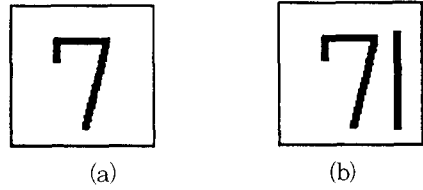


그림 1. (a) 학습된 원형패턴, (b) 입력패턴
Fig. 1. (a) Trained template pattern, (b) Input pattern.

이것은 '7'이 인식되기 위해서는 '7'에 대한 특징이 존재할 조건 뿐만이 아니라 '1'에 대한 특징이 없어야 한다는 조건도 필요하기 때문이다.

이러한 원형정합에서 나타나는 문제점은 네오코그니트론의 연결 가중치를 살펴보면 더욱 명확해진다. 그림 2에서 대표세포(representative cell)와 수용영역의 가중치를 보이고 있다. 학습단계에서 L번째 단(STAGE L)의 1번째 세포평면(P1)의 패턴으로 학습할 때, L-1번째 단(STAGE L-1)의 3번째 세포평면(P3)은 L번째 단(STAGE L)의 패턴에 필요치 않은 특징으로 발화값은 0이 되고 식 (4)에 따라 가중치 a 도 0으로 학습된다. 그러나 인식단계에서 그림 2와 같이 입력패턴(Input pattern)이 입력되면, L-1번째 단의 3번째 세포평면(P3)의 세포가 발화되므로 식 (2)에서와 같이 역제 가중치에 의해 V세포의 값이 커지므로 L번째 단의 발화값이 작아진다.

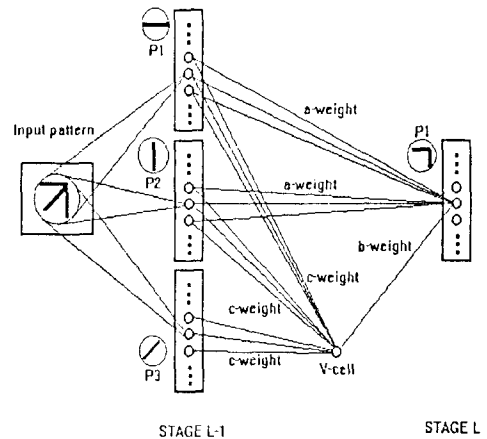


그림 2. 대표세포와 수용영역의 가중치(가중치가 0인 경우 표시하지 않음)

Fig. 2. Representative cell and weights from receptive regions (Zero weights is not represented).

그런데, 여러개의 입력패턴이 입력되었을 경우에

그런데, 여러개의 입력패턴이 입력되었을 경우에 어느 한 패턴이 원형과 유사도가 크게 나타나면 그 패턴을 추출할 수 있다. 특정 패턴을 추출한 후 그 패턴이 제거된 나머지 패턴으로 구성된 영상을 입력 영상으로 할 수 있다. 입력영상에 존재하는 모든 패턴이 추출될 때까지 반복적으로 수행하면 복수패턴에 대해서도 각 패턴의 특징을 모두 추출할 수 있다. 이러한 방법이 선택적 주의(Selective attention) 방법이다.^[14] 즉, 선택적 주의 방법은 인식된 값을 후방 연결에 의해 입력층으로 다시 전달하여 필요치 않은 특징들을 지워나간다. 이와 같은 과정의 반복으로 충분한 인식값이 나타나면, 입력층의 영상에서 다시 다음 패턴을 인식하는 과정을 반복한다. 선택적 주의 방법은 영상의 분할(segmentation)과 불완전한 입력영상의 복원 등의 장점을 지닌다. 그러나 복수개의 패턴들이 서로 붙어 있는 경우에 인식이 어려움이 있으며, 입력패턴의 수만개의 인식과정을 반복하므로 인식시간이 매우 길어진다는 단점을 가지고 있다.

Ⅲ. 변형된 네오코그니트론

이러한 단점을 개선하기 위해서 본 논문에서는 기존의 네오코그니트론을 복수패턴의 인식에 용이하도록 변형하였다. 변형된 네오코그니트론은 구조상으로 기존의 네오코그니트론과 비슷하지만, 세포들의 기능과 그들간의 연결이 다르다. S세포와 V세포들은 특정 부분만이 원형정합에 관여하도록 세포들을 연결함으로써 DCC가 이루어지도록 하여 영상에 존재하는 모든 패턴에 대한 특징을 추출하며, 추출된 패턴에 대한 위치 정보도 가진다. 위치정보는 DCC의 적용 결과로 나타나는 중속패턴을 제거하기 위하여 이용된다. 변형된 네오코그니트론은 위치정보와 DCC를 이용하므로 심하게 붙어있는 패턴에 대해서도 인식이 가능하다. 또 한번의 인식과정으로 복수패턴을 인식할 수 있으므로 인식시간을 줄일 수 있는 장점이 있다.

1. DCC의 적용

그림 3의 (a)에 보이는 영상을 한 세포평면에서 특징을 추출하기 위하여 만들어진 학습패턴이라 하자. 이 학습패턴과 그림 3의 (c)와 같은 입력패턴과 원형정합(template matching)을 할 때, 입력패턴에 다른 패턴이나 잡음 등이 공존할 경우 학습패턴 전체를 원형으로 취하여 정합을 하면 올바른 패턴의 추출이 어렵게 된다. 그러나 그림 3의 (b)와 같이 추출할 패턴부분만 원형으로 취하고 나머지 부분을 DCC를 적

용하면 DCC가 적용된 부분은 원형정합에 관여하지 않으므로 올바르게 원하는 패턴의 추출이 가능하다.

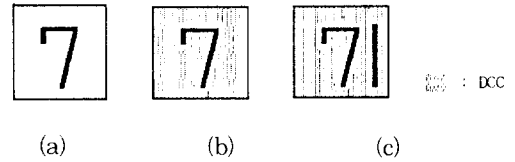


그림 3. (a) 학습패턴, (b) 학습패턴에서의 DCC (c) 입력패턴에서의 DCC

Fig. 3. (a) Training pattern, (b) DCC in a training pattern, (c) DCC in an input pattern.

네오코그니트론에서의 DCC의 구현은 원형정합이 S세포와 V세포의 역할로 이루어지므로 그림 4와 같이 DCC의 적용할 부분에 연결된 C층과 V세포 사이의 가중치 c 를 제거함으로써 DCC를 만들 수 있다. 각 층의 입력단에서 학습을 진행하여 가면서 순차적으로 가중치 c 를 정한다. 즉, 각 패턴에서 필요치 않은 특징들을 검출한 후, 그림 4와 같이 이러한 부분들을 DCC의 적용부분으로 결정하도록 한다. L단(STAGE L)의 학습패턴을 인가했을 때, L-1단(STAGE L-1)의 대표세포의 수용영역 중 발화값으로 0을 가지는 세포들은 인식에 필요치 않은 특징들을 나타내므로 이러한 부분을 DCC의 적용 영역으로 선정한다.

인식과정에서 그림 2에서와 같이 입력패턴이 입력되었을 때, L-1(STAGE L-1)단의 3번째 세포평면(P3)의 세포가 발화되며, 그 발화값은 가중치 c 를 통해 V세포로 전달된다. 이 부분에서 가중치 a 는 0으로 학습되므로 세포의 발화값에 영향을 주는 것은 가중치 c 이다. 이러한 부분의 가중치 c 를 0으로 결정함으로써 그림 4에서 보이는 바와 같이 DCC가 이루어진다. DCC를 적용한 억제 가중치 c 는 식 (7)을 만족하도록 결정한다. 즉, L번째 단의 κ 번째 세포평면의 대표세포 n 에 입력되는 L-1번째 단의 수용영역 중 발화값이 0인 부분에 연결되는 가중치는 $c_{\nu, k}$ (ν, k, κ) = 0이 되도록 설정하고, 발화값이 0이 아닌 세포들은 식 (7)을 만족하도록 결정한다.

$$\sum_{k=1}^{K_{c_{\nu,1}}} \sum_{\nu \in A_1} c_{\nu,1}(\nu, k, \kappa) = 1 \quad \text{for all } u_{c_{\nu,1}}(n, k) > 0, \quad (7)$$

$$\kappa = 1, 2, \dots, K_S$$

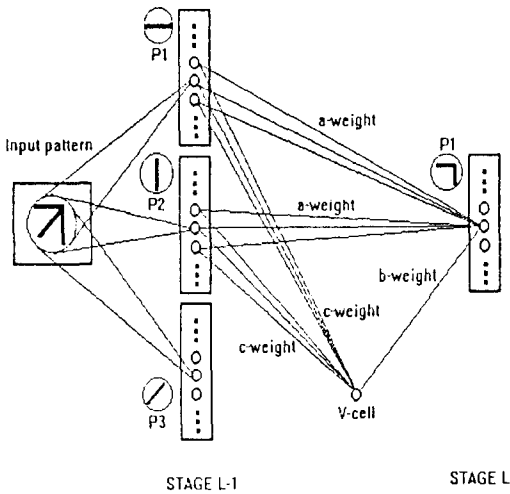


그림 4. DCC를 이용하여 학습한 가중치
Fig. 4. Weights trained using DCC.

2. 위치 정보의 획득 및 이용

변형된 네오코그니트론은 각 단계에서 DCC를 적용하므로 동시에 여러개의 패턴을 추출할 수 있다. 이때, 추출된 패턴들 중에서 어떤 패턴이 다른 패턴에 속해 있는 경우에는 동일한 위치에서 동시에 여러 세포평면에서 발화값이 높게 나온다. 이를테면, "7"이라는 패턴에서 "1"의 패턴이 속해 있다든가 "9"라는 패턴 속에 "0"의 패턴이 속해 있는 경우이다. 이렇게 한 패턴에 속한 다른 패턴을 종속패턴이라 정의하며, 종속패턴은 인식에서 필요하지 않는 패턴이므로 제거되어야 한다. 종속패턴의 제거를 위하여 본 논문에서는 위치정보를 이용한다. 위치정보는 첫단(first stage)에서 학습패턴의 유형에 따라 그림 5와 같이 첫번째 단의 학습패턴의 끝점이 되며 발화값이 임계치 이상이면 S세포에 저장된다. 첫번째단에서 획득된 위치정보는 다음 단으로 전달되면서 병합되어 저장된다.

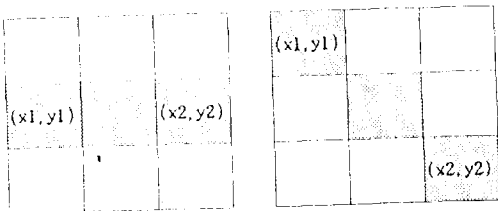


그림 5. 첫단에서의 위치정보 획득
Fig. 5. Acquisition of position informations in the first stage.

종속패턴을 제거하기 위하여 기억된 위치정보를 이용하여 종속패턴의 발화값을 억제시킨다. 수용영역에서 여러개의 패턴이 발화되면, 각 세포평면에서 발화된 세포들의 위치를 비교하여 동일한 위치이면 표 1과 같이 패턴에 대한 종속패턴의 발화값을 억제시킨다. DCC의 적용결과로 나타나는 종속패턴을 제거함으로써 복수패턴을 올바르게 인식할 수 있으며, 종속패턴의 제거는 마지막 단에서만 일어난다.

표 1. 패턴과 종속패턴
Table 1. Patterns and their subpatterns.

패턴	종속패턴
0	1
1	
2	1, 7
3	1, 7
4	1, 7
5	1
6	0, 1, 5
7	1
8	0, 1, 2, 3, 4, 5, 6, 7, 9
9	0, 1, 3, 4

IV. 실험 및 고찰

실험은 SUN SPARCstation 1+에서 행하였으며, C언어를 사용하였다. 실험 대상 패턴은 마우스로 입력된 오프라인(off-line) 필기체 숫자이다.

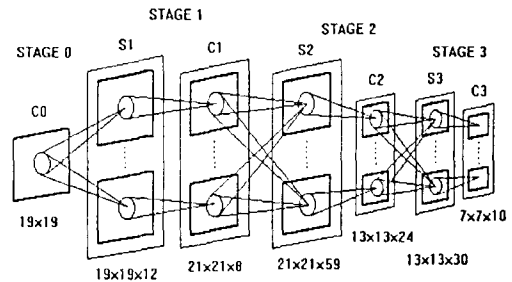


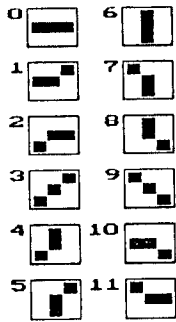
그림 6. 변형된 네오코그니트론의 구조
Fig. 6. Structure of the modified neocognitron.

신경망은 그림 6과 같이 입력단을 비롯하여 총 4단

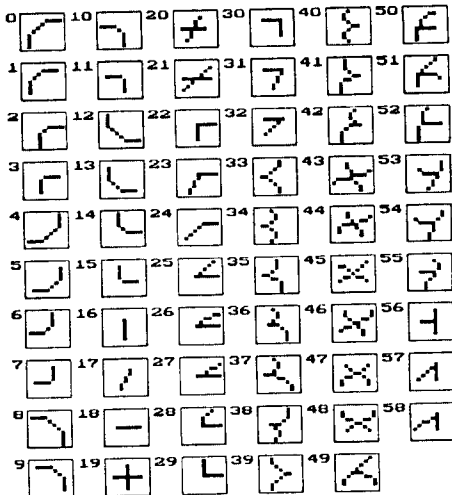
으로 구성된다. 각 단에서 추출되는 특징들이 서로 잘 구별되도록 하는 범위에서 세포수를 정하였다. 각 층의 수용영역의 크기는 표 2에 나타내었다. 1번째 단에서는 세밀한 특징까지 추출해야 하므로 세포수가 많으며, 인식단으로 갈수록 발화되는 세포수가 적게 나타나므로 세포수를 줄였다.

표 2. 수용영역의 크기
Table 2. Size of a receptive region.

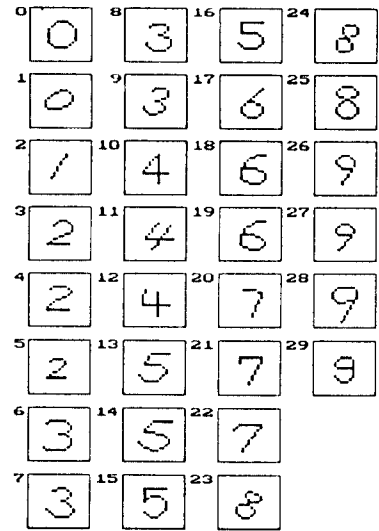
단	층	수용영역의 크기
1단	S1	3×3
	C1	3×3
2단	S2	5×5
	C2	5×5
3단	S3	5×5
	C3	5×5



(a)



(b)



(c)

그림 7. 학습패턴, (a) 1번째 단 (b) 2번째 단 (c) 3번째 단

Fig. 7. Training patterns, (a) for stage 1, (b) for stage 2, (c) for stage 3.

표 3. S층에서 C층으로의 연결
Table 3. Connections of from S layers to C layer.

C층의 세포 평면 번호	S층의 세포평면 번호		
	1단	2단	3단
0	0	0, 1, 2	0, 1
1	1, 2	3	2
2	3	4, 5, 6	3, 4, 5
3	4, 5	7	6~9
4	6	8, 9, 10	10, 11, 12
5	7, 8	11	13~16
6	9	12, 13, 14	17, 18, 19
7	10, 11	15	20, 21, 22
8		16, 17	23, 24, 25
9		18	26~29
10		19, 20, 21	
11		22	
12		23, 24	
13		25, 26, 27	
14		28	
15		29	
16		30	
17		31, 32	
18		33~38	
19		39~42	
20		43~48	
21		49~52	
22		53~56	
23		57, 58	

학습방법은 인식율이 높은 지도학습(Supervised learning)으로 하였으며, 학습패턴은 그림 7과 같이 각 인식패턴의 특징이 잘 나타나도록 만들었다. 또 DCC 때문에 생기는 원형의 열화를 줄이기 위하여 세포평면을 각 학습패턴에 하나씩 할당하였다. 그림 7에서 숫자는 S층의 세포평면 번호를 나타내며, S층에서 C층으로의 연결은 표 3에 보이는 바와 같다. 학습은 1번째 단부터 시작하여 그림 7의 학습패턴으로 2,3번째 단을 차례로 학습하였다. DCC를 적용하기 위하여 먼저 수용영역에 C세포의 발화값이 0인 부분을 DCC영역으로 잡고 식 (7)을 만족하도록 가중치 c를 결정하며, 식 (4)와 같이 가중치 a를 정하였다. 또 가중치 b는 식 (5)와 같이 정하였다. 여기서 학습 수정량 q는 원형정합이 잘 이루어지도록 하기 위해 표 4와 같이 비교적 큰 값으로 하였다. 선택성 계수 r은 표 4에 보이는 바와 같으며, 2단에서는 선택도를 높이기 위하여 비교적 큰 값으로 하였다.

표 4. 학습수정량 및 선택성 계수
Table 4. Magnitude of training update and selection coefficient.

	단	1	2	3
계수				
학습수정량 q		100	2400	1000
선택성계수 r		1.7	7	2

신경망은 학습패턴의 신경망 내부로의 사상(mapping)이 얼마나 균일하게 잘 되느냐에 따라 그 성능이 크게 다르게 나타난다. 네오코그니트론에서는 학습패턴의 신경망 내부로의 균일한 사상 정도는 포화함수에 많이 좌우된다. 세포의 출력값은 포화함수를 거쳐서 0과 1사이의 값으로 사상된다. 기존의 포화함수인 식 (3)의 $\psi [x]$ 는 그림 8의 (a)와 같이 일차함수 $\psi [x] =x$ 보다 전반적으로 낮은 값으로 사상시키며, 넓은 범위의 정의역에 대하여 치역이 분포한다. 그러나 포화함수를 거치기 전의 출력값은 0과 1사이의 값이 대부분이므로, 이 구간에서의 포화함수의 치역은 0과 0.5사이 분포하므로 사상이 균일하게 이루어지지 않는다. 따라서 발화값의 차이가 작아서 패턴들 간의 구별의 정도가 떨어진다.

포화함수의 균일한 사상을 위하여, 본 논문에서는 $\psi [x]$ 를 식 (8)과 같이 시그모이드(Sigmoid) 함수를 사용하였다. 이 함수는 그림 8의 (b)와 같이 비교적 균일한 사상을 시키면서도 1이상의 큰 값에 대해서는 포화함수의 역할을 충분히 한다.

$$\psi [x] = \frac{1}{e^{-6(x-0.5)} + 1} \tag{8}$$

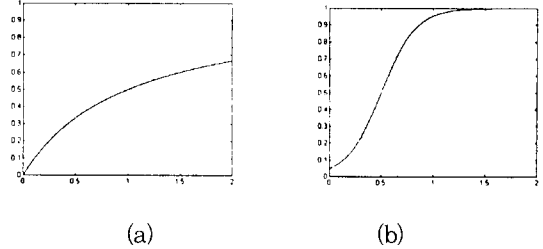


그림 8. 포화함수.
(a) 분수함수 (b) 시그모이드 함수
Fig. 8. Saturation function. (a) fractional function, (b) sigmoidal function.

본 논문에서 제안한 변형된 네오코그니트론으로 그림 9와 같은 복수패턴에 대하여 인식을 하였다. 최종단의 세포의 발화값이 0.5 이상이면 인식한 것으로 간주하였으며, 위치정보를 이용하여 중속패턴을 제거하였다. 그 결과, 대부분이 올바르게 인식되었다. 그러나 그림 9에서 “#” 표시를 한 패턴은 학습패턴과 크게 다른 패턴으로 발화값이 낮게 나와 미인식의 결과를 보였다. 그러나 발화값이 현저히 낮게 나오는 입력패턴을 학습 패턴에 추가시키면 이러한 문제는 해결된다.

인식된 숫자의 순서는 최종단의 발화세포의 위치를 조사하여 구한다.

인식시간은 표 5와 같이 약 3분 정도이며, 여러개의 패턴이 입력될 경우는 하나의 패턴만 입력될 경우와 비교할 때 수초간의 차이밖에 나지 않는데, 이것은 위치정보의 획득 및 이용에 걸리는 시간이다. 또, 본 논문에서는 입력영상의 크기를 19×19로 하였기 때문에 2개의 패턴밖에 입력할 수 없지만, 입력영상의 크기를 증가시키면 여러개의 패턴을 입력할 수 있다. 이 때의 인식시간도 DCC를 이용하므로 하나의 입력패턴의 경우와 거의 차이가 없다.

표 5. 학습 시간 및 인식 시간(단위:초)
Table 5. Training and recognition time (unit:second).

	단	1	2	3	합계
구분					
학습시간		14	2376	5334	7724
인식시간		8	144	64	216

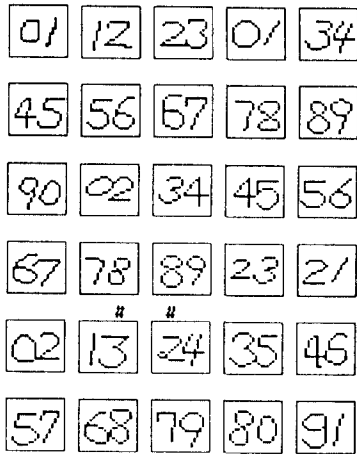


그림 9. 변형된 네오코그니트론에 의한 인식결과 (# : 미인식 패턴)

Fig. 9. Recognition results by the modified neocognitron. (# : unrecognized pattern)

그림 10의 입력 패턴에 대하여 선택적 주의 방법과 제한한 방법의 비교하였다. 입력패턴 #0은 학습패턴들로 구성되었으며, 나머지 입력패턴(#1~#4)은 학습패턴에서 변형된 숫자들이다. 비교의 결과는 표 6과 7에 나타내었는데, 선택성 계수 r 값에 따라서 두 방법이 모두 세포의 발화값이 크게 다르게 나옴을 알 수 있다. 표에서와 같이 DCC의 영향으로 약간의 원형의 열화가 생기지만, 인식에 큰 문제는 없었다. 입력패턴 #0, #1, #3에 대해서는 선택적 주의에서 다른 패턴이 잡음으로 작용하여 하나의 패턴만 발화값이 높게 나오지만, DCC를 이용한 본 논문의 방법은 두 패턴을 한번에 인식하였다. #2와 #4의 입력패턴은 학습패턴과 비교적 큰 차이가 있는 패턴임을 알 수 있다.

그림 10에 대한 결과를 통하여 기존의 선택적 주의 방법과 본 논문의 방법 모두 선택성 계수의 설정과 학습패턴의 구성이 인식율에 크게 영향을 줌을 알 수 있다.

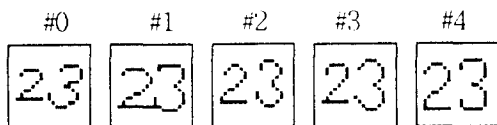


그림 10. 세포의 발화값 비교를 위한 입력 패턴 Fig. 10. Input patterns for comparison of the fire values of cells.

표 6. 그림 10의 패턴에 대하여 r 값에 따른 선택적 주의 방법에 의한 세포의 발화값의 비교

Table 6. Comparison of the fire values of cells for patterns of Fig.10 according to r values by the selective attention method.

패턴	r	1	2	3	4	5	6	7
		#0	2: 0.60	0.48	0.39	0.30	0.24	0.18
	3	0.94	0.93	0.93	0.92	0.92	0.91	0.90
#1	2	0.82	0.69	0.53	0.36	0.22	0.12	0.05
	3	0.24	0.10	0.05	0.00	0.00	0.00	0.00
#2	2	0.65	0.37	0.16	0.04	0.00	0.00	0.00
	3	0.40	0.12	0.05	0.00	0.00	0.00	0.00
#3	2	0.70	0.49	0.33	0.14	0.06	0.00	0.00
	3	0.75	0.57	0.36	0.19	0.09	0.00	0.00
#4	2	0.81	0.67	0.50	0.33	0.19	0.10	0.02
	3	0.54	0.24	0.08	0.00	0.00	0.00	0.00

표 7. 그림 10의 패턴에 대하여 r 값에 따른 제한한 방법에 의한 세포의 발화값의 비교

Table 7. Comparison of the fire value of cells for patterns of Fig.10 according to r values by the proposed method.

패턴	r	1	2	3	4	5	6	7
		#0	2: 0.85	0.76	0.64	0.49	0.35	0.21
	3	0.95	0.94	0.94	0.93	0.93	0.93	0.92
#1	2	0.78	0.60	0.47	0.28	0.15	0.04	0.00
	3	0.65	0.53	0.40	0.23	0.10	0.01	0.00
#2	2	0.63	0.35	0.11	0.00	0.00	0.00	0.00
	3	0.37	0.10	0.02	0.00	0.00	0.00	0.00
#3	2	0.69	0.51	0.25	0.10	0.00	0.00	0.00
	3	0.67	0.50	0.22	0.09	0.00	0.00	0.00
#4	2	0.80	0.65	0.49	0.29	0.16	0.05	0.00
	3	0.46	0.20	0.05	0.00	0.00	0.00	0.00

V. 결론

본 논문에서는 입력영상에 존재하는 복수 패턴을 한번의 패스로 인식할 수 있는 변형된 네오코그니트론을 제안하였다. 변형된 네오코그니트론은 패턴의 크기, 이동, 회전 등 변형에 덜 민감한 기존의 네오코그니트론의 장점을 지니면서 DCC와 위치정보를 동시에 이용하므로 서로 근접해 있거나 붙어 있는 복수개의 입력패턴에 대한 인식을 할 수 있었다. 또 피드백이 존재하는 선택적 주의방법과는 달리 단일패스로 인식이 가능하므로 인식시간이 짧다.

앞으로 선택성 계수의 설정과 학습패턴의 구성에 대한 연구가 필요하며, 복수 패턴을 인식할 수 있는 변형된 네오코그니트론은 여러개의 패턴이 한 문자를 이루는 한글, 한자의 문자인식에 적용될 수 있을 것으로 기대된다. 특히 한글과 같이 복잡한 문자의 자소 인식에 변형된 네오코그니트론을 적용하기 위하여 단순한 위치에 대한 지식만이 아니라 위상(phase) 등 다양한 지식을 이용하는 방법에 대한 연구가 필요하다.

參 考 文 獻

- [1] James A. Freeman, David M. Skapura, *Neural Networks : Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing Company, 1991.
- [2] 이 희건, 방 승양, "신경회로망을 이용한 오프라인 문자인식 연구 동향", 한국정보과학회지, 제9권 제1호, 1991년 2월, pp.33-44
- [3] 제 1회 문자 인식 워크샵 발표 논문집, 1993년 5월
- [4] K. Fukushima, S. Miyake, and T. Ito, "Neocognitron : A neural network model for a mechanism of visual pattern recognition", *IEEE Transactions on System, man, and Cybernetics*, SME-13, 1983.
- [5] K. Fukushima, "Neocognitron : A Hierarchical Neural Network Capable of Visual Pattern Recognition", *Neural Networks*, Vol. 1, 1988, pp.119-130.
- [6] Masato Okada and Kunihiko Fukushima, "Neocognitron Learned by Back-propagation", *Proceedings of International Conference on Fuzzy Logic & Neural Networks*, 1990, pp.667-670.
- [7] Etienne Barnard and David Casasent, "Shift Invariance and the Neocognitron", *Neural Networks*, Vol.3, 1990, pp.403-410.
- [8] K. Fukushima, "Improved Neocognitron with Bend-Detecting Cells", *Proceedings of International Joint Conference on Neural Networks*, Vol. 4, 1992. 6.
- [9] K. Fukushima and Nobuaki Wake, "Handwritten Alphanumeric Character Recognition by the Neocognitron", *IEEE Transactions on Neural Network*, Vol. 2, No.3, May 1991, pp.355-365.
- [10] Sung-Bae Cho and Jin H. Kim, "Hierarchically Structured Neural Networks for Printed Hangul Character Recognition", *Proceedings of International Joint Conference on Neural Networks*, Vol. 1, 1990. 7., pp.265-270.
- [11] Yillbyung Lee, Tae Cheon Kim, and Eun Jin Kim, "Hangul Recognition using Neocognitron", *Proceedings of International Joint Conference on Neural Networks*, Vol. 1, 1990. 1., pp. 416-419.
- [12] Lee, Y., Kim, T.C., & Kim, E.J. "Hangul recognition using neocognitron", *Proceedings of International Joint Conference on Neural Networks*, Vol. 1, 1990.
- [13] Eun Jin Kim and Yillbyung Lee, "Handwritten Hangul Recognition Using a Modified Neocognitron", *Neural Networks*, Vol. 4, 1991, pp.743-750.
- [14] 김 은진, 백 중현, 정 윤경, 이 일병, "수정된 Neocognitron을 사용한 필기체 한글 인식", 인지과학회 논문지, 제 3권 제 1호, 1991년 6

著 者 紹 介

金 泰 雨(正會員) 第31卷 B編 第3號 參照
현재 한양대학교 대학원 전자통신
공학과 박사과정 재학중

崔 炳 旭(正會員) 第31卷 B編 第7號 參照
한양대학교 전자통신공학과 교수