'92 韓國自動制御學術會議論文集 1992. 10. 19~21

회전량에 불변인 제한 신경회로망을 이용한 패턴인식

나 희승, 박 영진

Rotation - Invariant Pattern Recognition System With Constrained Neural Network

Hee-Seung Na and Youngjin Park

Department of Mechanical Engineering, KAIST
373-1 Kusong-dong, Yusong-ku, Taejon, KOREA

ABSTRACT

In pattern recognition, the conventional neural networks contain a large number of weights and require considerable training times and preprocessor to classify a transformed patterns. In this paper, we propose a constrained pattern recognition method which is insensitive to rotation of input pattern by various degrees and does not need any preprocessing. Because these neural network can not be trained by the conventional training algorithm such as error back propagation, a novel training algorithm is suggested. As such a system is useful in problem related to classify overse side and reverse side of 500 won coin. As an illustrative example, identification problem of overse and reverse side of 500 won coin is shown.

1. 서론

신경회로망 (Neural Network) 은 현재 많은 분야에 응용되고 있으며, 특히 패턴 (pattern) 인식문제에 성공적으로 응용되고 있다 [1]. 그것은 신경회로망이 기존의 Neumann 형 컴퓨터와는 달리 병렬처리, 학습 그리고 잡음 (noise) 에 강한 계산기능을 가지고 있기 때문이다. 특히 화상, 음성인식의 문제는 오랫동안 인공지능 분야에서 연구되어 왔으나 주로 부호 및 논리적 방법을 사용하므로써 큰 진전을 보지 못하였다. 그 이유는 패턴 인식문제가 근본적으로 병렬처리를 요구할 뿐만 아니라 그 문제의 해결에 있어 체계적인 알고리즘을 발견하기 어려우며, 또한 잡음에 민감하기 때문이다. 따라서 신경회로망은 이러한 문제점을 해결해 나감으로써 패턴인식 분야에 폭 넓게 응용되고 있다. 특히 오류 역전파 알고리즘은 지난 몇년 동안 많은 분야에 응용되고 있으며, 가장 일반적인 신경 회로망 모델 중의 하나로 제시되고 있다 [2].

기존의 패턴인식 방법은 입력패턴의 표준위치 (position), 방향 (orientation), 크기(size)가 필요하다. 그러나, 실제로 변환된 패턴이 입력될 경우 이를 인식하기 어렵기 때문에 신경회로망을 이용한 변환된 패턴에 둔감 (insensitive) 한 패턴인식 시스템들이 시도되어왔다. 그 예로 neocognition, ART, ADALINE, 고차 신경회로망 (high-order neural network) 등이 있다 [3, 4, 5, 6]. 그러나 이 모델들은 전처리 (preprocessor) 과정과 많은 학습시간을 필요로 하는 단점이 있다. 본 논문의 목적은 이러한 단점들을 해결할 수 있고, 회전량에 둔감한 새로운 연결 형태의 신경회로망인 제한 신경회로망 (constrained neural network) 을 설계하는데 있다. 또한 오류 역전파 알고리즘의 문제점을 살펴보고 새로 설계된 제한 신경회로망에 적합한 학습 알고리즘을 개발하고자 한다. 제시한 신경회로망은 특히 동전인식 (coin recognition)에 적합하며, 그 응용사례로 500원 동전의 앞면과 뒷면을 회전량에 관계없이 구별할 수 있음을 보이고 그 효율성을 검증하고자 한다.

2. 회전량에 불변인 제한 신경회로망 (constrained neural network) 의 설계

패턴인식에서 입력패턴의 병진 (translation), 회전 (rotation), 크기 (scale) 의 변화에 따른 정확한 인식 (recognition) 은 매우 중요하다. 이중에서 회전량에 불변인 패턴 인식 시스템의 경우, 90° 회전량에 불변인 시스템이 Widrow 에 의해 제안되었다 [5]. 최근 회전량의 자유도 (degree of freedom)가 90°에 국한되지 않고 많은 자유도를 갖는 시스템이 Fukumi 에 의해 개발되었다 [7]. 이 경우 그림 1과 같이 입력패턴을 직교좌표에서 극좌표로 바꿔 입력하고, 회전량에 불변이기 위한 전처리 과정을 거친후 기존의 신경회로망으로 학습하였다. 학습방법으로 오류역전파 알고리즘을 사용하였다.

본 논문에서는 회전량에 불변인 패턴 인식 시스템으로 극좌표 입력을 전처리 과정없이 바로 학습할수 있는 제한신경회로망의 설계및 학습 알고리즘을 제안하기로 한다. 그림 1의 경우는 3개/30°회전량을 갖는 극좌표 입력이다. 보다간단한 설

명을 위해 1개/120°회전량을 갖는 극좌표 입력을 생각하기로 한다. 이러한 극 좌표 입력에 적합한 신경 회로망으로 그림 2와 같은 구조를 갖는 신경 회로망을 설계하기로 한다. 그림 2에있는 공간상에서 설계된 신경망의 경우, x,y,z 축에서 (x,y,z)점을 입력층으로 (0,0,0)점을 출력층으로 나머지를 은닉층으로 볼때, x-y 평면, y-z 평면, z-x 평면의 가중치 (weight) 들이 모두 대칭임을 알수있다. 그러므로 입력패턴이 어떠한 회전을 하더라도 동일한 출력을 얻을 수있다. 이를 위하여 세 평면의 가중치가 모두 동일한 변수라는 제한조건이 필요하며, 이러한 신경회로망을 제한 신경회로망 (constrained neural network) 이라 부른다 [8].

이러한 제한 신경회로망의 설계에서는 우선 학습시켜야 할 패턴의 구조로 부터 가중치사이의 상호관계 즉 제한조건을 알아 낸후 이들을 고려하여 연결형태를 구성하여야 한다. 그림 2의 경우 가중치의 수는 12개 이나 독립적 가중치의 수는 w_1, w_2, w_3, w_4로 4개임을 알 수있다. 이 제한 신경망을 평면으로 펼쳐보면 그림 3과 같다. 또다른 특별한 형태의 신경망 구조는 (0,0,0)점의 출력층 밑에 은닉층이 숨어있고 완전연결인 경우로 공간상에서 펼쳐보면 그림 4와 같은 원뿔형태의 신경회로망이다. 단, 한 은닉층에는 한 개의 뉴론만이 존재한다. 만약 은닉층이 1개인경우, 입력이 6개, 출력이 1개인 네트워크의 가중치수는 13개이고 회전량에 불변일 독립적인 가중치의 수는 3개이다. 결과적으로 회전량에 불변인 제한 신경회로망은 전처리 (preprocessor) 과정 없이도 패턴인식을 할 수 있으며, 독립적인 가중치의 수를 감소시킴으로써 학습시간을 감소 시키는데 효과적이었다 [9]. 또한 입력의 수를 증가시켜서 회전량의 자유도 (degree of freedom)를 임의로 조절할 수 있는 장점이 있다.

3. 제한 신경회로망에 적합한 새로운 학습 알고리즘 제안

앞 절에서는 입력패턴의 회전량에 불변인 새로운 연결형태의 제한신경회로망을 제시하였다. 그러나 제한 신경망에서 기존의 학습규칙인 오류 역전파 알고리즘을 적용할 경우 가중치의 제한성을 유지할 수 없다. 따라서 여기에서는 이들 연결형태에 합리적인 새로운 학습 알고리즘을 제시하고자 한다. 먼저 다음과 같은 신경망을 가정한다.

1. 완전연결 (full connection)된 다층 퍼셉트론이다.
2. 같은층 안에서는 뉴론(neuron)끼리 연결되어 있지 않다.
3. 되먹임 (feedback)이 없다 ($W_{ii} = 0$).

앞으로 제안할 학습알고리즘은 다층퍼셉트론에 적용 가능하나 수식을 간단히 표현하기 위하여 은닉층이 두개인 신경망에 관한 알고리즘을 유도하기로 한다. 그림 5에 두개의 은닉층이 있는 신경망이 표현되어 있다. 즉, J층, K층, L층, M층의 4개의 층으로 구성되어 있다. 다층 퍼셉트론의 모든 가중치가 몇 개의 독립적인 가중치의 함수로 나타낼 수 있다면, 다음과 같은 제한 조건을 갖는다.

$$W_{jk} = G_{jk}(W_i)$$
$$W_{kl} = G_{kl}(W_i) \tag{1}$$
$$W_{lm} = G_{lm}(W_i)$$

여기서 G는 임의 함수이고, W_i 는 독립적인 가중치 벡터들이다 ($i = 1,2, \ldots ,n$). j, k, l, m은 J층, K층, L층, M층에서의 각 뉴우런들이고, W_{jk}, W_{kl}, W_{lm}은 j뉴우런과 k뉴우런사이의 가중치, k뉴우런과 l뉴우런사이의 가중치, l뉴우런과 m뉴우런사이의 가중치이다.따라서 독립적인 가중치의 학습규칙만 유도할 수 있다면, 제한 조건에 의해 모든 가중치의 변화량을 다음과 같이 구할 수 있다.

$$D_p W_{jk} = -\eta \frac{\partial G_{jk}}{\partial W_i} D_p W_i \tag{2}$$

각 패턴 P에 대한 오차함수 또는 에너지 함수 E_P^s 를 다음과 같이 정의한다.

$$E_P^s = \frac{1}{2} \sum_j [\, T_j - f(\sum_k W_{jk} O_k + \sum_l W_{jl} O_l + \sum_m W_{jm} O_m)\,]^2 \tag{3}$$

일반적으로 초기에 실제출력값은 목표값 T_j와 일치하지 않을것이다. 여기서 O는 각 뉴론에서의 출력을 나타내며, f(.)은 유니트에서의 비선형 전달함수이다. 비선형 전달함수로는 sigmoid함수를 사용하며 출력층에서의 출력 O_j는 다음과 같다.

$$O_j = \frac{1}{1 + e^{-(net_j + \theta_j)/\theta_0}} \tag{4}$$

여기서 θ_j 는 임계치이며, θ_0는 기울기를 변화시키는 상수이다. 각 뉴우런에서의 입력 net_j는 다음과 같이 정의한다.

$$net_j = \sum_k W_{jk} O_k + \sum_l W_{jl} O_l + \sum_m W_{jm} O_m \tag{5}$$

독립적인 가중치 W_i 에 대해 에너지 함수 E_P^s 를미분 하면 다음과 같다.

$$\frac{\partial E_P^s}{\partial W_i'} = -\sum_j [\, T_j - f()\,] \cdot f'() \cdot$$
$$[\, (\sum_k \frac{\partial W_{jk}}{\partial W_i'} O_k + \sum_k \frac{\partial O_k}{\partial W_i'} W_{jk}) + (\sum_l \frac{\partial W_{jl}}{\partial W_i'} O_l + \sum_l \frac{\partial O_l}{\partial W_i'} W_{jl}) + \sum_m \frac{\partial W_{jm}}{\partial W_i'} O_m \,]$$
$$= -\delta_j \Delta^j_j D_j - \delta_j \Delta^j_k D_k - \delta_j (\Delta^j_k \Delta^k_l + \Delta^j_l) D_l \tag{6}$$

여기서,

$$\Delta^j_k \equiv 1 \qquad , j=k$$
$$\Delta^j_k \equiv \sum_k W_{jk} f'(\sum_l W_{kl} O_l + \sum_m W_{km} O_m) \qquad , j \neq k$$
$$\delta_j = \sum_j [\, T_j - f()\,] \cdot f'()$$
$$D_j = \sum_k \frac{\partial W_{jk}}{\partial W_i'} O_k + \sum_l \frac{\partial W_{jl}}{\partial W_i'} O_l + \sum_m \frac{\partial W_{jm}}{\partial W_i'} O_m$$

이다. 일반적인 항으로 표현하기위해 j, k=j-1, l=j-2 라 하면 다음과 같다.

$$\frac{\partial E_P^s}{\partial W_i'} = -\delta_j \Delta^j_j D_j - \delta_j \Delta^j_{j-1} D_{j-1} - \delta_j (\Delta^j_{j-1} \Delta^{j-1}_{j-2} + \Delta^j_{j-2}) D_{j-2} \tag{7}$$

여기서,

$$D_J = \sum_{n=1}^{N} \sum_{J-n} \frac{\partial W_{j(j-n)}}{\partial W_i'} O_{j-n} \quad (\text{N : residual layers at J layer })$$

$$\Delta_{j-1}^{j} \equiv \sum_{j-1} W_{j(j-1)} f' (\sum_{n=1}^{N} \sum_{J-1-n} W_{(j-1)(j-1-n)} O_{j-1-n})$$

$$\frac{\partial W_{j(j-n)}}{\partial W_i'} = G_{j(j-n)}'(W_i) \frac{\partial W_i}{\partial W_i'}$$

이다

학습이란 수학적으로 총 에너지 함수가 0이 되는 가중치를 구하는 것이다. 수렴곡면이 전체적으로 오목 (globally concave)하다면 steepest descent 방법을 사용하여 가중치를 변화시켜 최소점으로 찾아갈 것이다. 따라서, 독립적인 가중치의 변화량을 다음과 같이 정의한다.

$$D_p W_i = - \frac{\partial E}{\partial W_i'} \tag{8}$$

이 학습방법의 특이한 점은 독립적인 가중치에 대한 시스템의 오차를 역방향으로 추정할 경우 이에 관계하는 모든 뉴런으로 부터 역방향으로 추정된다는 것이다. 이는 제한 신경망이 완전 연결된 다층 퍼셉트론이기 때문이다. 여기서 유도된 학습법칙은 기존의 알고리즘을 일반화 한것이라 할 수 있다. 예를 들면 기존의 오류 역전파 알고리즘은 모든 가중치가 독립적이고 근접한 층사이만 연결된 구조를 갖는경우, 제한 신경망의 학습방법은 기존의 학습방법과 일치한다.

4. 모의 실험 결과 및 고찰

앞에서 제안한 신경회로망의 재미있는 응용중의 하나는 동전 인식 (coin recognition) 시스템이다. 본 절에서는 그림 6과같은 500원 동전의 앞면과 뒷면을 회전량에 관계없이 구별 할 수있음을 보이고 그 효율성을 검증해보기로 한다.

모의 실험에서는 1개/30° 회전량을 갖는 3층구조의 제한신경회로망을 구성하였다. 첫째, 그림 2와 같은 형태의 제한신경망일경우, 입력층과 은닉층의 뉴론수는 각각 12개 이며 출력층의 뉴론수는 1개이다. 가중치의 수는 총 156개 이며 회전량에 불변인 시스템을 구성하기 위한 가중치 사이의 제한 조건은 다음과 같다.

$$\begin{aligned} W_{kl} \text{ 의 제한 조건}: \; & W_{11} = W_{22} = \cdots = W_{1212} \\ & W_{12} = W_{23} = \cdots = W_{121} \\ & \cdots\cdots\cdots\cdots \\ & W_{112} = \cdots\cdots = W_{1211} \\ W_{jk} \text{ 의 제한 조건}: \; & W_{11} = W_{12} = \cdots = W_{112} \end{aligned} \tag{9}$$

따라서 독립적인 가중치의 수는 13개로 줄일 수있다. 둘째, 그림 4와 같은 제한신경망일경우 가중치의 수는 총 25개이고 다음과 같은 제한조건에 따라 독립적 가중치의 수는 3개이다.

$$W_{jk} \text{ 와 } W_{kl} \text{의 제한 조건}: W_{11} = W_{12} = \cdots = W_{112} \tag{10}$$

입력패턴은 500원 동전의 앞면과 뒷면, 출력은 각각 1.0 으로 하였다. 입력패턴은 ccd카메라와 DT 2862 frame grabber 로 이미지를 받았다. 80 × 80으로 데이타를 받았으며, 1개/30° 회전량에 적합한 데이타로 평균화 (Averiging) 하였다.

다음은 두가지 형태의 신경회로망을 모의 실험 하였다. 식9과 10의 제한조건에따라 3절에서에서 유도한 새로운 학습 알고리즘을 적용하였다. 그림 7은 첫번째모델에 의한 학습곡선이다. 전처리 과정이 없음에도 불구하고 빠른 수렴성을 갖는 것을 알 수있다. 그림 8은 두번째 모델에 의한 학습곡선이다. 첫번째 모델에 비해 빠른 수렴성과 적은 오차를 보이고 있다. 이는 최적화 해야할 독립 가중치의 수가 적고 연결형태가 완전 연결 (full connection) 되어 있는데 기인한 것으로 보인다. 이 두 모델들은 입력 패턴들이 복잡할 경우 회전량의 자유도를 임의로 조절 할 수도 있다.

5 결론

본 연구에서는 패턴인식 분야 중 회전량에 불변인 인식 시스템을 구성하기 위하여 제한 신경망를 설계하였다. 또한, 이를 학습시키기 위한 새로운 학습 알고리즘을 개발하였으며, 이에 대한 응용사례로 동전인식에 관한 모의실험을 하였다.

회전량에 불변인 제한 신경망 시스템에 대하여 다음과 같은 결론을 얻었다. 새로 구성된 제한 신경망은 전처리 과정이 필요 없으며, 가중치의 제한조건은 독립가중치의 감소로 학습시간을 줄일 수 있다. 입력뉴론수를 조정하여 회전량의 자유도를 임의로 조정할 수 있다.

본 논문에서 제안한 신경망은 패턴인식에 좋은 성능을 보였으며 앞으로 보다 다양한 시스템에 적용되리라 기대된다.

참고문헌

[1] Pao,Y-H, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Co., Inc., 1989.

[2] Rumelhart, D.E. and Mc. Clelland, J.L., Parallel Distributed Processing, MIT Press, 1986.

[3] K. Fukushima, S. Miyake, and T. Ito, "Neocognitron: A neural network model for a mechanism of visual pattern recognition," IEEE Trans. Syst., Man, Cybem., vol. SMC-13, no. 5, pp. 826-834, 1983.

[4] G. A. Carpenter and S. Grossberg, "The ART of adaptive pattern recognition by a self-organizing neural networks," IEEE Computer, vol. 21, no. 3, pp. 77-88.1988.

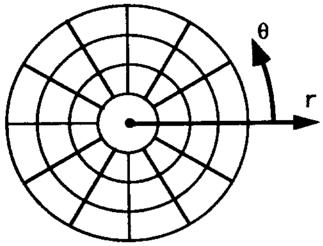
[5] B. Widrow, R. G. Winter, and R. A. Baxter, "Layered neural nets for pattern recognition," IEEE Trans. Acoust., Speech, Signal Process, vol. 36, no. 7, pp. 1109-1118,1988.

[6] M. B. Reid, L. Spirkovska, and E. Ochoa, "Rapid training of high-order neural networks for invariant pattern recognition," in Proc. Int. Joint Conf. Neural Networks, vol.1, 1989, pp. 689-692.

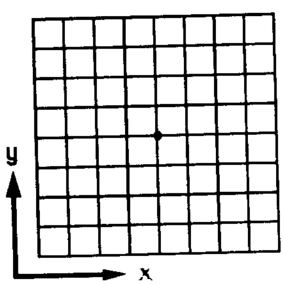
[7] M. Fukumi, S. Omatu, F. Takeda, and T. Kosaka, "Rotation - Invariant Pattern Recognition System with Application to Coin Recognition," IEEE Trans. Neural Networks., vol. 3, 1992, pp 272-279.

[8] 나 희 승, 박 영 진, "대칭 신경회로망과 그 응용에 관한 연구," 대한기계학회 논문집, 제6권 제7호, pp. 1322-1331,1992.

[9] Y. Park and H. S. Na, "Symmetric Neural Network and its Examples," IJCNN, Baltimore, 1992, pp 413-418



(a) 직교좌표 입력형식



(b) 극좌표 입력형식

그림 1. 입력패턴의 자료 입력형식

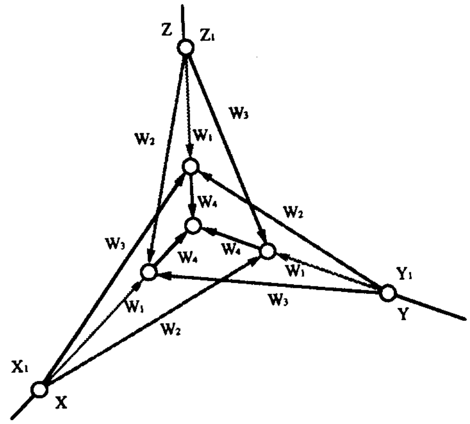


그림 2. 회전량에 불변인 제한 신경회로망 I

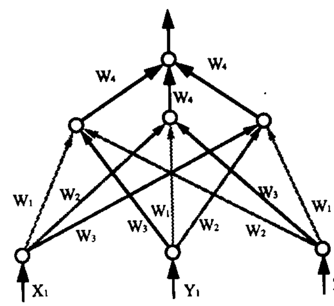


그림 3. 그림 2.를 평면상에서 펼쳐본 제한 신경회로망 I

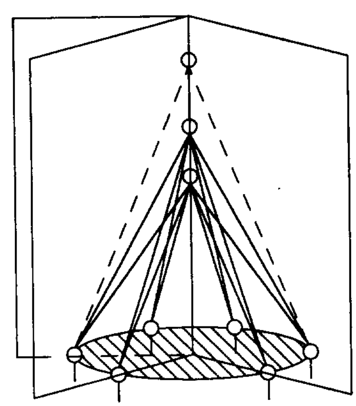


그림 4. 회전량에 불변인 제한 신경회로망 II

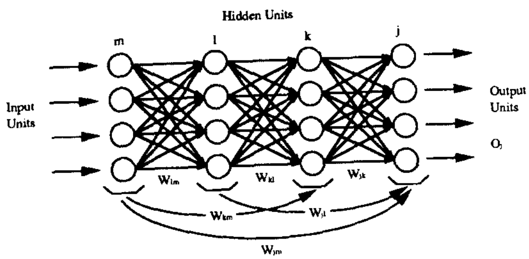


그림 5. 일반화된 제한 신경회로망모델

(a) 앞면

(b) 뒷면

그림 6. 500원 동전의 앞면과 뒷면

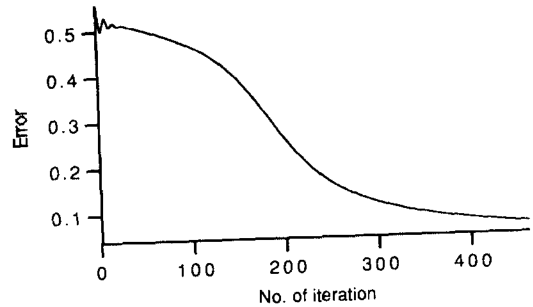


그림 7. 제한 신경회로망 I의 학습곡선

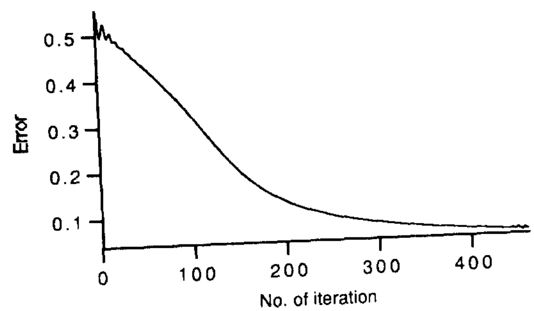


그림 8. 제한 신경회로망 II의 학습곡선