순환하는 레이어 연결을 갖는 개선된 On-line 신경회로망의 설계

여성원, 이종호
인하대학교 전기공학과

Design of an Improved On-line Neural Network with Circulating Layer Connections

Seong-Won Yeo, Chong-Ho Lee
Department of Electrical Engineering, Inha University

Abstract - In this paper, an improved on-line neural network model is suggested. This neural network is designed to store and recall sequence of key strokes in on-line. The network stores incoming patterns as weight connections between series of layers. The layer has a 2-dimensionally distributed neurons where the location of neurons are relevant to the actual location of computer keyboard. To store longer patterns, the network has circulating layer connections and different patterns can be superposed on the same layer. Also, when the patterns are stored over the layers, the starting layer is not fixed but changed by the characteristics of patterns to increases network capability. The ways how to choose the starting layer during the store and recall process are investigated.

1. 서 론

인간의 두뇌구조를 모방하여 구성한 인공 신경망(artificial neural networks)은 뉴런(neuron)들의 집합체인 레이어(layer)와 이들간을 상호 연결하는 가중치(weight)로 구성되며 처리해야할 정보는 이 가중치를 조정하는 학습(learning)과정을 통하여 저장된다. 그러나 현존하는 신경망 모델은 레이어에 포함된 뉴런의 수가 많아지게 되면 가중치의 수가 급속하게 증가하며 이에 따라 입력패턴으로부터 출력패턴을 생성하는데 많은 계산이 필요할 뿐 아니라 학습을 위해서는 더욱 많은 시간을 소모하게 되므로 온라인(on-line) 처리가 어려워진다. 본 논문에서는 2차원으로 분포된 많은 수의 뉴런들이 레이어 상에 존재하나 가중치의 수를 제한하고 가중치가 연결될 대상 뉴런을 필요에 따라 변경 가능하도록 함으로서 계산량을 줄여 온라인으로 동작이 가능한 새로운 신경망 모델을 제안한다. 제안하는 신경망 모델은 순환하는 레이어 연결 구조를 가지며 패턴이 다단계의 레이어에 걸쳐 저장되고 사용자의 요청에 따라 저장된 패턴을 재생한다. 이를 수행함에 있어 저장효율을 높이기 위해 여러 개의 레이어 중 적절한 위치를 선택하여 패턴을 저장시키며 재생 시에는 입력패턴을 근거로 저장된 위치를 찾아낸다. 제안하는 신경망 모델의 응용 예로서 컴퓨터 키보드 입력을 분산 저장시키고 재생시키는 과정을 보인다.

2. 본 론

2.1 구조

2.1.1 레이어의 연결과 패턴구조

제안하는 신경망의 레이어 연결구조는 그림 1과 같다.

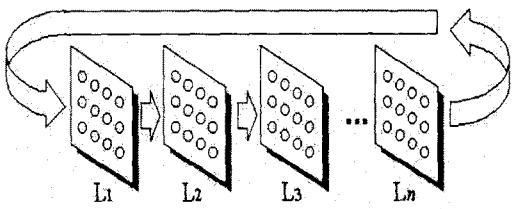

그림 1. 레이어의 연결구조

제안하는 신경망은 n개의 레이어를 가지고 있으며 레이어에는 2차원으로 분포된 뉴런들이 존재한다. 각 레이어상의 뉴런은 다음 레이어의 뉴런과 가중치로 연결되며 n번째 레이어의 뉴런이 첫 번째 레이어의 뉴런과 연결되는 순환구조를 갖는다. 신경망에 입력되는 패턴과 출력되는 패턴은 다음과 같이 나타낼 수 있다.

$$\text{입력패턴: } P=[\ p_1\ \ p_2\ \cdots\ p_k] \qquad (1)$$
$$\text{출력패턴: } O=[\ o_1\ \ o_2\ \cdots\ o_k]$$

입·출력 패턴은 k개의 세그먼트로 나뉘어지며 세그먼트 별로 레이어에 할당되어 처리된다. 세그먼트의 구성은 다음과 같다.

$$p_i=\begin{bmatrix} x_{i1} & y_{i1} \\ x_{i2} & y_{i2} \\ \vdots & \vdots \\ x_{is} & y_{is} \end{bmatrix} \qquad i=1,2,\cdots k \qquad (2)$$

여기서 x_{ij}, y_{ij}는 i번째 세그먼트에 의해 레이어 상에서 활성화될 j번째 뉴런의 x, y 좌표값이며 s는 세그먼트의 용량이다.

신경망에 입력패턴이 인가되어 출력패턴이 재생될 때에는 처리될 레이어들의 시작위치를 나타내는 색인(index)이 정해지고 이로부터 순서대로 레이어에 세그먼트가 할당되어 처리된다. 또한 k>n인 경우에도 레이어가 순환구조를 이루므로 한 레이어 상에 두 개 이상의 세그먼트가 중첩되어(superposed) 처리된다. 그림 2는 n=6, k=8, index = 4일 때 레이어에 세그먼트 별로 패턴이 할당되는 예를 보인다.

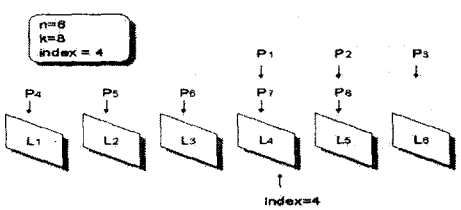

그림 2. 레이어별 세그먼트 할당 예

2.1.2 뉴런과 가중치

제안한 신경망 모델에서 사용되는 뉴런의 활성화 함수(activation function)와 가중치는 패턴의 재생과 온라인 동작을 위해 기존의 신경망과는 현저히 다른 구조를 갖는다.

뉴런의 활성화 함수는 다음과 같이 주어진다.[2]

$$f(net)=\begin{cases}1, & 1-\delta<net<1+\delta\\ 2, & otherwise\end{cases} \tag{3}$$

즉 net값이 1 근방일 때만 1을 출력하게 된다. 이를 그림으로 나타내면 그림 3과 같다.

뉴런은 가중치가 전단계 레이어와 연결된 것이 아니라 다음 단계의 레이어와 연결되어 출력값을 전달하는 outstar[1]구조를 갖는다. 따라서 전단계에서 누적된 net값은 활성화 함수를 취하여 가중치의 곱으로 다음 단계의 레이어로 전달되며 연결된 뉴런의 net값을 누적시킨다.

뉴런의 가중치는 다음단계의 레이어와 완전히 연결된(fully connected) 구조를 갖지 않고 10여 개의 제한된 수만큼을 갖는다. 그 대신 가중치가 연결되는 뉴런은 고정되어 있지 않고 필요에 따라 다음단계에 있는 레이어상의 임의의 뉴런과도 연결이 가능하다. 초기 뉴런의 가중치들은 아무런 뉴런과도 연결되어 있지 않고 그 값도 0이 되나 패턴을 저장시킴에 따라 필요한 수만큼 가중치를 할당하여 사용하며 만일 가중치가 부족한 경우에는 값이 작은 가중치의 연결을 끊고 새로운 뉴런과 연결시킨다. i번째 레이어상의 x, y위치에 있는 뉴런을 μ로 나타내면 이 뉴런의 가중치는 다음과 같다

$$W_{i,x,y}=\ W^{(\mu)}=\begin{bmatrix} w_1^{(\mu)} & x_1^{(\mu)} & y_1^{(\mu)}\\ w_2^{(\mu)} & x_2^{(\mu)} & y_2^{(\mu)}\\ \vdots & \vdots & \vdots\\ w_c^{(\mu)} & x_c^{(\mu)} & y_c^{(\mu)}\end{bmatrix} \tag{4}$$

여기서 $i=1,2,\cdots n$ 이며 n은 레이어의 개수, c는 하나의 뉴런이 가지는 제한된 가중치의 수를 나타낸다. 이러한 레이어 상의 뉴런의 연결을 그림 4에 나타내었다.

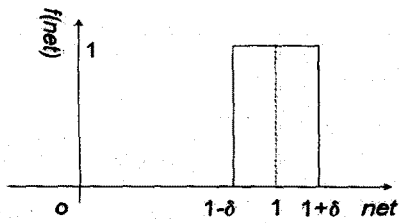

그림 3. 활성화 함수

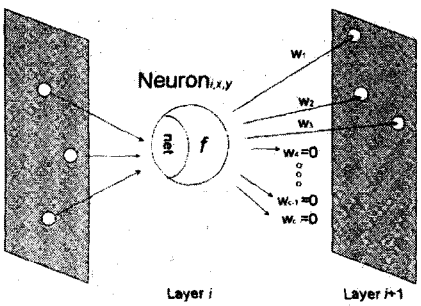

그림 4. 뉴런의 구조 및 연결

2.2 패턴의 저장

제안하는 신경망에서 패턴의 저장과 재생은 정해진 규칙에 따라 동시에 이루어진다. 패턴의 저장은 레이어간의 가중치의 연결로서 저장되는데 입력패턴이 인가되었을 때 연속하는 두 개의 세그먼트와 그들이 가리키는 두 개의 레이어상의 각각의 뉴런들을 가중치로 연결시키고 향후에 같은 입력이 들어오면 다음의 식에 따라 가중치를 증가시킨다.

$$w_j^{(\mu)}=w_j^{(\mu)}+\alpha\left(\frac{1}{N}-w_j^{(\mu)}\right) \tag{5}$$

여기서 $j=1,2,\cdots,c$(c는 뉴런이 갖는 가중치의 수)이며 α는 0 에서 1사이의 의 값을 갖는 학습계수이고 N은 세그먼트의 가 포함할 수 있는 정보의 개수 s중에 실제로 채워진 정보의 개수이다. 다시 말하면 N은 세그먼트가 레이어에 할당되었을 때 활성화 될 수 있는 뉴런의 개수가 된다. 단 가중치를 연결함에 있어 가중치의 순서가 세그먼트가 가리키는 순서와 동일하게 한다.

식(5)에서 나타내듯이 학습이 반복적으로 이루어지면 가중치는 $1/N$의 값으로 수렴하게 된다. 만일 3개의 뉴런이 활성화되어 각각의 가중치가 1/3으로 수렴해간다면 이들 3개의 뉴런으로부터 누적된 다음 레이어의 뉴런의 net값은 1에 수렴하게 되고 식(3)에 의하여 다음단의 뉴런도 성공적으로 활성화되어 출력을 전달할 수 있게 된다.

그림 5는 간단한 패턴의 저장 예를 보인다. 활성화된 뉴런을 검게 표시하였으며 그림에서 레이어 1은 3개의 뉴런에 의해 1/3의 가중치를 가지며 레이어 2는 2개의 뉴런에 의해 1/2의 가중치를 갖는다. 또한 패턴의 입력에 따라 같은 레이어에서 활성화되는 뉴런의 순서를 표시하였다. 이 경우에 활성화 순서를 고려한 네트웍에 인가되는 세그먼트들은 다음과 같다.

$$p_1=\begin{bmatrix}3&2\\2&3\\5&2\end{bmatrix},\quad p_2=\begin{bmatrix}3&3\\5&2\\0&0\end{bmatrix},\quad p_3=\begin{bmatrix}4&3\\1&4\\0&0\end{bmatrix} \tag{6}$$

패턴의 저장은 순환하는 연결을 갖는 레이어 중 임의의 레이어를 시점으로 저장시킬 수 있으나, 각 레이어를 시점으로 패턴을 저장하였을 때 이미 할당된 레이어와 중복되는 가중치의 개수를 계산하여 가장 중복이 적은 레이어를 선정하여 저장시킨다. 또한 두 개의 패턴이 전반부는 유사하나 후반부가 다른 경우에는 각각을 서로 다른 시점으로 하여 레이어에 저장시킨 뒤 재생을 할 때에는 전반부의 입력만 인가된 상태에서는 두 패턴을 모두 재생하고 재생이 진행됨에 따라 유사한 패턴 한 개만 재생한다.

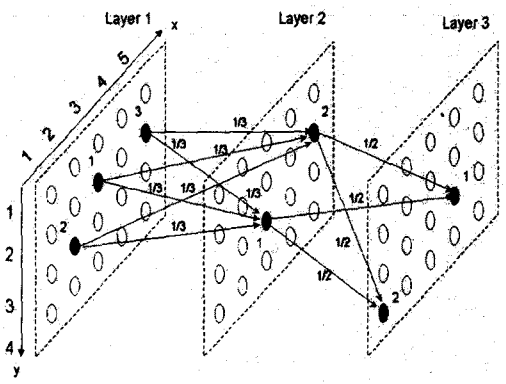

그림 5. 패턴의 저장 예

2.3 패턴의 재생

제안된 신경망은 학습된 패턴의 전반부를 적절한 레이어에 인가함으로서 레이어간의 전달을 통해 저장된 모든 세그먼트를 재생할 수 있다. 그러나 모든 가중치를 특정 레이어에 편중하지 않고 골고루 할당하려는 저장방법의 특성상 패턴이 어느 레이어를 시점으로 저장되었는지를 패턴입력 초기에 판단하여야 한다. 본 논문에서는 입력 패턴 초반부 세그먼트와 각 레이어의 가중치를 가지고 재생하려는 패턴이 저장된 레이어의 시점을 찾아내는 다음과 같은 방법을 제안한다.

방법 : 초반부의 세그먼트 $\mathbf{p}_1$과 레이어 $L_1, L_2, \cdots L_n$의 할당된 가중치로부터 다음세그먼트로 진행될 가능성이 높은 색인(index) j들을 결정

(i) $j \leftarrow 1$

(ii) $k \leftarrow 2$

(iii) $\boldsymbol{W}^{(k)}$를 L_j상의 $\mathbf{p}_1$이 가리키는 k번째 뉴런의 가중치($k=1,2,\cdots N$)로 정의

$$\boldsymbol{W}^{(1)}=\begin{bmatrix} w_1^{(1)} & x_1^{(1)} & y_1^{(1)} \\ w_2^{(1)} & x_2^{(1)} & y_2^{(1)} \\ \vdots & \vdots & \vdots \\ w_c^{(1)} & x_c^{(1)} & y_c^{(1)} \end{bmatrix}$$

$$\boldsymbol{W}^{(k)}=\begin{bmatrix} w_1^{(k)} & x_1^{(k)} & y_1^{(k)} \\ w_2^{(k)} & x_2^{(k)} & y_2^{(k)} \\ \vdots & \vdots & \vdots \\ w_c^{(k)} & x_c^{(k)} & y_c^{(k)} \end{bmatrix} \quad (7)$$

(iv) $\boldsymbol{W}^{(1)}$과 $\boldsymbol{W}^{(k)}$에 대하여 거리를 나타내는 함수를 정의하고 다음을 계산한다.

$$d_{1k}(l,m)=\sqrt{(x_l^{(1)}-x_m^{(k)})^2+(y_l^{(1)}-y_m^{(k)})^2} \quad (8)$$
$$l,m=1,2,\cdots,c$$

$$\boldsymbol{D}=\begin{bmatrix} d_{1k}(1,1) & d_{1k}(1,2) & \cdots & d_{1k}(1,c) \\ d_{1k}(2,1) & d_{1k}(2,2) & \cdots & d_{1k}(2,c) \\ \vdots & \vdots & \ddots & \vdots \\ d_{1k}(c,1) & d_{1k}(c,2) & \cdots & d_{1k}(c,c) \end{bmatrix} \quad (9)$$

$\boldsymbol{D}$의 각 행(row)의 성분 중 최소값과 그 색인(index) 다음과 같이 나타낸다.

$$\boldsymbol{MD}_k=\begin{bmatrix} \min D_1 & IndexD_1 \\ \min D_2 & IndexD_2 \\ \vdots & \vdots \\ \min D_c & IndexD_c \end{bmatrix} \quad (10)$$

(v) $k \leftarrow k+1$, $k<=N$까지 과정(iii) 반복

(vi) 과정 (iv)를 수행하여 얻어진 MD_k를 조사하여 최소값 $\min D_i$가 동일한 행(row)에서 나타나고 이때의 색인 $IndexD_i$가 연속하는 값을 가지면 다음 세 그먼트로 진행될 가능성이 높다. 즉 최소값이 나타나는 부분에 대하여 색인값들의 분산(variance)이 일정한 문턱값(threshold value)보다 작은 경우 j를 선택

(vii) $j \leftarrow j+1$, $j<=n$까지 과정(ii) 반복

예제 : 그림 5의 간단한 경우 세그먼트들은 식(6)과 같이 주어진다. 이때 $\mathbf{p}_1$에 대하여 위의 방법을 적용하였을 때 레이어1이 올바른 시작위치임을 보인다.

레이어 1의 뉴런 1,2,3의 가중치 $\boldsymbol{W}^{(1)}, \boldsymbol{W}^{(2)}\boldsymbol{W}^{(3)}$가 다음 식의 굵은 글씨로 표기된 것 과 같은 공통부분을 갖고 나머지는 임의의 값을 취한다. 편의상 $\boldsymbol{W}^{(1)}$의 성분 중 w성분은 생략하였고 $c=8$로 설정하였다.

$$\boldsymbol{W}^{(1)}=\begin{bmatrix} 3 & 3 \\ 4 & 3 \\ 2 & 1 \\ 4 & 5 \\ 2 & 2 \\ 1 & 2 \\ \mathbf{3} & \mathbf{3} \\ \mathbf{5} & \mathbf{2} \end{bmatrix}, \quad \boldsymbol{W}^{(2)}=\begin{bmatrix} \mathbf{3} & \mathbf{3} \\ \mathbf{5} & \mathbf{2} \\ 2 & 3 \\ 5 & 2 \\ 4 & 1 \\ 5 & 2 \\ 1 & 4 \\ 2 & 4 \end{bmatrix}, \quad \boldsymbol{W}^{(3)}=\begin{bmatrix} 1 & 4 \\ 4 & 4 \\ 4 & 4 \\ \mathbf{3} & \mathbf{3} \\ \mathbf{5} & \mathbf{2} \\ 4 & 2 \\ 4 & 4 \\ 3 & 4 \end{bmatrix} \quad (11)$$

이 경우 $\boldsymbol{W}^{(1)}$에 대한 $\boldsymbol{W}^{(2)}\boldsymbol{W}^{(3)}$의 거리에 대한 MD_2, MD_3는 다음과 같이 같은 위치에서 최소값과 연속적인 색인값이 나타난다.

$$\boldsymbol{MD}_2=\begin{bmatrix} 0.000 & 1 \\ 1.000 & 1 \\ 2.000 & 3 \\ 2.236 & 1 \\ 1.000 & 3 \\ 1.414 & 3 \\ \mathbf{0.000} & \mathbf{1} \\ \mathbf{0.000} & \mathbf{2} \end{bmatrix}, \quad \boldsymbol{MD}_3=\begin{bmatrix} 0.000 & 4 \\ 1.000 & 2 \\ 2.236 & 4 \\ 1.000 & 2 \\ 1.414 & 4 \\ 2.000 & 1 \\ \mathbf{0.000} & \mathbf{4} \\ \mathbf{0.000} & \mathbf{5} \end{bmatrix} \quad (12)$$

2.4 제안한 신경망의 응용

제안한 신경망의 응용으로 컴퓨터의 키보드 영문과 숫자 자판배열을 레이어의 뉴런배열과 일치시킨 뒤 사용자의 키보드 입력에 따라 자주 사용되는 단어를 실시간으로 재생하는 환경을 구현하였다. 그림 6은 응용에 사용된 레이어상의 실제 뉴런의 배치이다. 그림 7에서는 장문을 입력한 뒤 이중 wizet란 단어와 wizard란 단어를 사용자 입력에 따라 구분하는 과정을 보인다.

3. 결 론

본 논문에서는 순환하는 레이어 연결을 갖는 새로운 신경망 모델을 제안하고 그 응용 예를 보였다. 제안한 신경망 모델은 한 레이어에 중첩저장이 가능하며 실시간으로 저장과 재생이 이루어지므로 계산시간과 저장용량에 있어서 높은 성능을 보인다. 향후에는 제안하는 신경망 모델의 용량에 대한 정량적 분석과 오차를 극복하여 재생을 시키는 부분에 대한 연구가 이루어지고 이를 확장하여 지능적 키보드 입력처리기를 구현할 예정이다.

[참 고 문 헌]

[1] M. T. Hagan, H. B. Demuth, M. Beale, "Neural Network Design", PWS Publishing Company, pp.13-5 - 13-9, 1996

[2] 여성원, 이종호, "시변패턴의 저장과 인식을 위한 On-line연상 메모리의 설계", 하계학술대회 논문집, B권, pp.1395-1397, 1996

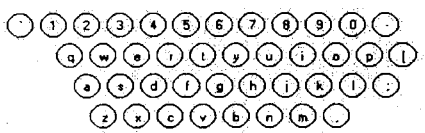

그림 6. 키보드와 관련된 레이어의 구성

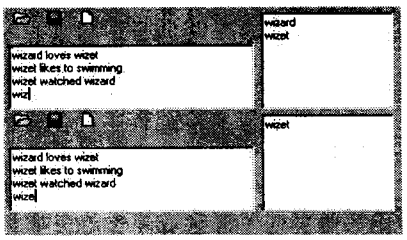

그림 7. 출력 패턴의 선택 예