

반경제한 퍼셉트론을 이용한 온라인 필기 영문자의 인식[†]

조 성배, 김 정현, 김 진형

한국과학기술원 전산학과 및 인공지능연구소

On-line Recognition of Boxed Handwritten English Characters Using Radius Limited Perceptron

Sung-Bae Cho, Jeong-hyun Kim and Jun Hyung Kim

Center for Artificial Intelligence Research and Computer Science Department, KAIST

요 약

본 논문에서는 태블릿을 통해서 입력된 손으로 쓴 영문자 및 숫자를 인식하기 위하여 신경망 모델의 일종인 반경제한 퍼셉트론을 제안하고 실험을 통하여 그 유용성을 입증한다. 반경제한 퍼셉트론은 기존 신경망 방식의 장점인 병렬분산 처리능력을 유지하면서 실시간 학습능력을 갖는데, 이는 온라인 필기인식 시스템에 유용한 모델이라고 할 수 있다. 실험에 의하면 제한된 학습 데이터에 대해서 사용자에 무관하게 높은 인식률을 내었으며, 오인식된 문자에 대해서도 즉시 학습시킬 수 있었다. 단지, 현재 설계된 모델에서는 학습이 진행됨에 따라 중간층 노드의 수, 즉 모델의 수가 계속 증가하는 문제가 있는데, 이를 해결하기 위한 모델관리기법에 대해서는 앞으로 좀더 연구되어야 할 것이다.

1 서론

키보드를 사용하는 대신에 필기에 의하여 컴퓨터에 정보를 입력하고자 하는 연구는 수년간에 걸쳐서 많은 사람들의 관심을 끌고 있다. 최근에는 하드웨어의 발전으로 저렴한 가격으로 노트북 크기의 노트패드 컴퓨터가 현실로 등장하고 있다. 하지만, 이러한 컴퓨터를 좀더 이상적인 것이 되도록 하기 위해서는 온라인의 손으로 쓴 문자를 인식할 수 있는 소프트웨어의 개발이 절실한 실정이다. 이를 위하여 여러 기관에서 오랫동안 연구를 계속해 오고 있는데 대표적인 것은 미국 IBM 연구소와 일본의 SONY사 및 영국의 Electronic Paper 프로젝트등을 들 수 있다. 이들은 각기 장기간 연구결과의 축적으로 나름대로 커다란 진보를 이루었다고 볼 수 있지만, 아직도 완전한 시스템의 구현은 요원한 상태이다.

이러한 소프트웨어 시스템의 주요한 요구사항은 크게 세가지로 높은 인식률, 실시간 응답속도 및 자연스러운 필기방식의 허용이라고 볼 수 있다. [4] 이와 같은 목적을 만족시키기 위해서는 필기 입력중에 발생하는 다양한 변형을 효과적으로 흡수할 수 있을 뿐만 아니라, 새로운 변형에 즉각적으로 적응해 나갈 수 있는 실시간 학습능력이 필요하다. 하겠다. 이 중에서 입력중의 다양한 변형을 처리할 수 있는 방법론으로 최근에 각광받고 있는 신경망[11]을 적용하고자 하는 연구가 활발히 진행되고 있다.

필기 문자 인식을 위한 가장 대표적인 신경망은 Fukushima

의 Neocognitron [5]을 들 수 있지만, 보다 실용적인 방법으로는 AT&T 등에서 연구하고 있는 시간지연 신경망 (Time-Delay Neural Network) [6] 이라든가, 네스터사의 RCE (Reduced Coulomb Energy) [12] 등이 있다. 국내에서도 연세대의 SEJONG Net [2]이나 ETRI의 연구 [7]가 있고, 서울대 [3]에서도 그 가능성에 대한 연구가 있었다. 실제로 상용 시스템도 개발되었는데, 네스터사의 NLS1000과 HNC사의 QuickStroke 등이 있다.

이제까지의 연구 중에서 실용화를 목적으로 한 것은 Kohonen Net 등을 전처리 과정에서 사용한 후에 Backpropagation 신경망을 사용하여 인식하는 것 [9]이 주류를 이루고 있다. 하지만, 이러한 방법의 가장 큰 문제는 기본적으로 초평면 (hyperplane)에 의하여 결정 경계면 (decision boundary)을 나누기 때문에 국부적인 정보의 변화에 즉각 반응하지 못하게 되어 실시간 학습이 불가능하다는 것이다. 이를 해결하기 위해서는 원을 기본 결정 경계면으로 하는 방법이 유용할 것인데, 이러한 모델로는 RBF (Radial Basis Function), RCE (Reduced Coulomb Energy) 등이 있다.

본 논문에서는 이러한 원을 기본으로 하는 신경망의 하나로 반경제한 퍼셉트론 (Radius-limited Perceptron)을 정의하고 이를 온라인의 손으로 쓴 문자를 인식하는 문제에 적용하여 그 결과를 알아 보기로 한다. 이 논문의 구성은 다음과 같다. 2장에서는 반경제한 퍼셉트론의 구조 및 이의 학습법칙에 대해서 알아보고, 이를 문자별로 나누어진 필기 영문자의 인식에 적용한 예를 3장에서 소개한다. 그리고 나서 4장에서는 그 결과를 논의하고 5장의 결론으로 맺기로 한다.

[†] 본 논문은 인공지능연구소 전산학과에서 지원을 받았다.

2. 반경제한 퍼셉트론

결정 경계면이 원으로 구성되는 반경제한 퍼셉트론의 주요 기능은 신경망의 특성인 병렬 분산처리 기능과 새로운 데이터의 실시간 학습기능을 들 수 있다. 이러한 모델의 이론적인 성능은 Perceptrons [8]에 자세히 기술되어 있는데, 생물학적인 관점에서 카테고리 학습의 일종으로 구현한 경우 [10]나 패턴인식의 관점에서 nearest neighbor 분류기의 일종으로 구현한 경우 [1]가 있었다. 본 장에서는 신경망의 관점에서 구현된 네트워크의 구조와 전이함수 및 그 학습법칙에 대해서 기술하고자 한다.

2.1. 구조와 전이함수

반경제한 퍼셉트론은 기본적으로 <그림 2.1>과 같이 입력층, 중간층, 출력층으로 구성된 3층 구조로 이루어져 있다. 입력층의 노드들은 N차원의 입력벡터를 저장하고, 출력층은 해당 입력이 속하는 패턴의 클래스를 나타내며, 중간층의 노드들은 이 사이의 사상관계를 나타내게 된다. 이들 사이의 연결방식은 비재환 (non-recurrent)의 전방향 네트워크 (feedforward network)로 입력층과 중간층 사이는 완전히 연결되어 있고, 중간층과 출력층 사이는 중간층 노드가 하나의 출력층 노드에만 연결되는 부분 연결의 형태를 띄고 있다. 이를 정리하면, 각각의 중간층 노드는 N차원 입력벡터를 나타내는 연결강도의 벡터 w_i 와 연결된 입력을 받아들여서 하나의 출력노드에 1의 연결강도로 연결되어 있음을 볼 수 있다.

각 노드의 전이함수는 중간층의 경우와 출력층의 경우에 각기 다르다. 먼저 중간층 노드의 경우에는 코호넨 (Kohonen)의 자기구성 특성지도에서와 같이 주어진 입력 x 와 해당 중간노드 w_i 의 거리 d_i 를 다음과 같이 계산한다.

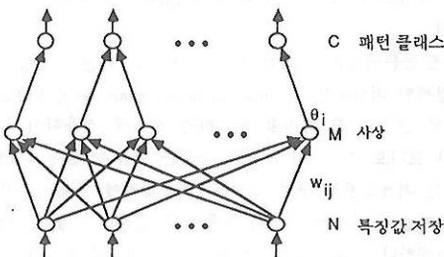
$$d_i = \sqrt{\sum_{j=1}^N (x_j - w_{ij})^2} \quad (1)$$

그리고 나서 이 노드의 출력 y_i 는 다음과 같이 1 또는 0의 값을 내게 된다.

$$y_i = f(d_i, \theta_i) = \begin{cases} 1, & \text{if } d_i \leq \theta_i \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

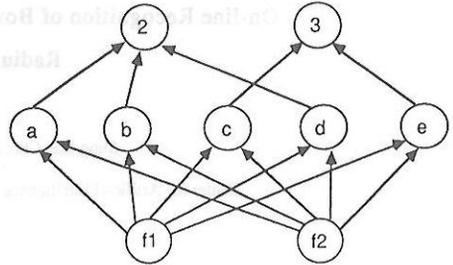
출력노드의 경우에는 단순히 논리적 OR기능을 하여 입력되는 중간노드의 값중에서 하나라도 1인 경우에 1을 출력한다.

반경제한 퍼셉트론내의 파라미터는 결국 중간노드에 연결된 연결강도 w_i 와 각 중간노드의 임계값 θ_i 인데, 이것은 N차원 공간상의 한 점 w_i 를 중심점으로 하여 반경이 θ_i 인 원형의 지역을 나타낸다. 즉, 전체 네트워크는 특징공간 (feature space)상에서 각

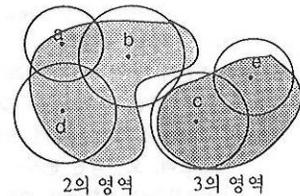


<그림 2.1> 반경제한 퍼셉트론의 기본구조

클래스를 여러개의 원형 지역으로 나타낸 것으로, 주어진 입력벡터의 클래스는 이것이 어느 원형지역에 속하는지를 병렬적으로 검토함으로써 알 수 있게 된다. 예를 들어서, 두개의 특징값을 이용하여 입력문자가 2인지 또는 3인지를 구별하는 네트워크가 <그림 2.2>와 같다고 할 때 각 노드의 기능을 2차원 특징공간 상에서 도시하면 <그림 2.3>과 같다.



<그림 2.2> 2와 3을 구별하는 네트워크의 예



<그림 2.3> 2차원 공간상의 가상적 결정 경계면들

2.2. 학습법칙

반경제한 퍼셉트론이 제대로 작동하기 위해서는 입력노드와 중간노드 사이의 연결강도 w_i 와 각 중간노드의 임계값 θ_i 를 적당히 조정하여야 한다. 반경제한 퍼셉트론은 이를 위하여 주어진 입력의 올바른 출력값을 제공하는 지도학습 (supervised learning) 방법을 사용하는데, 기본적으로 새로운 중간노드에 연결강도 w_i 를 할당하는 과정과 임계값 θ_i 를 조정하는 과정으로 이루어진다.

학습이 발생하는 경우는 다음의 두가지 경우를 고려할 수 있다. 첫째는 주어진 입력에 대하여 i번째 출력노드가 1이 되어야 하는데도 불구하고 0이 출력된 경우이다. 이 때는 새로운 중간노드를 하나 할당하여 주어진 입력벡터 x 로써 연결강도 w_i 를 지정하고 초기의 임계값으로 θ_i 를 지정하여 i번째 출력노드와 연결시킨다. 즉, 특징공간 상에서 해당 클래스가 주어진 입력을 포함하도록 새로운 원형의 영역을 할당하는 것이다. 둘째는 0이 출력되어야 할 노드가 1이 출력된 경우이다. 이 때는 이 출력노드와 연결되어 있는 중간노드 중에서 1을 출력하는 노드들의 임계값을 적당히 감소시킨다. 즉, 주어진 입력을 더이상 포함하지 않도록 원형 영역을 줄이는 것이다.

이상과 같은 학습과정을 좀더 쉽게 이해하기 위해서 앞에서 예로 들었던 2와 3을 분류하는 문제를 다시 생각하여 보자. 현재 어느정도 학습된 반경제한 퍼셉트론에 새로운 2가 입력되었다고 가정하자. 만일 이 입력이 기존의 2 클래스를 나타내는 영역에 포함되었다면 올바르게 인식된 것이므로 아무런 행동도 취하지 않는다. 만일 어떠한 원형 영역도 해당 입력을 포함하지

문장에 대해서 서로 다른 사용자가 어떤 인식률을 나타내는지 알아본 후에 특정 사용자에게 대해서 학습 데이터를 늘려 나가면서 그 인식률이 어떻게 변화하는지 살펴본다 그리고 나서 일반적인 경우로 임의의 데이터를 학습시킨 후에 인식률을 알아본다

먼저 한 사용자가 쓴 특정 문장을 학습시킨 후 다른 사용자의 인식률을 알아 보는 실험에서는 <표 41>과 같은 결과를 얻었다 즉, 172자로 이루어진 01번 문장을 04번 사용자가 쓴 데이터로 학습시킨 후에 5명의 다른 사용자가 쓴 같은 문장을 인식하여 보았다 결과적으로 평균 인식률이 90.01%를 기록하여 사용자에게 따른 변형을 어느정도 흡수함을 볼 수 있다. 여기에서 오인식을 일으킨 많은 경우가 대소문자나 0과 O를 혼동하여 발생하였으며, 그 이외에도 다음과 같은 혼동이 일어났다

(/ -) J (L + (C ; 3

두번째 실험으로 특정 사용자에게 대하여 제한된 데이터만으로 학습시킨 후에 그 인식률을 알아본 것이 <표 42>에 나와 있다 이 표는 04번 사용자에게 대하여 01번 문장에 포함된 172자로 학습시킨 후에 5개의 다른 문장에 대한 인식률을 알아 본 것이다 이 경우에 평균 95.78% 라는 높은 인식률을 나타냈는데, 이로 미루어 반경제한 퍼셉트론이 제한된 데이터만으로 학습한 경우에도 일반화를 잘함을 알 수 있다

마지막으로 일반적인 경우로 두번째 실험에서 사용한 반경제한 퍼셉트론을 이용하여 학습에 사용하지 않은 문장을 다른 사용자가 쓴 경우의 인식률이 <표 43>에 있다 이때의 인식률은

평균 93.67%로 학습한 데이터를 쓴 사용자의 결과에 비해서는 다소 낮은 인식률을 냈지만 전반적으로 높은 인식률을 얻을 수 있음을 알 수 있다

5 결론

본 논문에서는 신경망의 장점인 병렬분산 처리능력을 유지 하면서 실시간 학습기능을 갖는 반경제한 퍼셉트론의 구조와 학습방법에 대하여 알아본 후 온라인 필기체 문자를 인식하는데 적용하여 보았다 여러가지 상황에 따른 실험결과에 의하면 문자 단위로 분리된 영문자 및 숫자에 대해서 높은 인식률을 얻을 수 있었다. 하지만, 이제까지의 실험은 다소 제한된 범위에서 수행되었기 때문에 보다 다양한 경우에 대해 반경제한 퍼셉트론의 패러미터를 변경시키면서 그 능력을 평가해 보아야 할 것이다

앞으로 좀더 연구하여야 할 내용은 보다 안정된 특장을 추출하는 방법에 대한 것이다 물론 약간의 입력변형은 실시간 학습능력을 갖고있는 신경망인 반경제한 퍼셉트론이 어느정도 흡수할 수 있겠지만, 필기 데이터에서 발생하는 많은 변형들을 모두 처리하기는 불가능할 것이다 이를 위하여 본 연구팀은 코호년의 자기구성 특성지도 방법을 사용하여 주어진 입력을 고정된 수의 코드북중의 하나로 바꾸어서 입력을 반도록 시스템을 보완하고 있다

참고 문헌

- [1] 박영환, 방승양, "패턴 분류용 신경회로망 APC의 설계 및 구현," 제 2회 신경회로망 연구회 연차 총회 및 학술대회, pp 30-35, 1991년 6월
- [2] 이일병, "동적 문자인식 모형 사례연구," 정보과학회지, 제 9권 1호, pp 64-78, 1991년 2월
- [3] 최정훈, 권희용, 김춘석, 황희용, "신경망 모델을 이용한 한글 필기체 온라인 인식," 정보과학회 논문지, 제 17권 5호, pp 540-549, 1990년 9월
- [4] T Fujisaki, T. E. Chafalas, J Kim, C C Tappert, and C G Wolf, "Online run-on character recognizer Design and Performance," Pattern Recognition and Artificial Intelligence, 1991
- [5] K Fukushima, "Neocognitron A hierarchical neural network capable of visual pattern recognition," Neural Networks, vol 1, pp 119-130, 1988
- [6] I Guyon, P Albrecht, Y Le Cun, J Denker, and W Hubbard, "Design of a neural network character recognizer for a touch terminal," Pattern Recognition, vol 24, no 2, 1991
- [7] M. W Kim, J-H Kim, C-D Lim, and G L Lee, "Context dependent Pattern Recognition Using a Hierarchical Interaction Architecture," 제 2회 신경회로망 연구회 연차 총회 및 학술대회, pp 18-23, 1991년 6월
- [8] M. Minsky and S Papert, Perceptrons, 2nd Edition, The MIT Press, 1988
- [9] P. Morasso, "Neural models of cursive script handwriting," in Proc International Conference on Neural Networks, vol II, pp 539-542, San Diego, 1989
- [10] D L Reilly and L N Cooper, "An overview of neural networks Early models to real world systems," An Introduction to Neural and Electronic Networks, 1990
- [11] D E Rumelhart, G E Hinton, and R J Williams, "Learning internal representations by error propagation," in Parallel Distributed Processing Explorations in the Microstructure of Cognition, vol I, pp 318-362, Bradford books, 1986
- [12] C L Scofield, "Learning internal representations in the coulomb energy network," in Proc International Conference on Neural Networks, vol 1, pp 271-276, 1988

<표 41> 사용자 무관성에 대한 실험

학습화일	인식화일	인식률	인식률
01-04	01-04	160 / 172	93.02
	01-05	160 / 173	92.49
	01-06	141 / 172	81.98
	01-07	153 / 172	88.95
	01-08	161 / 172	93.60

<표 42> 일반화 정도에 대한 실험

학습화일	인식화일	인식률	인식률
01-04	02-04	189 / 194	97.42
	03-04	184 / 200	92.00
	05-04	190 / 198	95.96
	06-04	195 / 201	97.01
	07-04	194 / 201	96.52

<표 43> 일반적인 경우의 인식률 실험

학습화일	인식화일	인식률	인식률
01-04	02-05	190 / 195	97.44
	03-06	176 / 201	87.56
	04-07	179 / 191	93.72
	05-08	195 / 198	98.48
	06-09	175 / 192	91.15