

論文96-33B-7-19

필기체 숫자인식을 위한 병렬 자구성 계층 신경회로망

(Parallel, Self-organizing, Hierarchical Neural Networks for Handwritten Digit Recognition)

房 極 準 *, 趙 南 信 *, 康 昌 彦 *, 洪 大 植 *

(Keuk Joon Bang, Nam Shin Cho, Chang Eon Kang, and Dae Sik Hong)

요 약

이 논문에서는 필기체 숫자에서의 다양한 변형을 흡수할 수 있고, 학습 시간을 줄이기 위하여 병렬 자구성 계층 신경회로망을 이용한 필기체 숫자인식 방법을 제안한다. 일반적인 다층 신경회로망의 역전파 학습 알고리즘이 과다한 학습 시간을 필요로 하는데 반하여 제안된 방법은 은닉층이 없는 2층 신경회로망의 계층적 연결에 의하여 학습시간을 줄일 수 있으며 특히 필기문자에서와 같이 다양한 변형을 흡수해야 하는 분류문제에서는 여러가지 선형 및 비선형 변환과정을 각 계층에 부여하여 패턴의 변형을 흡수하였다. 또한 각각의 출력노드의 특성 차이로 인하여 출력노드 출력값의 의미가 서로 다르기 때문에 발생하는 차이점을 흡수하기 위한 출력노드 신뢰성 구간을 모든 출력노드에 개별적으로 적용하였다. 제안된 방법으로 NIST 데이터베이스를 이용하여 실험한 결과 정인식 97.48%, 오인식 1.72%, 거부 0.78%의 인식율을 얻었으며, 또한 병렬 자구성 계층 신경회로망의 여러 특징을 항목별로 비교 실험하여 제안된 방법의 유효함을 보였다.

Abstract

In this paper, we propose the Parallel, Self-Organizing, Hierarchical Neural Networks as a handwritten digit recognition system. This system can absorb the various shape variations of handwritten digits by using the different methods of extracting the features in each stage neural network(SNN) of the PSHNN, and can reduce training time by using the single layer neural network as the SNN, and can obtain high rate of correct recognition by using the certainty area in all the output nodes individually. Experiments have been performed with NIST database. In which we use 21,315 digits (10,652 digits for training and 10,663 digits for testing). The results show that the correct rate is 97.48%, the error rate is 1.72% and the reject rate is 0.78%.

I. 서 론

문자 인식에 있어 인식성능을 향상시키기 위하여 일반적으로 두가지 측면을 주로 고려하게 된다. 그 첫 번째가 실제 인식 단계에서의 입력으로 사용하게 될 인식문자 패턴에 대한 전처리 및 특징추출시 인식에 필요한 정보의 손실이 최소가 되도록 하는 방법에 관한

연구이고, 두번째가 특정한 패턴 특징에 국한되지 않은 일반화 및 객관적 적용 능력을 가지면서 유사 패턴에 대한 인식율을 극대화 할 수 있는 분류기의 적용에 관한 연구이다^[1, 2, 3, 4, 5, 6].

전처리 단계에서는 일반적으로 보다 우수한 인식성능을 얻기 위하여 패턴의 형태를 다양하게 변형하는 방법이 많이 사용되며, 이 변형 방법은 크게 선형 형태변형(Linear Shape Variation)과 비선형 형태변형(Non-linear Shape Variation)으로 분류할 수 있다. 선형 형태변형의 방법으로는 이동(Translation), 신축(Scale), 회전(Rotation), 밀림(Shear) 등의 정규화 과

* 正會員, 延世大學校 電子工學科

(Dept. of Electronic Engineering Yonsei University, Seoul, Korea)

接受日字: 1996年1月5日, 수정완료일: 1996年6月21日

정이 포함되며, 비선형 형태 정규화 방법으로는 특징이 많이 모여 있을 수 있는 부분을 강조하기 위한 점밀도, 선밀도, 획밀도 정규화 등 다양한 비선형 형태 변형 과정이 포함된다. 특히 매우 불규칙적이면서 부분적인 형태왜곡이 심한 필기체 문자에서의 전처리과정은 선형 형태 변형과 더불어 비선형 형태 변형의 적용이 필수적이다^[1, 2, 3].

또한 특징추출단계에서는 전처리 과정에 의하여 정규화되어진 패턴정보가 인식부에서 보다 효율적으로 사용되어질 수 있는 인식부의 입력정보를 추출하는 단계로서, 일반적으로 가장 많이 사용하는 매쉬(Mesh)특징 이외에 2차원 허스토그램 분포특징, 자형 유크선 특징, 푸리에변환 등에 의한 주파수특성 등 많은 방법이 이용되고 있다^[2, 3].

한편 신경회로망을 이용한 필기체 문자인식에 대한 연구는 신경회로망이 갖는 적응적 특성으로 인하여 활발히 연구되고 있는 분야이다. 그러나 대부분의 신경회로망 분류기의 경우 패턴의 전처리 및 특징추출 방법의 적용에 있어 한두가지 정도의 극히 제한된 방법을 이용함으로 인하여 다양한 형태 왜곡을 충분히 흡수하지 못하는 문제점이 있으며, 또한 다층 신경회로망을 문자인식에 적용할 경우 역전파 학습 알고리즘이 과다한 학습 시간이 필요하다는 문제점도 지적되고 있다^[1, 2, 3, 4, 5, 7].

따라서, 본 논문에서는 필기체 문자가 내포하고 있는 다양한 형태 변형 및 다층 신경회로망이 갖는 학습 시간의 문제점을 등을 해결하고자 병렬 자구성 신경회로망을 이용한 필기체 숫자인식 시스템을 제안한다.

병렬 자구성 계층 신경회로망은 은닉층이 없는 2층 신경회로망의 계층적 연결 및 각 계층에서 학습된 데이터를 다음 계층에서 제외하는 학습 방법에 의하여 학습 시간을 줄일 수 있으며 특히 필기문자에서와 같은 문제에서는 여러가지 선형 및 비선형 변환 등의 전처리과정 및 다양한 특징추출방법 등을 각 계층에 부여하여 다양한 서체 변형을 흡수할 수 있다. 또한 각각의 출력노드의 특성 차이로 인하여 출력노드 출력값의 의미가 서로 다르기 때문에 발생하는 차이점을 흡수하기 위한 출력노드 신뢰성 구간을 모든 출력노드에 개별적으로 적용하였다^[1, 8, 9, 10].

본 논문의 구성은 모두 4장으로 구성되어 있으며 제2장에서는 병렬 자구성 계층 신경회로망을, 제3장에서는 제안된 방법에 의한 실험 및 고찰을 그리고 제4

장에서는 결론을 기술하였다.

II. 병렬 자구성 계층 신경회로망(PSHNN : Parallel, Self-organizing, Hierarchical Neural Networks)

1. PSHNN 개요

다층 신경회로망은 일반화된 델타규칙(Generalized Delta Rule) 학습방법을 사용함으로써 신뢰성 면에서는 많은 빌전을 했지만, 학습과정에서 국부최소점에 빠질 가능성이 있기 때문에 성능을 높여야 할 경우 매우 많은 학습 시간이 소요되고, 기억 패턴의 수정이나 추가학습이 불가능하다는 문제점은 여전히 존재한다^[1, 2, 3, 8, 9, 10]. 따라서 PSHNN의 목적은 신경회로망 시스템의 복잡도를 줄여 학습시간을 줄이고, 학습 데이터의 복잡도에 따라 신경회로망의 계층 수를 스스로 조정하여 다양한 변형을 계층적으로 흡수할 수 있도록 하는 데 있다^[1, 8, 9, 10].

PSHNN의 기본적인 동작원리는 다음과 같다. 즉, 일반적인 분류 문제에서 경계영역이 아닌 분류영역은 단순한 분류기로도 분류가 잘 되지만 경계영역 주변의 데이터는 분류에 어려움이 존재한다. 이러한 문제에 대하여 PSHNN은 쉽게 분류할 수 있는 데이터에 대하여는 먼저 분류를 하고 남은 데이터에 대하여 데이터 변환을 통한 특정 분산을 시도한 다음 새로운 분류를 시도하는 방법으로 모든 데이터가 분류될 때까지 반복하며 계층을 스스로 생성해 가는 방법이다^[8, 9, 10].

PSHNN의 시스템 구조는 그림 1과 같이 독립된 N개의 SNN(Stage Neural Network)을 병렬로 자구성(Self Organizing)하도록 된 구조를 갖는다. 각각의 SNN은 은닉층 없이 입력층과 출력층으로만 구성되어 있는 2층 신경회로망으로 은닉층을 포함하는 일반적인 다층 신경회로망에 비하여 매우 간단하다. 또 학습방법으로써는 델타규칙 학습 알고리즘을 통하여 오류 역전과 학습방법을 사용한다. 각각의 개별 SNN마다는 서로 상이한 변환방법에 의한 전처리 및 특징추출과정을 수행함으로써 바로 윗단에서 제대로 인식하지 못한 패턴에 대해서 다른 방법으로 변형된 입력 데이터를 사용하여 다시 인식을 시도함으로 인해 비선형적인 특징이 강한 필기체 숫자의 인식을 새로운 측면에서 시도할 수 있다^[1, 8, 9, 10].

한편 인식 처리부는 각각의 SNN의 출력단에서의 출

력을 거부 또는 인식으로 받아들일 것인지를 검사하는 인식 알고리즘이 적용되어 있으며, 거부될 경우에만 다음단의 SNN이 수행된다. 그림 1에서 점선 연결의 경우가 거부의 경우이며 이는 다음단의 수행에 대한 제어 연결선의 의미도 갖는다^[8, 9, 10].

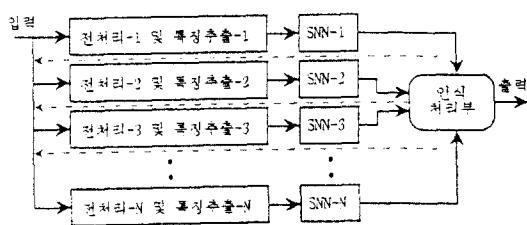


그림 1. PSHNN 시스템 구조도
Fig. 1. Architecture of PSHNN.

2. PSHNN 학습 알고리즘

PSHNN의 학습과정은 먼저 첫번째 SNN을 어느정도 학습 시킨 후 제대로 학습된 것과 잘못 학습된 것들을 분류하여 신뢰성 구간을 결정하고 개별 학습 데이터의 결과값이 받아들일 수 있는 구간에 속하면 통과시키고 그렇지 않으면 거부한다. 거부된 학습 데이터들은 다음단의 전처리 및 특징추출과정을 거쳐 다음단의 SNN을 학습시킨다. 이러한 반복과정은 학습 데이터가 모두 학습되어 학습이 완료되거나, 거부 데이터가 모두 특정한 한 부류에 속하는 경우이거나, 또는 시스템에서 정한 SNN의 최대갯수 M에 도달할 경우의 3 가지 중 한 조건이 만족되면 학습을 종료한다^[9, 10]. 이러한 학습 과정을 정리하면 그림 2와 같다.

3. 신뢰성 구간 설정

신뢰성 구간의 설정은 개별 SNN마다 학습환경과 출력노드의 특성으로 인하여 각각의 출력노드의 출력값의 의미가 서로 다르기 때문에 발생하는 차이점을 흡수하기 위하여 사용한다^[9, 10].

개별 SNN의 학습이 끝나면 각각의 학습 데이터에 대한 출력값을 종합하여 각각의 노드에 대한 정(正)의 신뢰구간인 상한한계(Upper Boundary) 및 오(誤)의 신뢰구간인 하한한계(Lower Boundary)를 구하고 상한한계와 하한한계의 사이 영역이 거부영역 즉, 해당 출력값의 정/오 판단이 불분명한 영역을 설정하여 인식 실행시 판단을 거부하는 영역으로 사용한다.

조건 1 : 개별 SNN의 학습 종료조건 설정. (최대

epoch 횟수, MSE의 임계값 등)

조건 2 : 최대 SNN 자구성 갯수 = M

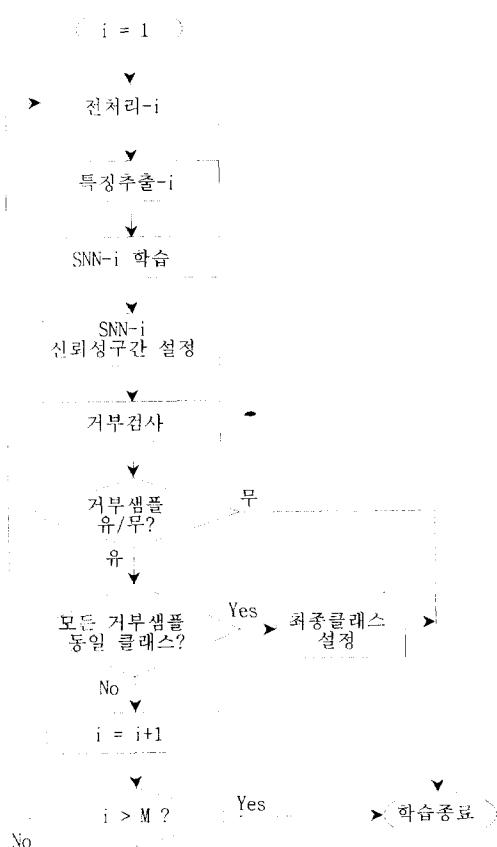


그림 2. PSHNN 학습 알고리즘

Fig. 2. Training Algorithm of PSHNN.

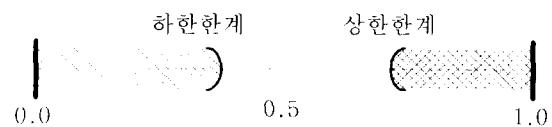


그림 3. 출력노드의 신뢰성 구간

Fig. 3. Certainty Area in a Output Node.

그림 3은 하나의 출력 노드에 대한 신뢰성 구간의 예이다^[9, 10].

일반적으로 신뢰성 구간의 설정 방법은 학습이 끝난 데이터 중에서 상하위 중간값인 0.5를 상하 한계로 초기 설정한 후 인식 검사를 수행하여 조건에 부합되는 데이터만을 취합하고 취합된 데이터 중에서 상하한 영역 각각에 대하여 경계에서부터의 일정비율에 위치한 값을 경계값으로 재설정하여 해당 출력노드의 신뢰성

구간으로 정하게 되며, 이 신뢰구간의 재설정을 위하여 사용하는 조정비율값은 시스템의 결과에 대한 신뢰도에 적절적인 영향을 주게 된다. 즉, 이 값은 시스템의 신뢰도를 높일 필요가 있을 때 값을 높여 사용하게 된다.

4. 인식 처리부

인식 처리부에서는 개별 SNN별로 신경회로망 결과값에 대해서 인식검사를 수행한다. 인식의 판단 방법은 각각의 노드에 대한 출력값이 하한한계 이하이면 0, 상한한계 이상이면 1의 신뢰성을 갖는다고 보고, 하한 및 상한 한계 사이의 값이면 거부한다. 인식의 결과는 10개의 출력 노드에 대하여 오로지 하나만 1의 신뢰구간에 속해야 하고 그외의 9개 노드에서 모두 0의 신뢰구간에 속할 때만 인식으로 판단하고 그 외의 결과에 대하여는 판단을 거부한다. 그러므로 인식 판단을 얻기 위해서는 10개의 출력노드중 오직 하나의 출력노드가 상한한계보다 크며 나머지는 모두 하한한계보다 작아야 한다.

각각의 개별 SNN에서 인식이 거부된 경우 바로 다음 단의 SNN이 수행되어 인식의 판단을 구하는 과정을 반복하여, 학습시에 구성된 모든 개별 SNN을 통하여 인식이 거부된 경우에만 해당 패턴은 인식 거부로 판정하게 된다.

III. PSHNN의 필기체 숫자인식에의 적용

1. 데이터베이스

실험에서 사용한 데이터베이스는 NIST에서 제공하는 필기체 데이터베이스 NIST(National Institute of Standards and Technology) Special Database 19를 사용하였다.

사용된 데이터베이스에는 여러 가지 분류 방법에 의하여 데이터가 제공되고 있으며 본 논문에서는 필기자에 따라 분류된 f0000~f0499의 500명에 의한 데이터 분류를 학습 데이터로 사용하였으며 f0500~f0999의 500명에 의한 데이터 분류를 테스트 데이터로 사용하였다. 제공된 데이터베이스 중 학습데이터로 사용된 분류에는 53,449자, 테스트데이터로 사용된 분류에는 53,312자의 데이터가 제공되고 있으며 이 중에서 임의로 추출된 21,315자(학습데이터 10,652자, 테스트 데이터 10,663자)를 본 논문에서 사용하였다.

2. 전처리

본 논문에서 사용한 전처리 방법은 세선화, 기울기교정, 덧씌우기 등의 선형 형태변환 및 점밀도를 이용한 비선형 변환, 교차횟수를 이용한 비선형 변환, 회간격에 바탕을 둔 획밀도를 이용한 비선형 변환 등 6가지의 기본 변환 방법과 그들의 조합 방법을 사용하였다.

1) 세선화

세선화 과정은 입력패턴의 형태를 쉽게 볼 수 있을 뿐만 아니라 입력 데이터의 양도 줄일 수 있다는 장점이 있다. 따라서 많은 문자인식 시스템에서 세선화 과정은 빼놓을 수 없는 중요한 전처리 과정에 속하고 있다.^[12, 3, 11, 12]

일반적인 세선화의 방법은 패턴의 경계에 위치한 화소들을 양파껍질 벗기듯이 한겹씩 벗겨가는 방법을 사용하며 이러한 방법을 반복적 세선화 알고리즘이라 한다. 이 반복적 세선화 알고리즘은 다시 한화소의 제거 여부를 판단하기 위한 이웃화소 검사 방법에 따라 순차 세선화와 병렬 세선화의 방법으로 분류된다.

또한 화소의 제거 검사를 위하여 사용하는 방법은 그림 4와 같이 8-이웃 방법이 사용되며 이 경우 화소의 연결성 보장을 위하여 8-이웃의 교차수, 연결수를 이용한 알고리즘이 사용된다.

본 논문에서는 참고문헌 [11. Wang]의 알고리즘을 사용하였다.

8	1	2
7	P	3
6	5	4

그림 4. 화소의 8-이웃 관계

Fig. 4. 8-Neighborhood of a pixel.

2) 기울기 교정

필기체 문자에서는 문자의 기울기가 매우 다양하여 기울기를 교정하지 않고 인식한다는 것은 매우 어려운 일이다. 따라서 필기문자 인식분야에서는 기울기의 교정이 필수적인 문제이다.^[12, 3, 13]

일반적으로 가장 많이 사용하는 기울기 교정 방법은 Hough 변환 방법이다. 즉, 하나의 패턴에 대하여 직선성분들의 기울기를 추정하여 최 우도 기울기를 교정할 각도로 정하여 회전시키는 방법이다.^[13] 그러나 이

Hough 변환 방법은 계산량이 많다는 단점이 있다.

본 논문에서는 처리속도를 고려하여 그림 5와 같이 패턴 상하부의 무게중심점을 지나는 선분의 각도를 패턴의 기울기로 정하는 방법을 적용하였으며, 기울기 교정방법은 x축방향으로의 밀립방법을 사용하였다.



그림 5. 상하영역 중심을 이용한 기울기 교정

Fig. 5. Slant normalize using center points of upper and lower area.

3) 덧씌우기

패턴인식에 있어 세선화 과정을 거친 데이터를 이용하여 인식을 실행할 경우 일반적인 패턴인식 방법인 구문론적 방법 또는 통계적 방법에서는 유효한 정보를 추출하기 위하여 크게 도움이 되지만 신경회로망을 이용한 인식 방법에서는 역효과가 발생할 수 있다. 따라서 패턴의 두께에 대한 균일화 목적의 측면에서 세선화 된 데이터의 덧씌우기 과정은 신경회로망을 이용한 패턴인식의 방법에서 도움이 된다^[1].

덧씌우기 방법은 패턴 전체에 대하여 흰화소 중 8-이웃 화소 중에 흑화소가 존재하는 경우 해당 흰화소를 흑화소로 변환하는 방법을 사용하였다.

4) 점밀도를 이용한 비선형 변환

입력 패턴의 각 점에서의 특정 특징을 수평축 또는 수직축에 투영하여 특정 투영 히스토그램을 구하고 이를 균일화 함으로써 입력 패턴의 특징을 균일화하는 방법이 입력 패턴에 대한 비선형 정규화 방법이다^[2, 3].

이 과정에서 특정 투영은 수평축 투영함수 $H(i)$ 와 수직축 투영함수 $V(j)$ 를 계산하는 과정이라 할 수 있고 특정 밀도 균일화는 입력패턴 $f(i, j)$ 의 좌표 (i, j) 에 대응하는 정규화 패턴 $G(m, n)$ 의 좌표 (m, n) 을 구하는 과정이라 할 수 있다.

점밀도를 이용한 비선형 변환 방법의 특징 투영 함수는 입력 영상 중 흑화소의 분포를 수평 또는 수직 투영한 후 균일화 하여 얻어지며 식으로 표현하면 다음과 같다.

$$H(i) = \sum_{j=1}^J f(i, j) + \alpha_H \quad (1)$$

$$V(j) = \sum_{i=1}^I f(i, j) + \alpha_V \quad (2)$$

여기서 α_H 와 α_V 는 상수이며 정규화 패턴의 좌표 (m, n) 은 다음 식으로 구한다.

$$m = \sum_{k=1}^i H(k) \times \frac{M}{\sum_{k=1}^I H(k)} \quad (3)$$

$$n = \sum_{l=1}^j V(l) \times \frac{N}{\sum_{l=1}^J V(l)} \quad (4)$$

5) 교차횟수를 이용한 비선형 변환

교차횟수를 이용한 비선형 변환 방법의 특징 투영 함수는 입력 영상 중 백화소에서 흑화소로 변하는 교차횟수를 수평 또는 수직 투영한 후 균일화 하여 얻어지며 식으로 표현하면 다음과 같다^[2, 3].

$$H(i) = \sum_{j=1}^J \overline{f(i, j-1)} \times f(i, j) + \alpha_H \quad (5)$$

$$V(j) = \sum_{i=1}^I \overline{f(i-1, j)} \times f(i, j) + \alpha_V \quad (6)$$

여기서 $\overline{f(i, j)}$ 는 $f(i, j)$ 의 논리역이며, 정규화 패턴의 좌표 (m, n) 은 다음 식으로 구한다.

$$m = \left\lfloor \sum_{k=1}^i H(k) \times \frac{M}{\sum_{k=1}^I H(k)} \right\rfloor \quad (7)$$

$$n = \left\lfloor \sum_{l=1}^j V(l) \times \frac{N}{\sum_{l=1}^J V(l)} \right\rfloor \quad (8)$$

6) 획간격에 바탕을 둔 획밀도를 이용한 비선형 변환

이 방법은 단순히 획의 수만을 이용하는 교차횟수에 바탕을 둔 획밀도를 이용한 방법을 개선하기 위하여 획간격(Line Interval)을 사용하여 획밀도를 정의한다. 즉, 획밀도란 패턴의 한 점 (i, j) 가 백화소의 점이면 특징함수 $F_H(i, j)$ 와 $F_V(i, j)$ 는 각각 이웃한 두 획 사이의 간격에 대한 역수로 정의되며 다음과 같이 표현된다^[2, 3].

$$F_H(i, j) = \frac{1}{h(i, j)} \quad (9)$$

$$F_V(i, j) = \frac{1}{v(i, j)} \quad (10)$$

여기서 $h(i, j)$ 는 점 (i, j) 의 좌우에 이웃하는 두 획 사

이의 수평 간격을 나타내며, $v(i, j)$ 는 점 (i, j) 의 상하에 이웃하는 두 획 사이의 수직 간격을 나타낸다. 이때 점 (i, j) 가 흑화소이면 $F_H(i, j)$ 와 $F_V(i, j)$ 는 매우 작은 값으로 설정된다. 이와같은 특징함수 $F_H(i, j)$ 와 $F_V(i, j)$ 를 수평, 수직 축에 투영한 투영함수는 다음과 같다.

$$H(i) = \sum_{j=1}^J F_H(i, j) \quad (11)$$

$$V(j) = \sum_{i=1}^I F_V(i, j) \quad (12)$$

이 투영함수에서 정규화 패턴의 좌표는 다음 식과 같이 투영함수를 누적하는 것으로 정의되는 뒤틀림함수 (warping function) $\phi(i)$ 와 $\psi(j)$ 를 사용하여 정규화 패턴의 좌표 (m, n) 을 구한다.

$$m = \phi(i) = \frac{\sum_{k=1}^i H(k) \times \frac{M}{\sum_{k=1}^I H(k)}}{} \quad (13)$$

$$n = \psi(j) = \frac{\sum_{l=1}^j V(l) \times \frac{N}{\sum_{l=1}^J V(l)}}{} \quad (14)$$

3. 특징추출

전처리 과정이 끝난 입력패턴은 이진 영상의 형태를 갖고 있으며 이로부터 신경회로망의 입력으로 사용할 수 있는 특징벡터를 추출하게 된다.

일반적으로 가장 많이 사용하는 특징으로는 매쉬 (Mesh)특징, 2차원 히스토그램 분포특징, 자형 윤곽선 특징, 푸리에변환 등에 의한 주파수특성 등 많은 방법이 이용되고 있다.



그림 6. 논문에서 사용한 특징추출

- (a) 매쉬특징 (b) 2차원 히스토그램 (c) 자형 윤곽선

Fig. 6. Features that used in this paper.

- (a) Mesh (b) Histogram (c) Outline of pattern

매쉬특징은 그림 6(a)와 같이 주어진 패턴을 정해진 크기의 모눈형태에 매핑한 후 각각의 모눈 영역에 발생하는 흑화소의 양을 입력으로 사용하는 방법으로 이

러한 특징을 사용하기 위하여 전처리 후의 데이터에 대한 크기 정규화 과정이 포함된다. 2차원 히스토그램은 그림 6(b)와 같이 가로 및 세로축에 대하여 패턴의 분포곡선을 이용하여 흑화소의 밀집도에 따른 정보를 특징으로 사용하는 방법이다. 또한 자형 윤곽선 특징은 그림 6(c)와 같이 패턴의 외부 윤곽선 정보를 특징으로 이용하는 방법이다.

본 논문에서는 모든 SNN에서 이와같은 세가지의 특징을 모두 사용한 특징추출 방법을 채택하였다.

4. 전처리 및 특징추출의 성능평가

본 논문에서는 2절에서 설명한 6가지의 기본 변형방법을 사용한 전처리 과정을 이용하여 5개의 개별 SNN에 각각 상이한 방법으로 입력 데이터의 전처리 과정에 의한 변형방법을 적용하기 위하여 6가지 기본 변형방법의 특징을 잘 보여줄 수 있으리라 판단되어 선정된 12가지 조합된 변형방법을 첫 계층부터 순서적으로 해당 계층에서의 최상의 성능을 내는 변형을 찾는 방법을 사용하였다.

각 계층에서 최상의 성능을 내는 변형방법을 찾는 과정은 각 계층에서 모든 변형방법을 모두 적용한 후 정인식 및 오인식 그리고 수령속도 등을 종합한 인식 성능의 척도를 사용하여 최상의 인식 결과를 보인 방법을 해당 계층의 변형 방법으로 적용하고 다음 계층에서는 첫 번째 계층에서부터 바로 전 계층까지 중 선정되지 않은 변형 방법만을 사용하여 위의 성능 평가 방법을 적용해 가는 방법을 사용하였으며 이 때 선정을 위한 데이터는 학습에 이용된 데이터의 반을 학습용으로 나머지 반을 테스트용으로 사용하여 결정하였다.

표 1이 각각의 전처리방법에 대한 개별 계층에서의 적용 결과를 평가한 순위표이며 표 내의 숫자는 해당 계층에 각각의 변형 방법을 적용하였을 때의 순위를 나타내고, 표의 비고란에 표시된 ○표가 5개의 SNN에서의 전처리방법으로 사용된 표시이다.

또한 본 논문에서는 각각의 SNN에 적용하기 위한 전처리의 개별 방법 뿐만 아니라 특징추출의 서로다른 방법을 찾으려 하였으나 실험결과 본논문에서 사용한 매쉬특징, 2차원 히스토그램 분포특징, 자형 윤곽선 특징 등 세가지 특징을 모든 SNN에 동시에 적용한 경우가 항상 높은 인식성능을 보여 특징추출에서의 변형은 적용하지 않았다.

표 1. 데이터 전처리방법과 각 계층에서의 성능평가

Table 1. Performance of the preprocessings in each stage.

[단위 : 순위]

전처리방법		계층	1	2	3	4	5	선정
1	기		4	5	2	1	-	○
2	세		11	10	3	2	8	×
3	세+기		10	7	7	5	2	×
4	기+세		9	11	3	2	3	×
5	세 + 기 + 덧		2	2	9	9	6	×
6	기 + 세 + 덧		1	-	-	-	-	○
7	세 + 기 + 덧 + 점		8	3	1	-	-	○
8	세 + 기 + 덧 + 교		6	4	8	7	5	×
9	세 + 기 + 덧 + 회		3	1	-	-	-	○
10	기 + 세 + 덧 + 점		5	6	5	4	1	○
11	기 + 세 + 덧 + 교		7	9	10	8	4	×
12	기 + 세 + 덧 + 회		12	8	6	6	6	×

기 : 기울기 교정

세 : 세선화

덧 : 덧씌우기

점 : 점밀도 비선형 변환

교 : 교차회수 비선형 변환

회 : 회밀도 비선형 변환

5. 실험환경 및 조건

실험 환경은 팬티움 PC 90MHz 시스템 및 Windows-95 운영체제 그리고 Boland C++ 4.5 컴파일러 환경에서 구현하였다.

개별 SNN의 학습 종료조건은 표 2와 같이 해당 SNN 계층에 대한 최대 반복학습횟수, MSE 임계값 또는 MSE불변임계 조건 중 하나가 만족되면 SNN의 학습을 종료하며, 학습시 PSHNN의 최대 SNN자구성 개수는 5로 정하였다.

표 2. SNN 학습 종료조건

Table 2. Stop Criteria of training for SNN.

SNN 계층	1	2	3	4	5
최대 반복횟수	100				
MSE 임계값	0.0035	0.007	0.014	0.02	0.03
MSE 불변 임계값	0.0001이내 변화 6회 반복시 종료				

개별 SNN의 입력층 노드의 수는 매쉬특징(120개), 2차원 히스토그램 분포특징(44개), 자형 유파선 특징(88개) 등 세가지 특징을 모두 사용한 252개의 노드를

사용하였으며, 개별 SNN의 출력층 노드의 수는 숫자 0~9의 10개를 사용하였다.

또한 시스템의 결과에 대한 신뢰도에 직접적인 영향을 주는 출력노드의 신뢰구간의 설정시 사용하는 경계에서의 조정비율값은 1%를 사용하였다.

6. 실험결과 및 고찰

제안된 방법의 성능 평가를 위하여 실험결과를 정인식율, 오인식율, 거부율 및 신뢰도(식 15) 등의 평가점수를 사용하여 이미 발표된 필기숫자 인식 시스템과 비교하여 보았다^[3]. 비교의 대상으로는 신경회로망을 이용한 필기숫자 인식시스템에 국한하였으며, 제안 시스템은 PSHNN의 기본 신경회로망인 SNN에 대하여 은닉층이 없는 2층 신경회로망(입력노드 252개, 출력노드 10개)의 방법을 사용하였다. 실험 결과 표 3에서 보는 바와 같이 PSHNN방법은 거의 모든 측면에서 비교 시스템보다 좋은 성능을 보였다.

$$\text{신뢰도} = \frac{\text{정인식율}}{\text{정인식율} + \text{오인식율}} \times 100 \quad (15)$$

표 3. 필기체 숫자인식 성능비교

Table 3. Comparison of recognition results.

방법	정인식	오인식	거부	신뢰도	데이터수		해상도 (dpi)
					학습	테스트	
Le Cun 90	92.00	2.00	6.00	97.80	7,291	2,007	300
Krzyak 90	94.85	5.15	0.00	94.85	4,000	2,000	166
Knerr 92	90.30	1.00	8.70	98.90	7,200	1,800	
제안방법	97.48	1.72	0.78	98.26	10,652	10,663	300

또한 PSHNN은 여러 가지 일반적인 신경회로망과 비교되는 특징이 있다. 따라서 이러한 특징을 세가지 측면에서 제안하는 기본 시스템과 비교 분석하였다. 즉, 첫 번째 은닉층이 있는 3층 퍼셉트론을 단독으로 사용한 경우, 두 번째 PSHNN의 SNN에 은닉층이 있는 3층 퍼셉트론 사용한 경우, 그리고 마지막으로 학습시 계층간 학습데이터 탈락을 적용하지 않은 경우 등의 각각에 대하여 비교 실험을 수행하였다. 비교 방법은 정/오인식율 및 거부율, 인식 신뢰도, 학습시간, PSHNN 계층 수, 신경회로망 가중치에 소요되는 메모리 용량 등이다. 표 4는 이러한 비교 항목에 대한 비교 실험 결과이다.

실험결과 은닉층이 있는 3층 퍼셉트론을 단독으로

사용한 경우 학습시간 및 사용 메모리의 측면에서 제안시스템보다 나은 점이 없음에도 불구하고 인식 결과가 제안시스템에 비하여 좋지 않았으며, SNN에 은닉층을 포함한 방법은 제안 방법보다 정인식율은 우수하였으나 신뢰도가 떨어짐에 따라 사용 메모리의 양 및 학습시간의 희생에 비하여 좋은 결과를 내지 못하였다. 그리고 학습시 계층간 학습데이터 탈락을 적용하지 않은 경우의 방법에 대한 실험 결과는 제안 방법에 비하여 신뢰도는 약간 좋으나 정인식율 및 학습시간 등에서 제안방법보다 좋지 못한 결과를 보였다. 이상과 같은 실험결과, 비교의 대상이 동일 데이터를 사용한 경우가 아니므로 수치의 직접적 의미는 없겠으나 제안시스템의 결과가 일반적으로 비교하는 척도의 측면에서 타당성이 있다고 하겠다.

표 4. 비교실험 결과

Table 4. Results of experiments for comparison.

구분	정인식 (%)	오인식 (%)	거부 (%)	신뢰도 (%)	학습시간 (초)	계층수	메모리 (KB)
제안방법	97.48	1.72	0.78	98.26	707	5	50.40
비교방법 I	94.89	2.16	2.94	97.77	1342	1	41.92
비교방법 II	97.74	2.06	0.18	97.93	1889	5	209.60
비교방법 III	97.16	1.67	1.15	98.30	1315	5	50.40

- 제안방법 : SNN에 은닉층이 없는 2층 퍼셉트론 사용 (입력층 252개, 출력층 10개)
- 비교방법 I : 은닉층이 있는 3층 퍼셉트론의 단독 사용 (은닉층 40개)
- 비교방법 II : SNN에 은닉층이 있는 3층 퍼셉트론 사용 (은닉층 40개)
- 비교방법 III : 학습시 계층간 학습데이터 탈락을 적용하지 않은 경우

한편, 제안 시스템은 II-3절에서 설명한 바와 같이 출력노드의 신뢰구간의 조정비율을 변경함에 따라 시스템의 결과에 대한 신뢰도를 조절할 수 있다. 즉, 조정비율을 높여줌에 따라 정인식율 및 오인식율이 저하되지만 이에따라 신뢰도는 향상되도록 되어 있다. 이러한 관계를 보기 위하여 신뢰구간 조정비율에 따른 신뢰도의 변화 추이를 실험해 보았으며 그 결과가 표 5에서 보여주고 있다. 실험결과 1%의 조정비율 적용시 98.26%의 신뢰도를 얻었으며 이때의 오인식은 1.72%이다. 또한 16%의 조정비율을 적용하였을 때 99.00%의 신뢰도를 얻었으며 이때의 오인식은 0.97%이다.

그리고 PSHNN이 기본적으로 계층구조에 의한 인식시스템이므로 각 계층의 역할 정도를 판단할 필요가 있으며 그 결과는 표 6과 같다. 실험결과 계층이 증가

함에 따라 정인식 및 오인식율은 증가하나 신뢰도는 점차 저하됨을 보여준다. 이는 잘 분류되는 데이터가 앞 계층에서부터 차례로 분리되어 뒷 계층에서는 분류가 어려운 데이터가 집중 처리되기 때문으로 당연한 결과라 할 수 있다. 또한, 이 결과는 초기계층에의 데이터 변형방법 설정이 전체 시스템의 성능향상에 결정적인 영향을 줄 수 있음을 확인할 수 있었다.

표 5. 신뢰구간 변경 실험 결과

Table 5. Results for changing the certainty area.

신뢰구간 설정비율	정인식 (%)	오인식 (%)	거부 (%)	신뢰도 (%)	학습시간 (초)	계층수	메모리 (KB)
제안방법(1%)	97.48	1.72	0.78	98.26	707	5	50.4
2%	97.20	1.66	1.12	98.31	659	5	50.4
4%	97.09	1.52	1.37	98.44	676	5	50.4
8%	96.90	1.11	1.97	98.86	673	5	50.4
16%	96.07	0.97	2.94	99.00	697	5	50.4

표 6. 계층별 인식문자 수

Table 6. Recognition results in each stage.

구분	계층					비율 (%)
	1	2	3	4	5	
정인식	단위	9,701	453	132	82	27
	누계	9,701	10,154	10,286	10,368	10,395
	누계비율	90.97	95.22	96.46	97.23	97.48
오인식	단위	83	39	26	23	13
	누계	83	122	148	171	184
	누계비율	0.77	1.14	1.38	1.60	1.72
신뢰도	99.16	98.81	98.58	98.38	98.26	98.26
거부					84	0.78

또한 본 논문의 실험은 그래픽 사용자 인터페이스 환경에서 구현되어 실험 결과를 시작적으로 볼 수 있도록 구현하였다. 그럼 7은 실제 인식결과의 예에 대한 각 계층별 출력 상황을 보여주고 있다. 그림에서는 각각의 출력노드 그래프가 출력노드의 신뢰성 구간에 따라 다른 색으로 표시되고 있다.

그림 7에서의 위 두가지 경우는 정인식의 예이며, 그 중 '3'의 문자에 대하여는 첫 번째 계층에서 자기 노드에서 최상의 값을 냈지만 신뢰구간에 도달하지 못하여 다음계층으로 진행되어 정인식 된 예이다.

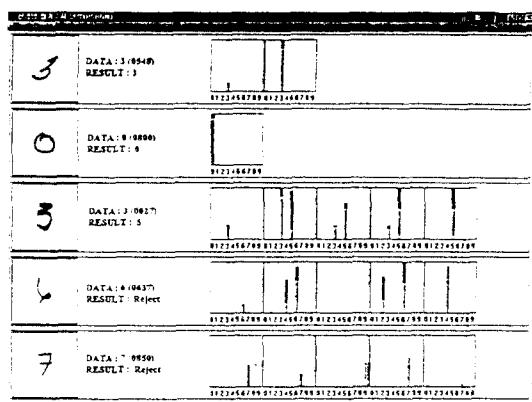


그림 7. 인식결과의 예
Fig. 7. Recognition Samples.

또한 가운데 '3'의 경우는 처음 두단계에서 '3'의 노드에 최상의 값을 내었지만 상한한계를 넘기지 못하였거나 두 번째 상위의 값이 불확실구간의 값을 내어 인식되지 못하여 다음단계에서 '5'의 값으로 오인식 된 것이다. 마지막 두가지 경우는 각각 첫 번째 계층에서 자신의 출력 노드에 최대의 값을 내었지만 상한한계 이상의 값을 내지 못하였거나 다른 노드가 불확실 구간의 값을 내어 인식하지 못하고 다음 계층으로 진행하여 인식이 거부된 것이다.

한편 오인식 중 '7'의 경우는 입력 데이터가 다른 데이터와는 달리 변형된 특징을 가지고 있는 경우에 해당되며 실험결과 이와같은 데이터의 경우가 대표적인 오인식 패턴으로 분류된 것이다.

IV. 결 론

본 논문에서는 필기체 숫자인식의 방법으로 병렬 자구성 계층 신경회로망을 제안하였다. 제안된 방법은 필기체 문자의 다양한 형태변형을 계층 신경회로망의 방법으로 흡수할 수 있었으며, 각각의 계층 신경회로망이 은닉층을 포함하지 않은 단순구조를 가지고 있음으로 하여 학습시간의 단축효과도 얻을 수 있었다. 또한 실험 결과는 기존의 다른 방법에 의한 인식결과와 비교하여 고른 우수성이 있음도 보였으며 병렬 자구성 계층 신경회로망이 갖는 계층 구조 및 신뢰성 구간의 적용 등의 특징이 일반적인 신경회로망 방법에 비하여 좋은 인식성능을 얻을 수 있는 특징임도 볼 수 있었다. 그러나 실험결과 오인식의 입력패턴을 볼 때 인간의

눈으로 판단할 때 정인식이 될 수 있는 많은 데이터가 있음은 불만스러운 현상이었으며 실험 결과에서도 보았듯이 하부계층에서의 오인식을 증가 등은 본 연구에 이어 계속적인 연구가 필요한 부분이라 판단된다.

참 고 문 헌

- [1] 방극준, 조남신, 홍대식, 강창언, “병렬 자구성 계층 신경회로망을 이용한 필기체 숫자인식”, 한국통신학회 추계종합학술발표회 Vol. 14, No 2, pp. 951-954, 1995. 11
- [2] 이성환, “패턴인식의 원리” II권 1부 신경망을 이용한 패턴인식, 홍릉과학출판사, 1994
- [3] 이성환, “문자인식 - 이론과 실제” II권 13장 무제약 필기체 숫자의 오프라인 인식, 홍릉과학출판사, 1994
- [4] Keiji Yamada, “Inverse Recall Neural Network Model and Feedback Pattern Recognition”, *IEEE International Conference on Neural Networks*, pp. 399-406 1993.
- [5] Gwang Hoon Park, “Handwritten Digit and Script Recognition using Density Based Random Vector Function Link Network”, Ph. D. Dissertation, case Western Reserve University, Cleveland, Ohio, U.S.A., January, 1995.
- [6] Lianwen JIN, Kwokping CHAN and Bingzheng XU, “Off-line Chinese Handwriting Recognition using Multi-Stage Neural Network Architecture”, *IEEE International Conference on Neural Networks*, pp. 3083-3088 1995.
- [7] James A. Freeman and David M. Skapura, “Neural Networks algorithms, Applications, and Programming Techniques”, Addison-Wesley Publishing Company, 1991.
- [8] D. Hong and O. K. Ersoy, “Parallel, Self-organizing, Hierarchical Neural Networks”, *IEEE Trans. on Neural Networks*, Vol. 1, No. 2, pp. 167-178 June, 1990.
- [9] D. Hong and O. K. Ersoy, “Parallel, Self-organizing, Hierarchical Neural Networks-II”, *IEEE Trans. on Industrial*

- Electronics*, Vol. 40, No. 2, pp. 167-178 April, 1993.
- [10] Daesik Hong, "Parallel, Self-organizing, Hierarchical neural networks", Ph. D. Dissertation, Purdue University, U.S.A., August 1990.
- [11] P. S. P. Wang and Y. Y. Zhang "A fast and flexible thining algorithm", *IEEE Trans. on Computers*, Vol. 38, No. 5, pp. 741-745
- 1989.
- [12] P. C. K. Kwok, "Customising thinning algorithms", *Proc. IEEE International Conf. on Image Processing Applications*, pp. 633-637 1989.
- [13] J. Illingworth and J. Kittler, "A Survey of the Hough Transform", *Computer Vision, Graphics, and Image Processing*, Vol. 44, pp. 87-116 1988.

저자 소개



房 極 準(正會員)

1961년 6월 30일생. 1985년 2월 연세대학교 전자공학과 공학사. 1995년 2월 연세대학교 산업대학원 전자공학과 공학석사. 1995년 3월 ~ 연세대학교 대학원 전자공학과 박사과정. 1984년 12월 ~ 1991년 10월 삼성전자 응용시스템개발실 주임연구원. 주관심분야는 신경회로망, 영상처리, 패턴인식, GIS



趙 南 信(正會員)

1971년 9월 27일. 1995년 2월 연세대학교 전자공학과 공학사. 1995년 3월 ~ 연세대학교 대학원 전자공학과 석사과정. 주관심분야는 신경회로망, 패턴인식, GIS

康 昌 彦(正會員) 第 31 卷 A 編 第 4 號 參照
현재 연세대학교 전자공학과 교수

洪 大 植(正會員) 第 33 卷 A 編 第 4 號 參照
현재 연세대학교 전자공학과 교수