

오프라인 필기체 숫자 인식을 위한 다양한 특징들의 성능 비교 및 인식률 개선 방안

박 창 순[†] · 김 두 영^{††}

요 약

본 논문에서는 오프라인 필기체 숫자의 변형을 흡수할 수 있는 효과적인 특징을 찾기 위해서 여러 가지 특징의 성능을 비교하였다. 실험적인 성능 비교 결과는 윤곽선을 이용한 4방향성 특징 그리고 교차 거리 +교차 +망 +투영 특징이 오프라인 필기체 숫자 인식에서 인식률과 인식 시간 측면에서 효과적인 것으로 나타났다. 그리고 단일 신경회로망에서 인식률의 한계점을 극복하기 위하여 효과적인 특징을 조합한 복합 특징으로 다수결 투표와 신뢰도 지수를 이용한 모듈화된 신경회로망을 제안한다. 제안된 방식의 성능을 검증하기 위해서 캐나다 Concordia 대학교와 한국의 Dong-A 대학교 오프라인 필기체 숫자 데이터베이스에 대하여 실험을 하였다. Concordia 대학교의 데이터 베이스는 97.1%의 정인식률, 1.5%의 기각률, 1.4%의 오인식률 그리고 98.5%의 신뢰도가 나타났으며, Dong-A 대학교의 데이터 베이스는 98%의 정인식률, 1.2%의 기각률, 0.8%의 오인식률 그리고 99.1%의 신뢰도가 나타났다.

Performance Comparison of Various Features for Off-line Handwritten Numerals Recognition and Suggestions for Improving Recognition Rate

Chang-Soon Park[†] · Doo-Young Kim^{††}

ABSTRACT

In this paper, in order to find effective features which can handle variations in off-line handwritten numerals, we performed a comparative study on various sets of features. Results of experimental performance comparison shows that 4-directional features using contours and features which combined cross distance, cross, mesh and projection features are very effective for off-line handwritten numerals recognition in terms of recognition rates and recognition times. And in order to surmount limitation of recognition rate by a single neural network, we proposed a modularized neural network using majority voting and reliability factor with complex feature that mix effective features together. In order to verify the performance of the proposed method, the handwritten numeral databases of Concordia University of Canada and Dong-A University of Korea are used in the experiments. With the database of Concordia University, the recognition rate of 97.1%, the rejection rate of 1.5%, the error rate of 1.4% and the reliability of 98.5% are obtained; and with the database of Dong-A University, the recognition rate of 98%, the rejection rate of 1.2%, the error rate of 0.8%, the reliability of 99.1% are obtained.

† 정 회 원: 장신전문대학 전산정보처리과 조교수

†† 정 회 원: 동아대학교 전자공학과 교수

논문접수: 1996년 5월 11일, 심사완료: 1996년 7월 16일

1. 서 론

현대 사회에서 발생하는 정보의 양은 급속히 증가하고 있는 반면 수작업을 통한 컴퓨터와의 대화는 시간의 낭비를 초래하고 있어서 정보 입력의 자동화에 대한 필요성이 오래 전부터 대두되고 있다. 최근에 활발히 연구되고 있는 문자 인식은 인쇄체 문자 인식과 필기체 문자 인식으로 나눌 수 있으며 정해진 활자체의 인쇄체 문자 인식은 거의 실용화 단계에 이르렀고 필기체의 경우도 온라인 인식에서는 많은 진전을 이루었다. 오프라인 필기체 문자 인식에서는 필자에 따라 필체가 달라지고 동일한 사람도 환경에 따라 변화가 불규칙하기 때문에 인쇄체 문자 인식에 비하여 인식의 수준이 초보적이라 볼 수 있다. 그러나 문서 자동화의 궁극적인 목적이 인간과 컴퓨터와의 자연스러운 통신이 가능하도록 하는 것이라면 필기체 문자 인식은 필수적인 과제이다.

오프라인 필기체 문자 인식의 기본 단계라고 할 수 있는 오프라인 필기체 숫자 인식은 그 응용 분야가 다양하여 우편번호나 각종 전표 자동 분류 등과 같은 분야에 적용 가능하지만 실용화 단계까지는 해결해야 할 많은 문제를 안고 있다[1]. 그 중의 한가지로 필기체 숫자의 다양한 형태 변형을 흡수할 수 있는 최적의 특징을 찾는 문제가 제기된다. 그리고 단일 특징만으로는 높은 인식률을 기대하기 어려우며, 인식률의 개선을 위해서 다양한 특징을 조합한 복합 특징을 다층 구조로된 단일 신경회로망의 입력으로 사용한다[2]. 이와 같이 복합 특징을 사용하여 다층 구조의 단일 신경회로망을 구성하면 어느 정도 인식률은 증가하지만 장시간의 학습과 인식률의 한계가 발생하는 문제가 제기된다.

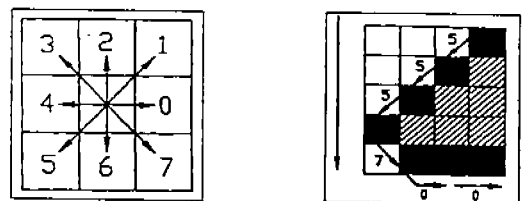
따라서 본 논문에서는 필기체 숫자의 다양한 형태 변형을 흡수할 수 있는 최적의 특징을 찾기 위해서 다양한 특징 추출 알고리즘을 구현하였고, 이와 같은 특징을 신경회로망의 입력으로 학습시킨 후 인식 결과를 비교 분석하여 최적의 특징을 찾고자 하였다. 특징 추출을 위해서 구현한 알고리즘은 윤곽선을 이용한 4방향(수평, 사선, 수직, 역사선) 특징[3], 모포로지(Morphology)를 이용한 4방향(수평, 사선, 수직, 역사선) 특징[4, 5], 고리(Ring)를 이용한 투영(Projection) 및 부채꼴 망(Fan mesh) 특징[6, 7, 8, 9, 10], 교차 거리

(Cross Distance) +교차(Cross) +망(Mesh) +투영(Projection) 특징[11], 숫자 영상을 압축한 특징이다. 이와 같은 특징을 비교 분석한 결과는 윤곽선을 이용한 4방향 특징과 교차 거리 +교차 +망 +투영 특징이 구현한 다섯 가지의 특징 중에서 최적한 특징으로 나타났다. 그리고 다층 구조의 단일 신경회로망으로 구현시 발생하는 문제점을 해결하기 위하여 복합 특징을 소규모 신경회로망의 입력으로 모듈화시킨 모듈화된 신경회로망(Modularized neural networks)[12]로 다수결 투표(Majority Voting)와 신뢰도 지수(Reliability factor)를 이용하여 인식률이 개선되는 점을 제안하였다. 제안된 방법의 타당성을 검증하기 위하여 Concordia 대학교의 필기체 숫자 데이터 베이스에 적용한 결과, 정인식률은 97.1%, 기각률은 1.5%, 오인식률은 1.4%, 신뢰도는 98.5%가 나타났으며, 자체적으로 구축한 Dong-A 대학교의 필기체 숫자 데이터 베이스는 정인식률이 98%, 기각률은 1.2%, 오인식률이 0.8%, 신뢰도가 99.1%의 실험적 결과로 제안된 방식의 유용성을 확인하였다.

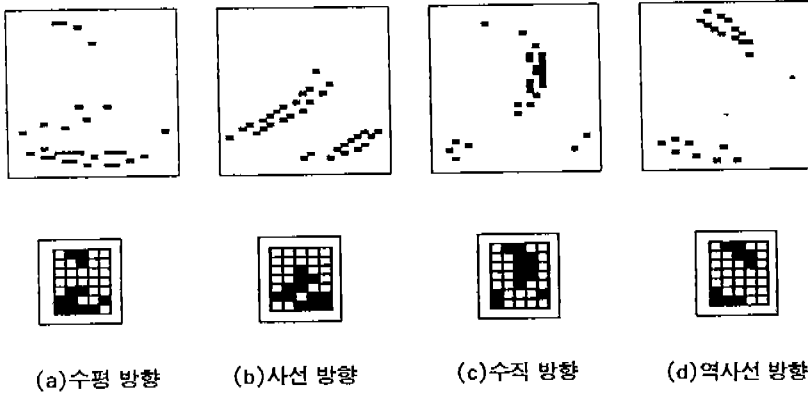
2. 특징 추출

2.1 윤곽선을 이용한 4방향 특징

윤곽선이란 화소가 0과 1로 구분되어 지는 선을 말한다. 윤곽선 추출은 (그림 1. (a))의 3×3 창인 체인 코드(Chain code)를 이용하여 숫자 영상을 왼쪽 위에서 아래로 주사(Scan)를 하여 검은 화소가 만나는 점을 시작점으로 코드화가 시작되고, 선을 따라 중심 화소에 이웃한 화소들을 관찰하면서 다음으로 진행할 화소를 결정한다. 예를 들면(그림 1. (a))의 8방향



(그림 1) 8방향 체인 코드 및 방향 추출
 (a) 8방향 체인 코드 (b) 방향 추출
 (Fig. 1) 8-Directional chain code and direction extracting
 (a) 8-Directional chain code (b) Direction extracting



(그림 2) 윤곽선을 이용한 4방향 특징
 (Fig. 2) 4-Directional feature using contour
 (a) Horizontal direction (b) Slant direction
 (c) Vertical direction (d) Reverse slant direction

체인 코드를 이용하면 (그림 1. (b))에서 555700으로 윤곽선의 방향 성분이 추출된다.

4방향 특징 추출은 숫자의 윤곽선을 따라가며 이웃하는 두 화소에 대한 수평, 수직, 사선, 역사선의 방향 성분을 추출한 다음, 각 방향 성분별로 7×5의 크기로 압축하여 (그림 2)와 같이 특징을 구한다.

$$\beta(A) = A - (A \oplus B) \tag{1}$$

$\beta(A)$ 를 (그림 4)와 같은 Structuring element(B)로써 각각 Opening하면 4방향 성분을 구할 수 있다.

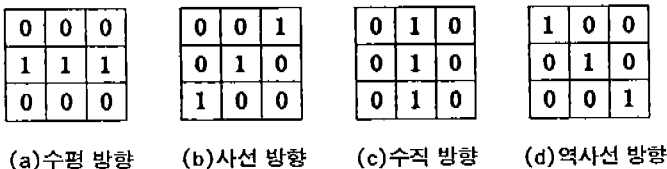
$$\beta(A) \circ B = (\beta(A) \ominus B) \oplus B \tag{2}$$

2.2 모포로지를 이용한 4방향 특징

모포로지 연산자들은 영상에서 기하학적 형태와 패턴으로 정의되어 있고 윤곽선 추출, 에지(Edge) 검출, 세션화동 영상 처리의 많은 분야에서 발전되고 있으며 간단한 논리적 연산과 병렬 처리 및 실시간 응용이 가능한 장점을 가지고 있다. 숫자 영상의 집합 A 의 윤곽선을 $\beta(A)$ 라면 다음 수식과 같이 (그림 3)의 Structuring element(B)로써 윤곽선을 구할 수 있다.

0	1	0
1	1	1
0	1	0

(그림 3) 스트럭처 엘리먼트 B
 (Fig. 3) Structuring element B



(그림 4) 4방향 스트럭처 엘리먼트 B
 (Fig. 4) 4-Directional structuring element B
 (a) Horizontal direction (b) Slant direction
 (c) Vertical direction (d) Reverse slant direction

추출된 4방향 성분을 7×5의 크기로 압축하여 특징으로 이용한다.

2.3 고리를 이용한 투영 및 부채꼴 망 특징

입력 영상의 위치, 크기의 확대 및 축소, 그리고 회전 변형에 비교적 강한 이 방법은 위치 정규화, 크기 정규화, 특징 추출을 한다.

2.3.1 위치 정규화

숫자 영상 위치에 상관없이 그 숫자를 영상 평면상의 원점을 중심으로 변환함으로써 위치의 변형을 흡수할 수 있다. 먼저 숫자 영상의 무게중심을 구하여 숫자 영상을 영상 평면상의 원점으로 변환시키도록 한다. 임의의 2차원 숫자 영상은 다음과 같이 2차원 좌표계로 표현될 수 있다.

(단, M 은 영상 패턴 크기이다.)

$$f(x, y) = \begin{cases} 1 & \text{(문자부분)} \\ 0 & \text{(배경부분)} \end{cases} \quad (1 \leq x, y \leq M) \quad (3)$$

이때, 숫자 영상의 무게중심 (\bar{X}, \bar{Y}) 는 다음과 같이 정의한다.

$$\begin{cases} \bar{X} = m_{10}/m_{00} \\ \bar{Y} = m_{01}/m_{00} \end{cases} \quad (4)$$

여기서 $m_{p,q} = \sum_{x=1}^M \sum_{y=1}^M x^p y^q f(x, y)$.

2.3.2 크기 정규화

일반적으로 필기체 숫자 영상은 크기가 다양한 형태 변형을 포함하고 있다. 이러한 변형에 무관하게 일관성 있는 특징 추출을 위해서는 고리를 이용한 크기 정규화 과정이 필요하다. 숫자 영상의 크기는 무게 중심으로부터 가장 먼 거리에 있는 검은 화소까지의 거리로 정의되는데 고리를 이용하여 숫자의 크기를 구한다. 정규화 하기 위한 크기를 D , 숫자 영상의 크기를 d_j 라 가정하면 D/d_j 의 크기로 정규화 된다.

본 논문에서는 Cheng[13]의 크기 정규화 알고리즘을 사용하였다.

2.3.3 고리를 이용한 투영 특징

먼저 입력 패턴의 무게 중심으로부터 거리를 1씩 증가시키면서 반지름이 r 인 고리를 만든다. 무게 중심 (\bar{X}, \bar{Y}) 을 극좌표계의 원점으로할 때 $(x_0 = \bar{X}, y_0 = \bar{Y})$, 패턴 영상은 다음과 같이 표현될 수 있다.

$$f(r, \theta) = \begin{cases} 1 & \text{(문자부분)} \\ 0 & \text{(배경부분)} \end{cases} \quad (5)$$

단, $0 \leq r \leq n, 0 \leq \theta \leq 2\pi$ 이며,

$$n = \sqrt{x_1^2 + y_1^2}, x_1 = \text{Max}(\bar{X}, M - \bar{X}), y_1 = \text{Max}(\bar{Y}, M - \bar{Y}).$$

입력 영상에 고리(Ring)를 이용하여 검정 화소를 고리층 별로 투영할 때 p_i 는 다음과 같다.

$$p_i = \sum_{\theta=0}^{2\pi} f(i, \theta), i = 0, 1, 2, \dots, n \quad (6)$$

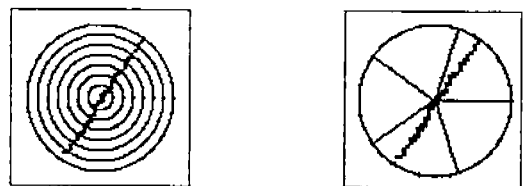
p_i 를 i 번째 고리(Ring)인 r_i 에 투영되는 검정 화소들의 수라 하면 p_i 를 (그림 5. (a))에서 보는 바와 같이 r_i 를 대표하는 특징이 된다.

2.3.4 고리를 이용한 부채꼴 망 특징

입력 영상에 고리를 이용하여 원하는 각도 $(\Delta\theta)$ 별로 검정 화소를 누적한 부채꼴 망 특징 M_i 는 다음과 같다.

$$M_i = \sum_{\theta=\Delta\theta i}^{\Delta\theta(i+1)} f(r, \theta) \quad r=0, 1, \dots, n \quad i=0, 1, \dots, (2\pi/\Delta\theta-1) \quad (7)$$

(그림 5. (b))는 필기체 숫자의 기울기를 감안하여 $\Delta\theta$ 값을 $2\pi/5$ 로 한 경우이다.



(a) 고리 투영 특징

(b) 부채꼴 망 특징

(그림 5) 고리 투영 특징과 부채꼴 망 특징
(Fig. 5) Ring projection feature and fan mesh feature
(a) Ring projection feature (b) Fan mesh feature

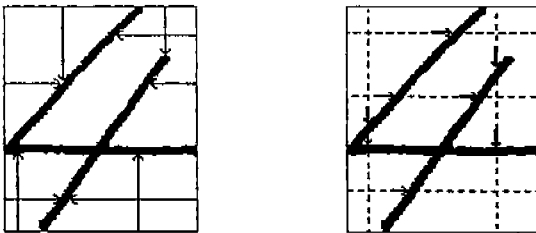
2.4 교차 거리 + 교차 + 망 + 투영 특징

2.4.1 교차 거리(Cross Distance) 특징

숫자 영상의 주어진 경계선으로부터 상, 하, 좌, 우로 검출선을 주사하였을 때 획과 만나는 거리를 특징으로 한다. 이 특징은 서로 다른 윤곽을 가진 숫자 영상의 구별은 쉽게 되나 비슷한 윤곽을 가진 숫자 영상의 구별이 어려운 단점을 가지고 있다.

2.4.2 교차(Cross) 특징

숫자 영상의 주어진 경계선으로부터 수직축과 수평축으로 검출선을 주사하였을 때 0에서 1로 바뀌는 수를 누적하여 특징으로 한다. 숫자의 기울어짐이나 위치 이동에 강한 특징을 가지고 있지만 비슷한 형태의 수치 구별은 어려운 단점을 가지고 있다.



(a) 교차 거리 특징

(b) 교차 특징

(그림 6) 교차 거리 특징과 교차 특징

(Fig. 6) Cross distance feature and Cross feature

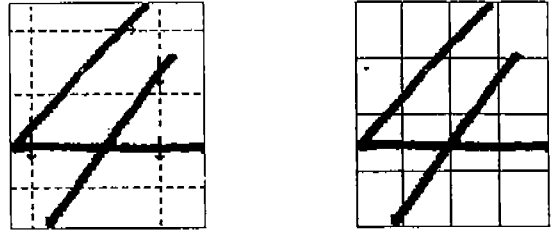
(a) Cross distance feature (b) Cross feature

2.4.3 투영(Projection) 특징

숫자 영상을 수직과 수평으로 검출선을 주사하였을 때 검출되는 획의 수를 세어서 특징으로 한다. 숫자의 위치 이동이나 약간의 기울어짐에도 강한 특성을 보이지만 비슷한 형태의 숫자를 구분하기 어려운 단점도 있다.

2.4.4 망(Mesh) 특징

숫자 영상을 수직과 수평으로 균등한 크기나 비균등적 크기의 부 영역으로 나누어 각 부 영역에 대하여 획의 수를 누적하여 특징으로 이용한다. 이 특징은 획의 이동에 약한 단점을 가지고 있다.



(a) 투영 특징

(b) 망 특징

(그림 7) 투영 특징과 망 특징

(Fig. 7) Projection feature and Mesh feature

(a) Projection feature (b) Mesh feature

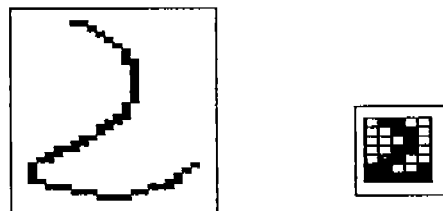
2.5 숫자 영상을 압축한 특징

2.5.1 세선화

숫자 영상을 압축하기 위해서는 데이터 양을 최소화시켜야 할 필요성이 있다. 따라서 숫자 영상을 압축하기 이전에 세선화가 전처리로 필요하다. 그러나 세선화 과정에서 필요한 정보를 잃어버리는 경우가 있으며 이에 대한 방안으로 세선화 알고리즘 연구가 많이 되어 왔다. 본 논문에서는 잔가지의 발생이 드물고 교차점과 분기점의 처리가 비교적 잘된다고 알려진 T. Y. Zhang[14]이 개발한 알고리즘을 사용하였다. (그림 8. (a))는 Zhang의 알고리즘에 의하여 추출된 세선화의 결과이다.

2.5.2 숫자 영상을 압축한 특징

필기체 숫자 영상은 크기가 다양한 형태로 변형을



(a) 세선화의 결과

(b) 7×5 크기로 압축한 특징

(그림 8) 세선화 및 7×5 크기로 압축한 특징

(Fig. 8) Thinning and Compressed feature by 7×5 magnitude

(a) Result after thinning

(b) Compressed feature by 7×5 magnitude

포함하고 있으며 이러한 변형에 무관하게 일관성 있는 특징 추출을 위해서는 정규화 과정이 필요하다. 입력 영상의 x축 길이를 M_x , y축 길이를 M_y 로 가정하고 정규화하기 위한 x축 길이를 N_x , y축 길이를 N_y 로 가정하면 다음 수식과 같이 표준 크기로 압축된다.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} N_x/M_x & 0 \\ 0 & N_y/M_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (8)$$

세션화된 숫자 영상을 식 (8)에 의해서 (그림 8. (b))와 같이 7×5 의 크기로 압축하여 특징으로 이용한다.

3. 신뢰도 지수 및 모듈화된 신경회로망

3.1 신뢰도 지수

신경회로망의 특징 중의 하나는 인식하고자 하는 패턴이 학습된 패턴과 비교해서 근사한 값을 나타낼 때 출력 노드의 값이 학습할 때의 목표치와 일치하려고 작용한다는 점이다. 이러한 특징을 이용하면 인식이 가능한 범위와 인식 불능에 대한 판별식을 구할 수 있다. 본 논문에서는 신경회로망의 출력층은 숫자의 종류가 10개로 구성되므로 출력층의 노드는 10개이며 분류된 활성화 노드는 '1'로 나머지 노드는 '0'으로 나타나도록 학습시켰다. 가장 간단하면서 오인식에 대한 부담이 적은 경우에는 출력층의 10개 노드 중 최고값 M_1 로 신뢰도 지수를 구할 수 있다.

$$RF_1 = M_1 \quad (9)$$

두 번째로 출력층의 10개 노드중 최고값 M_1 과 두 번째로 큰 값 M_2 의 차로서 신뢰도 지수를 구할 수 있다.

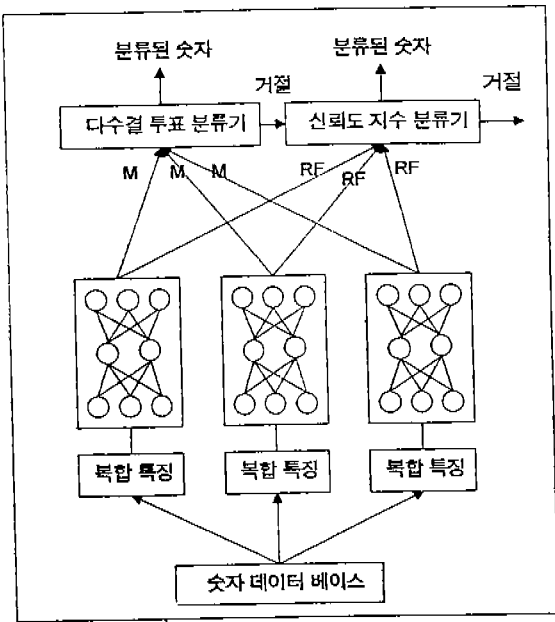
$$RF_2 = M_1 - M_2 \quad (10)$$

Takahashi에 의해 제안된 신뢰도 지수[15]는 학습이 잘된 숫자는 해당되는 노드에서 '1'에 가까운 값이 나타나고 나머지 노드는 작은 값으로 분포되는데 근거를 두고 신뢰도 지수를 다음 식과 같이 나타내었다.

$$RF_3 = M_1 + (1 - M_2/M_1) \quad (11)$$

3.2 모듈화된 신경회로망

필기체 숫자는 환경이나 필자에 따라 변화가 불규칙하기 때문에 한 두개의 특징을 사용한 신경회로망은 높은 인식률을 기대하기 어렵다. 단일 특징 자체로 인식률이 우수한 몇 개의 특징을 합쳐서 만들어지는 복합 특징을 다층 구조의 단일 신경회로망에 입력으로 사용하면 단일 특징을 다층 구조의 단일 신경회로망에 입력으로 사용하는 것보다 일반적으로 인식률이 약간 증가하지만, 인식률의 한계가 나타나고, 신경회로망의 연결 강도가 증가하므로 장시간의 학습이 필요하다. 이러한 문제점을 해결하기 위해서 소규모의 모듈화된 신경회로망을 여러개 구성한 다음 이들로부터 최종 결과를 선택하는 방법으로 인식기를 구성한다. 이렇게 구성된 시스템을 모듈화된 신경회로망(Modularized neural network)이라고 한다. 모듈화된 신경회로망의 각 모듈을 부 신경회로망으로 정의하며, 부 신경회로망의 입력은 성격이 서로 다른 특징을 이용한다. 이때 개별적인 부 신경회로망의 인식률이 너무 낮으면 전체 인식률이 떨어지므로 부 신경회로망 자체의 인식률이 높은 복합 특징을 선택하는 것도 중요하다. 본 논문에서는 부 신경회로망의 입력으로 사용되는 복합 특징을 (그림 9)와 같이 세개의 그룹으로 나누었다. 대상 숫자에 대하여 모듈화된 신경회로망의 최종 분류 방법은 독립적으로 인식을 수행하는 부 신경회로망의 출력으로부터 구할 수 있으며, 다수결 투표(Majority Voting) 그리고 신뢰도 지수(Reliability Factor)순으로 처리한다. 일차적으로 대상 숫자에 대하여 각 부 신경회로망의 출력 노드에서 최대값(M)으로 나타나는 숫자를 인식하고, 세개의 부 신경회로망의 인식 결과가 모두 같으면 다수결 투표에 의하여 대상 숫자를 분류하고, 부 신경회로망의 출력 결과중 한 개라도 틀리면 분류를 거절한다. 이차적으로 거절된 숫자만 각각의 부 신경회로망에서 신뢰도 지수를 구하고, 구해진 신뢰도 지수가 정해진 임계값의 범위 내이면 대상 숫자를 분류하고 임계값의 범위를 벗어나면 분류를 거절한다. 대규모 복합 특징을 사용하는 다층 구조의 단일 신경회로망에서 단점으로 나타나는 장시간의 학습 시간을 모듈화된 신경회로망은 부 신경회로망이 독립적으로 구성되어 있기 때문에 부 신경회로망의 입력 수를 줄일 수가 있다. 따라서 신경회로망의 연결 강도가 감소하므로 학습 시간을 줄일 수가 있는 장점을 가지고 있다.



(그림 9) 모듈화된 신경회로망
(Fig. 9) Modularized Neural Network

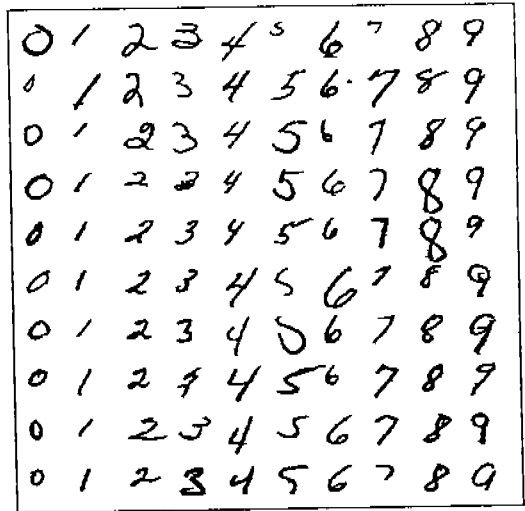
4. 실험 결과 및 고찰

4.1 실험 환경

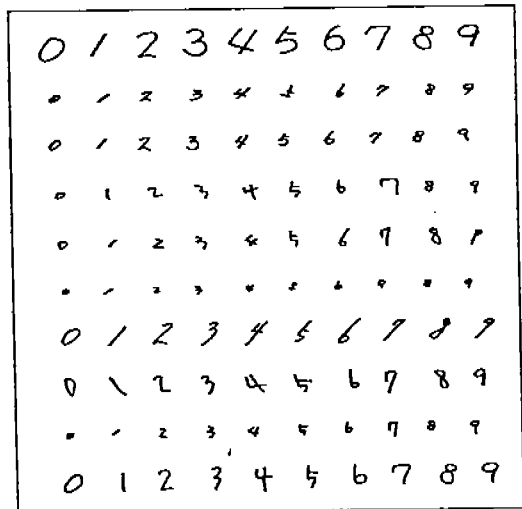
본 논문에서 사용된 필기체 숫자는 자체적으로 구축한 Dong-A 대학교(Dong-A) 필기체 숫자 데이터 베이스의 2000자와 성능을 객관적으로 검증하기 위해

<표 1> 특징들의 정의
(Table 1) Define of features

정 의	특 징
FEATURE 1	윤곽선을 이용한 4방향 특징
FEATURE 2	모포로지를 이용한 4방향 특징
FEATURE 3	고리를 이용한 투영 및 부채꼴 망 특징
FEATURE 4	교차 거리 +교차 +망 +투영 특징
FEATURE 5	숫자 영상을 압축한 특징
COMPLEX 1-5	FEATURE 1 +FEATURE 5
COMPLEX 4-5	FEATURE 4 +FEATURE 5
COMPLEX 1-4	FEATURE 1 +FEATURE 4
COMPLEX 1-4-5	FEATURE 1 +FEATURE 4 +FEATURE 5



(그림 10) Concordia 대학교의 필기체 숫자 데이터 베이스
(Fig. 10) Handwritten numeral database of Concordia university



(그림 11) 동아 대학교의 필기체 숫자 데이터 베이스
(Fig. 11) Handwritten numeral database of Dong-A university

서 캐나다 Concordia 대학교(CENPARMI) 필기체 숫자 데이터베이스중 일부분인 2000자를 사용하였다. 각 데이터 베이스의 1000자는 학습용 데이터로 사용하였고, 학습에 불참한 1000자를 테스트용 데이터로 사

용하였으며, 인식기는 3층 구조의 단일 신경회로망과 모듈화된 신경회로망을 사용하였다. 실험을 위해서 사용된 시스템은 IBM-PC Pentium(120Mhz)이며 C언어를 사용하여 알고리즘을 구현하였다. 그리고 표기의 편의상 각각의 특징을 <표 1>과 같이 정의한다.

<표 1>에서 COMPLEX 1-5는 FEATURE 1과 FEATURE 5를 합쳐서 형성되는 복합 특징으로 정의하였다. (그림 10)은 실험에 사용된 CENPARMI 데이터 베이스를 나타내었고, (그림 11)은 Dong-A 데이터 베이스를 나타내었다.

4.2 필기체 숫자의 인식 실험 및 결과 분석

CENPARMI의 필기체 숫자 데이터 베이스에 대하여 FEATURE 1을 추출하여, 3층 구조의 단일 신경회로망으로 학습한 후 인식한 결과를 <표 2>에 나타내었다.

<표 2> CENPARMI 데이터 베이스에 대하여 FEATURE 1을 사용한 인식 결과

<Table 2> Recognition results using FEATURE 1 for CENPARMI database

in	out	0	1	2	3	4	5	6	7	8	9
0	96	0	0	0	0	1	2	0	1	0	0
1	0	97	1	0	1	0	0	1	0	0	0
2	0	0	90	1	0	0	0	5	4	0	0
3	0	0	1	92	0	2	0	0	3	2	0
4	0	0	0	0	92	0	3	1	1	3	0
5	0	0	0	2	0	92	4	0	2	0	0
6	1	0	0	0	0	1	97	0	1	0	0
7	0	0	0	0	1	0	0	91	2	6	0
8	0	0	1	2	0	0	1	0	92	4	0
9	0	0	0	0	1	0	0	0	3	96	0
Total recognition rate 93.5%											

<표 2>에서 비슷한 형태의 획을 갖는 (2, 7), (5, 6), (4, 7, 9), (8, 9)등이 상호간의 오인식의 분포가 큰 것을 알 수 있으며 (0, 1, 6, 9)는 윤곽선을 이용한 방향성 특징으로 분류가 잘되는 것으로 나타났다. 유사한 획을 갖는 필기체 숫자의 인식은 방향성만으로는 해

결할 수 없으며, 보다 정확한 분류를 위해서는 복합적인 특징이 요구된다는 것을 알 수 있다. 상용화되는 인식기에서 중요한 부분을 차지하는 것은 오인식률을 최소화시키는 것이다. 따라서 오인식률을 최소화하기 위해서는 신뢰도 지수의 임계값을 설정하여 임계값 이하인 경우 인식을 거부함으로써 오인식률을 줄일 수 있으며 신뢰도를 높일 수 있다. <표 3>은 CENPARMI의 필기체 숫자 데이터 베이스에 대하여 FEATURE 1을 추출하여, 3층 구조의 단일 신경회로망으로 인식할 때 신뢰도 지수(RF₃)를 이용하여 분류한 결과를 나타내었다.

<표 3> FEATURE 1에 대하여 RF₃를 사용한 인식 결과
<Table 3> Recognition results using RF₃ for FEATURE 1

신뢰도 지수 (RF ₃)	정인식률 (%)	기각률 (%)	오인식률 (%)	신뢰도 (%)
0.0	93.5	0.0	6.5	93.5
0.5	93.0	1.1	5.9	94.0
0.7	91.7	3.6	4.7	95.2
0.9	88.9	8.3	2.8	96.9
1.1	87.2	10.5	2.3	97.4
1.3	79.0	19.9	1.1	98.6
1.5	66.4	33.0	0.6	99.1

RF₃의 값에 따라서 인식률의 차이가 나타났으며, 이와 같은 특성을 모듈화된 신경회로망에서 이차 분류를 할 때 이용한다. <표 4>와 <표 5>는 CENPARMI 필기체 숫자 데이터 베이스, Dong-A 필기체 숫자 데

<표 4> CENPARMI 데이터 베이스에 각각의 특징들을 사용한 성능 비교

<Table 4> Performance comparison using each features for CENPARMI database

특징	특징 벡터 수	평균시간 (Sec)	오인식률 (%)	정인식률 (%)
FEATURE 1	140(7×5×4)	0.008	6.5	93.5
FEATURE 2	140(7×5×4)	0.031	7.3	92.7
FEATURE 3	32(16+16)	0.157	11.0	89.0
FEATURE 4	104(36+6+32+30)	0.005	5.5	94.5
FEATURE 5	35(7×5×1)	0.011	9.0	91.0

〈표 5〉 Dong-A 데이터 베이스에 각각의 특징들을 사용한 성능 비교

〈Table 5〉 Performance comparison using each features for Dong-A database

특징	특징 벡터 수	평균시간 (Sec)	오인식률 (%)	정인식률 (%)
FEATURE 1	140(7×5×4)	0.008	5.5	94.5
FEATURE 2	140(7×5×4)	0.042	6.6	93.4
FEATURE 3	32(16+16)	0.104	9.8	90.2
FEATURE 4	104(36+6+32+30)	0.005	4.2	95.8
FEATURE 5	35(7×5×1)	0.015	7.9	92.1

이더 베이스에 대하여 본 논문에서 구현한 다섯 가지 특징으로 성능을 비교하였으며, 인식기는 3층 구조의 단일 신경회로망을 이용하였다. 여기서 평균 시간은 한 개의 숫자에 대하여 특징을 추출하는데 소요되는 시간을 나타내었다.

본 논문에서 구현한 특징 중에서 FEATURE 1과 FEATURE 4가 정인식률 그리고 인식 시간 면에서 필기체 숫자 인식에 최적한 특징으로 나타났으며 FEATURE 5는 특징 벡터수에 비하여 인식률이 높은 것으로 나타났다. Dong-A 데이터 베이스가 CENPARMI 데이터 베이스를 이용한 경우보다 전체적으로 인식률이 높게 나타났다. 〈표 6〉은 실험적 결과에서 우수한 특징으로 판단된 FEATURE 1, FEATURE 4, FEATURE 5를 조합해서 형성되는 복합 특징을 3층 구조의 단일 신경회로망으로 인식한 결과를 나타내었다.

비교적 복합 특징의 인식률은 〈표 4〉, 〈표 5〉의 특징 보다는 인식률이 높은 것으로 나타났으며, 이것은 신

경회로망의 특징 벡터수를 증가하면 어느 정도 인식률이 향상하는 것을 알 수 있다. 그러나 COMPLEX 1-4-5를 사용한 3층 구조의 단일 신경회로망은 특징 벡터수에 비하여 더이상 인식률이 증가하지 않는 한계점이 발생했다. 이와 같은 문제를 개선하기 위하여 모듈화된 신경회로망에 COMPLEX 1-5, COMPLEX 4-5, COMPLEX 1-4를 부 신경회로망의 입력으로 하고, 일차로 다수결 투표에 의하여 분류하면서 거절된 숫자는 이차로 신뢰도 지수(RF₃)가 0.7이상이면 분류하고 아니면 거부하여 인식한 실험적 결과를 〈표 7〉에 나타내었다.

〈표 7〉 모듈화된 신경회로망의 인식 결과

〈Table 7〉 Recognition results of modularized neural network

데이터 베이스	정인식률 (%)	기각률 (%)	오인식률 (%)	신뢰도 (%)
CENPARMI database	97.1	1.5	1.4	98.5
Dong-A database	98.0	1.2	0.8	99.1

모듈화된 신경회로망을 이용한 인식률이 COMPLEX 1-4-5를 사용한 3층 구조의 단일 신경회로망보다도 높은 것으로 나타났다. 이와 같은 결과는 다층 구조의 단일 신경회로망에서 발생하는 인식률의 한계점을 극복할 수 있다는 것을 알 수 있었다. 그리고 〈표 7〉의 기각된 숫자에 대한 인식은 유사한 형태의 획을 갖는 숫자만을 분류할 수 있는 전용 인식기를 개발함으로써 인식률을 높일 수 있을 것으로 기대된다.

5. 결론

본 논문에서는 필기체 숫자의 다양한 형태 변형을 흡수할 수 있는 특징을 찾기 위해서 다섯 종류의 특징을 구하여 비교 분석하였고, 분석 결과 성능이 우수한 특징을 조합한 복합 특징을 모듈화된 신경회로망의 입력으로 이용하였으며, 이 방법이 다층 구조의 단일 신경회로망에서 발생하는 문제점을 해결하고 인식률이 증가하는 점을 제안하였다. 필기체 숫자의 다양한 형태 변형을 흡수할 수 있는 특징을 찾기 위한 실험에서 FEATURE 1과 FEATURE 4가 인식률,

〈표 6〉 복합 특징의 인식 결과

〈Table 6〉 Recognition results of complex features

복합 특징	특징 벡터 수	정인식률(%)	
		CENPARMI database	Dong-A database
COMPLEX 1-5	175(140+35)	95.4	96.3
COMPLEX 4-5	175(104+35)	95.3	97.1
COMPLEX 1-4	244(140+104)	95.8	96.6
COMPLEX 1-4-5	279(140+104+35)	95.7	97.0

인식 시간 측면에서 효율적인 특징으로 나타났다. 그리고 FEATURE 5는 특징 백터수에 비하여 인식이 높게 나타났다. 인식률의 향상을 위해서 효율적인 특징을 조합해서 형성되는 복합 특징(COMPLEX 1-4, COMPLEX 1-5, COMPLEX 4-5)을 다층 구조의 단일 신경회로망의 입력으로 학습시킨 후 인식 결과는 FEATURE 1과 FEATURE 4를 이용한 것보다는 약 2%정도의 정인식률이 향상되는 점을 확인할 수 있었다. 이것은 신경회로망의 특징 백터수를 증가하면 어느 정도 인식이 증가하는 것을 알 수 있다. 그러나 다층 구조의 단일 신경회로망에 COMPLEX 1-4-5를 사용하여 인식한 결과는 Concordia 대학교의 숫자 데이터 베이스는 정인식률이 95.7%, Dong-A 대학교의 숫자 데이터 베이스는 정인식률이 97%로 나타났다. 이 결과를 COMPLEX 1-4, 1-5, 4-5와 비교하면 인식률의 증가는 거의 없었으며 학습이 장시간 소요되었다. 이와 같은 장시간의 학습과 인식률의 한계점을 극복하기 위하여 다층 구조의 단일 신경회로망을 모듈화하였다. 모듈화된 신경회로망에서 부 신경회로망의 입력은 COMPLEX 1-5, COMPLEX 4-5, COMPLEX 1-4를 사용하였으며, 최종 출력은 다수결 투표 그리고 신뢰도 지수를 사용하여 인식한 결과, Concordia 대학교의 숫자 데이터 베이스는 정인식률 97.1%, 기각률 1.5%, 오인식률 1.4%, 신뢰도 98.5%가 나타났으며, Dong-A 대학교의 숫자 데이터 베이스는 정인식률 98%, 기각률 1.2%, 오인식률 0.8%, 신뢰도 99.1%의 실험적 결과가 나타나 제안된 방식의 우수성을 확인하였다.

참 고 문 헌

[1] 김영준, 이성환, "유전자 알고리즘과 결합된 다층 클러스터 신경망을 이용한 무제 약 필기체 숫자의 오프라인 인식", 한국 정보 과학회 논문지, 제 21권, 제8호, pp.1468-1478, 1994.
 [2] J. S. N. Jean and Y. C. Chan, "Input Representation and Output Voting Considerations for Handwritten Numeral Recognition with Back-Propagation," Proc. of IJCNN, Sandiego, Vol.1, pp.408-411, 1990.
 [3] C. Y. SUEN and T. A. MAL, "Computer Recog-

inition of Unconstrained Handwritten Numerals," Proceedings of the IEEE, Vol.80, No.7, pp. 1162-1180, july. 1992.
 [4] R. R. Haralick, S. R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology," IEEE trans. Patt. Anal. Mach. Intell., Vol.PAMI-9, No.4, pp.532-550, 1980.
 [5] J. Serra, Image Analysis and Mathematical Morphology, Academic Press:San Diego, CA, chap. XII, pp.23-56, 1982.
 [6] A. Khotanzad and Y. H. Hong, "Rotation Invariant Image Recognition using Features selected via systematic method," Pattern Recognition, Vol.23, No.10, pp.1089-1101, 1990.
 [7] K. Mersereau and G. M. Morris, "Scale, Rotation, and Shift Invariant Image Recognition," Applied Optics, Vol.20, No.13, pp.2238-2342, 1986.
 [8] Y. Y. Tang, H. D. Cheng and C. Y. Shen, "Transformation-Ring-Projection(T RP) Algorithm and Its VLSI Implementation," Word-Scientific Series in Computer Science, Vol.30, 1991.
 [9] R. H. Wu and H. Stark, "Rotation and Scale Invariant Recognition of Images," Proc. 8th. int. Conf. On Pattern Recognition, Paris, France, pp. 92-94, Oct. 1986.
 [10] 이성환, 박희선, "고리 투영을 이용한 위치, 크기 회전 변형에 무관한 패턴 인식", 인지 과학회 논문지, 제13권, 제1호, pp.139-160, Jun. 1986.
 [11] 김우태, 윤병식, 박인규, 진성일, "인쇄체 한글 문자 인식을 위한 피쳐 추출 및 성능 비교," 한국 정보과학회 논문지, 제20권, 제8호, pp.1103-1111, 1993.
 [12] 류강수, 김우태, 진성일, "다층 특징과 모듈화된 신경회로망을 이용한 인쇄체 및 필기체 혼용 숫자 인식", 전자공학회 논문지, 제32권, 제10호, pp.101-111, 1995.
 [13] H. D. Cheng, Y. Y. Tang and C. Y. Suen, "Parallel image transformation and its VLSI implementation," Pattern Recognition, Vol.23, No. 10, pp.1113-1129, 1990.
 [14] T. Y. Zhang and P. A. S. P. Wang. "A Maxi-

mum Algorithm for Thinning Digital Pattern," Proc. 9th Int. Conf. on Pattern Recognition, Rome, Italy, pp.942-944, 1988.

[15] H. Takahashi, "A Neural Network OCR using Geometrical and Zonal Pattern Features," proc. of ICDAR '91, Paris, Vol.2, pp.821-828, 1991.



김 두 영

1972년 부산대학교 전자공학과 졸업(공학사)

1975년 부산대학교 대학원 전자공학과 졸업(공학석사)

1983년 경북대학교 대학원 전자공학과 박사 수료

1977년 3월~현재 동아대학교 전자공학과 교수

관심분야: 디지털신호처리, 영상처리 및 컴퓨터비전



박 창 순

1983년 경남대학교 전자공학과 졸업(공학사)

1986년 동아대학교 대학원 전자공학과 졸업(공학석사)

1993년 동아대학교 대학원 전자공학과 박사과정 수료

1990년~현재 창신전문대학 전산정보처리과 조교수
관심분야: 문자 인식, 패턴 인식, 영상 처리