

論文95-32B-10-11

다중 특징과 모듈화된 신경회로망을 이용한 인쇄체 및 필기체 혼용 숫자 인식

(Recognition of Printed and Handwritten Numerals Using Multiple Features and Modularized Neural Networks)

柳 江 洙 *, 金 佑 泰 **, 秦 成 一 ***

(Gang Soo Ryu, Woo Tae Kim, and Sung Il Chien)

요 약

본 논문에서는 인쇄체와 필기체가 혼용된 숫자를 인식하기 위한 인식기로서 신경회로망을 사용하였다. 인식기는 총 4개의 부인식기로 모듈화되어 있으며, 각 부인식기의 입력은 서로 다른 특징을 사용하였다. 필기체 숫자의 경우 부인식기를 결합한 인식기의 인식률은 각 부인식기의 인식 결과보다 높게 나타났다. 인식기 출력은 신뢰도 지수에 의한 최고치 RF 선택법으로 결정하였다. 인쇄체 1,080자와 필기체 3,000자에 대한 인식 실험 결과는 98.2%의 인식률을 보였다. 또한 임계치로 사용한 RF가 1.5일때 인식 결과는 인식률 97%, 오인식률 1% 및 인식 거절률 2%로 나타났다.

Abstract

In this paper, we describe a modularized neuroclassifier for enhancing the recognition accuracy of mixed printed and handwritten numerals. This classifier combines four modularized subclassifiers using multi-layer perceptron module. The input of each subclassifier is comprised of a group of specialized feature sets. On applying this method to combining several subclassifiers for unconstrained handwritten numerals, the experimental result shows that the performance of individual subclassifier can be improved. In winner-take-all voting method, the result of subclassifier having the highest RF value is selected as the output. The generality of this classifier is tested with 1,080 printed and 3,000 handwritten numerals that was not shown in training the neural networks. Experimental results show 98.2% recognition rate. The typical recognition test with a threshold value(RF=1.5) has shown 97% recognition, 1% substitution and 2% rejection rates.

* 正會員, 龜尾專門大學 電子通信科

(Dept. of Elec. Comm., Kumi Junior College)

** 正會員, 昌信專門大學 電算情報處理科

(Dept. of Computer Science, Changshin Junior College)

*** 正會員, 慶北大學校 電子工學科

(Dept. of Elec. Eng., Kyungpook Nat'l Univ.)

接受日字: 1994年8月27日, 수정완료일: 1995年10月2日

I. 서론

최근 컴퓨터의 급속한 보급 및 자동화의 영향으로 문자 인식에 관한 연구가 활발히 진행되고 있다. 숫자의 인식은 인쇄체 숫자 인식과 필기체 숫자 인식으로 구별된다. 인쇄체 숫자 영상의 변화가 필기체 숫자의 영상에 비해 적으므로 인쇄체 숫자의 인식은 필기체 숫자에 비해 쉽다. 그러나, 인쇄체 숫자는 형태가 명조체, 로만체, 이탤릭체 및 은행 지로용 OCR체 등 다양

하며, 숫자 영상은 영상 획득 과정에서 회전의 가능성 및 문서 취급시 오손등에 기인한 잡음의 영향으로 변형될 수 있다. 필기체 숫자의 인식은 이상의 인쇄체 숫자 인식시에 발생할 수 있는 문제점외에도 필기자의 습관 및 건강 상태등에 관련된 필기자 특성 및 필기 도구, 필기 용지의 질감 특성등 영상 획득 방법에 따라 달라질 수 있다. 따라서 숫자의 인식은 숫자의 형태 및 숫자의 영상에 다양한 변형이 발생하므로 어려움이 따른다. 문자 인식 분야에 종사하는 사람이 필기체 문자를 완성된 문장중의 문자가 아닌 개별 문자로 판독할 때 오인식 확률은 약 4%로 조사되었다.^[11]

인식 시스템의 성능은 인식기의 종류, 인식기를 구성하는 방법 및 인식기에 사용되는 특징에 따라 좌우된다. 최근의 문자 인식 분야에서 인식기의 구성은 하나의 인식기를 사용하기보다는 여러개의 인식기를 결합함으로써 신뢰성이 높은 인식기를 구현하는 경향이 있다.^[12] 인식기의 종류는 문자를 구성하는 화소의 특징에 근거한 인식기(feature vector based classifier)^[13,41], 문자의 형태 특성에 근거한 구조적인 인식기(syntactic-and-structural based classifier)^[15,61] 및 이들 두가지를 혼합한 형태의 인식기^[7,8]로 나눌 수 있다. 이들 인식기들은 하나의 인식기로 구성하거나 또는 다수개의 부인식기를 결합하여 구성할 수 있다. 고차의 입력을 가지는 단일 인식기는 구현의 어려움, 계산의 복잡성 및 정확도 문제등이 발생할 수 있다. 이러한 문제점은 단일 인식기를 다수개의 부인식기들로 나누고, 각 부인식기의 입력으로 고차의 입력 벡터를 다수개의 저차 벡터로 분할하여 사용함으로써 해결할 수 있다. 이와같은 부인식기를 결합한 인식기는 구현 및 계산의 용이성뿐만 아니라 부인식기 각각의 인식 결과에 비해서 우수한 성능을 나타낸다.^[21] 그러나, 결합형 인식기가 좋은 성능을 가지기 위한 조건은 다음과 같다. 첫째, 결합형 인식기에 사용된 부인식기는 부인식기들 사이에 상호 보완 관계를 가져야 한다. 둘째, 각 부인식기는 전문성을 가지는 모듈(expert module)로 구성되어야 한다. 셋째, 각 부인식기의 입력으로 사용하는 특징들은 상호 보완적이어야 한다. 따라서 이상의 조건을 충족시키는 결합형 인식기는 구현시에 많은 어려움이 따른다.

본 논문은 신용 카드의 전표 및 편지 봉투의 우편 번호등과 같이 필기체 숫자와 인쇄체 숫자가 혼용된 응용 분야에 적용이 가능한 범용 숫자 인식 시스템에

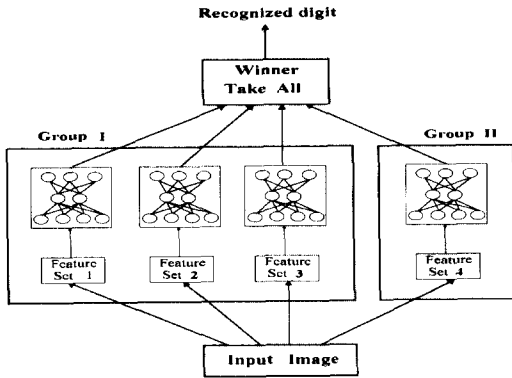
관한것이다. 필기체와 인쇄체가 혼용된 문서를 인식하기 위한 기존의 방법은 필기체와 인쇄체를 사전에 구분하여 별개의 인식기를 사용하였다.^[7] 이 방법은 전처리 과정에서 사람이 숫자체를 구분하여야 하며, 숫자체에 따라 별개의 인식기를 사용하므로써 인식 시스템의 규모가 커지는 문제점이 있다. 따라서 본 논문에서는 이러한 문제점을 해결하기 위하여 필기체와 인쇄체를 구분하지 않고 인식할 수 있는 범용 숫자 인식 시스템을 구성하였다. 인식기는 영상의 화소 특징에 근거한 인식기의 일종인 신경회로망을 이용하였다. 신경회로망은 적응학습 능력과 결합 복구 능력으로 인해 문자 인식에 많이 응용되고 있다. 이러한 신경회로망의 성능은 문자의 특징을 잘 나타낼 수 있는 특징을 선택하는 문제와 직결된다. 신경회로망의 입력은 하나의 특징으로 숫자의 특성을 잘 표현하기 어려움에 따라 다양한 특징을 추출하여 이들 특징들을 조합한 복합 특징을 사용한다. 복합 특징을 사용하여 구성된 하나의 대규모 신경 회로망은 시스템의 규모가 커져 학습에 어려움이 있을뿐만 아니라 잘못 선정된 특징이 전체 시스템에 나쁜 영향을 미칠 수 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 신경회로망을 모듈화하여 소규모 부인식기를 여러개 구성한 다음 이들의 결과를 조합하는 방법을 사용하였다. 모듈화된 신경회로망은 모든 부인식기가 같은 종류이므로 서로 다른 여러 종류의 부인식기를 모듈화 구조로 사용할때 발생할 수 있는 문제점을 해결할 수 있다. 또한 여러 종류의 인식기들을 결합한 인식기는 최종 출력을 결정하는데 어려움이 있다.^[2] 본 논문에서 출력을 결정하기 위해 사용한 신뢰도 지수(reliability factor : RF)는 오인식을 최소화하는 방법으로서 각 부인식기의 출력값이 정해진 신뢰도 지수 이상이면 인식, 이하이면 인식을 거부하는데 사용된다. 신뢰도 지수는 인식기의 인식률, 인식 거절률 및 오인식에 큰 영향을 미치므로 신뢰도 지수의 정의 및 신뢰도 지수의 임계치를 설정하는 것이 중요하다.^[9] 본 논문은 신경회로망의 입력으로 영상의 통계적 특징을 나타내는 6종류의 특징을 사용하였다. 신경회로망은 입력으로 사용하는 특징의 종류가 많음에 따라 특징별로 부인식기를 4개 구성하여 모듈화된 구조를 갖는다.

인식기의 출력은 부인식기의 출력에 대한 신뢰도 지수를 비교하여 최고의 신뢰도 지수를 나타내는 부인식기의 출력값을 선택하였다.

II. 모듈화된 신경망 기반의 숫자 인식 시스템

1. 인식 시스템의 구성

일반적으로 문자 인식 시스템은 영상의 획득 과정, 입력된 문서 영상으로부터 문자 영상의 추출 및 특징 추출의 전처리 과정, 인식 과정 및 오인식에 대한 후처리 과정으로 구성된다. 본 논문에서 사용한 인식 시스템의 구성은 그림 1에 나타난 것과 같다. 영상의 입력은 비교적 적은 스캔 시간으로 양질의 숫자 영상을 획득할 수 있는 플랫폼 베드 스캐너를 사용하였다. 전처리 과정은 입력된 영상으로부터 개별 숫자의 분리 및 분리된 숫자로부터 특징을 추출하는 것이다. 개별 숫자의 분리는 문서 영상에 대해 먼저 수평축을 따라 투영값을 구하여 숫자열을 추출한다. 이로부터 수직축으로 검출선을 주사하였을 때 화소의 누적 분포를 조사하여 숫자 영역과 숫자 사이의 공백 영역을 구별함으로써 수행된다.



Subclassifier group I: trained by hand written numerals

Subclassifier group II: trained by printed numerals

그림 1. 숫자 인식 시스템의 구성도

Fig. 1. Block diagram of numerals recognition system.

특징의 추출은 분리된 개별 숫자로부터 숫자를 대표할 수 있는 특성을 찾는 과정으로서, 추출된 특징은 인식기의 입력으로 사용된다.

신경회로망은 필기체 숫자의 특징으로 학습한 3개의 부인식기와 인쇄체 숫자의 특징으로 학습한 하나의 부인식기로 구성되어진 모듈화된 구조를 갖는다. 이때 개

별 신경회로망은 1개의 은닉층을 갖는 다층 퍼셉트론을 사용하였다. 인식기에서 숫자의 인식은 각 신경회로망의 출력으로부터 구해진 신뢰도 지수를 기준으로 결정하였다.

2. 다중 특징 추출

특징은 인식 시스템의 성능에 결정적인 영향을 미치므로 특징의 선택은 어려운 문제이다. 본 논문에서는 망 특징(mesh feature)^[9], 교차 특징(cross feature)^[9], 투영 특징(projection feature)^[9], 교차 거리 특징(cross distance feature)^[9] 및 특성 궤적 특징(characteristic-loci feature)^[10,11]을 사용하였다.

1) 망 특징

망 특징은 개별 숫자 영상의 MBR에 대해 가로와 세로를 일정 비율의 부분 영역으로 나누었을 때 각 부분 영역의 검은 화소를 합한 것이다. 이 특징은 부분 영역내 화소의 수에 따라 결정되므로 약간의 잡음에 영향을 적게 받는다. 각 부분 영역을 3x3 크기를 가지는 윈도우로 사용할 때 각 부분 영역의 특징값은 0과 9 사이의 값중 하나가 되고 신경회로망의 입력으로 사용하기 위해 0과 1사이의 값으로 정규화시켰다.

2) 교차 특징

교차 특징은 숫자의 구조적 형태를 잘 나타내는 특징으로서, 숫자의 수직축과 수평축으로 검출선을 주사하였을 때 검출선과 숫자의 검은 영역과 만나는 횟수로 표현된다. 이 특징은 특징 추출 방법이 간단하여 숫자와 같이 분류할 그룹이 적을 때 많이 사용된다. 또한 이 특징은 수직 및 수평축과의 만나는 횟수만을 고려함에 따라 숫자의 이동이나 약간의 기울어짐에도 강한 특성을 나타내는 장점이 있다. 그러나 이 특징은 비슷한 형태의 숫자를 구분하기 어려운 단점을 가지고 있어 단일 특징으로 사용하기보다 보조 특징으로서 다른 특징과 함께 복합 특징으로 많이 사용된다. 이 특징은 잡음에 민감할 수 있으므로, 본 논문에서는 망 특징 추출때 얻어진 두꺼워진 영상에서 교차 특징을 추출하였다.

3) 투영 특징

투영 특징은 숫자 영상의 MBR에 대하여 수직축과 수평축을 따라 검출선을 주사하였을 때 만나는 검은 화소수를 더함으로써 특징을 추출한다. 이 특징은 교차 특징에서 나타나지 않는 획의 두께를 나타낼 수 있으나 두께가 서로 다른 영상을 입력으로 사용할 때 좋은 특성을 나타내지 못하는 단점이 있다. 따라서 이 특징

은 주로 보조 특징으로 많이 사용되고 있으며, 신경회로망의 입력으로 사용하기 위하여 투영값중 최대값으로 모든 투영값을 나눔으로써 0과 1사이의 값으로 정규화하였다.

4) 교차 거리 특징

교차거리는 영상의 MBR을 기준으로 설정하고, 이 기준점으로부터 수직 및 수평 방향으로 검출선을 주사할때 검은 화소를 처음 만날때 까지의 거리를 나타낸다. 이 특징은 숫자의 4개의 경계선인 위, 아래, 좌측 및 우측의 경계선으로부터 검출선을 주사하면 숫자의 외곽 형태 특징을 추출할 수 있으나, 외곽 형태가 비슷한 숫자를 구별하기가 어려운 단점이 있다. 따라서 본 논문에서는 그림 2와 같이 숫자의 수직축 중간 지점에 하나의 경계선을 추가하여 숫자의 내부 형태도 표현할 수 있도록 하였다. 이렇게 형성된 8개의 경계선에서 검출선을 주사하여 교차 거리 특징을 추출하였다.

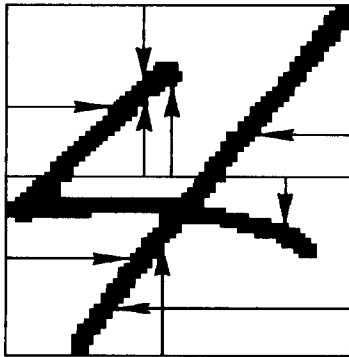


그림 2. 교차거리 특징
Fig. 2. Extraction of cross distance feature.

5) 특성 궤적 특징

이 특징은 Knoll^[10]에 의해 체계화된 특징으로서, 대부분의 특징들이 이진화된 숫자 영상의 검은 화소를 기준으로 정의한 것과는 다르게 숫자 영상의 배경 영역인 흰 화소를 기준으로 특징을 추출한다. 숫자의 모든 흰 화소로부터 동,서,남,북 방향으로 숫자 영상의 경계까지 검출선을 주사하였을때, 새로운 검은 영역과의 만나는 횟수를 특징으로 추출한다.

이 특징은 교차 특징과 마찬가지로 숫자 영상의 두께에는 무관하게 특징을 추출함에 따라 영상의 두께 및 회전에 강한 특성을 나타내지만, 잡음에는 약한 단점이 있다. 그러나 숫자 영상이 아주 두꺼운 경우, 예

를 들어 '8'과 같은 숫자는 내부의 배경 영역이 모두 채워질 가능성이 있으므로 이 특징의 추출이 어렵다. 따라서 숫자 영상의 두께에 무관한 특성을 갖는 특성 궤적 특징을 추출하기 위해 본 논문에서는 숫자 영상에 대해 Jang이 제안한^[12] 모폴로지(morphology)를 이용한 세선화(thinning) 과정을 거친후 이 특징을 추출하였다. 모폴로지에 의한 세선화 알고리즘은 구현이 쉽고 계산 속도가 비교적 빠른 장점이 있다. 본 논문에서는 두 종류의 마스크 집합들을 사용하여 세선화를 수행하였다. 두개의 마스크 집합은 영상의 컨벡스 코너(convex corner)를 제거하는 마스크의 집합인 $D = \{D1, D2, D3, D4\}$ 와 영상의 컨캐브 코너(concave corner)를 제거하는 마스크 집합인 $E = \{E1, E2, E3, E4\}$ 로 구성되어 있으며 그림 3에 나타낸 것과 같다. 여기서 D1은 북서쪽 코너, D2는 북동쪽 코너, D3는 남동쪽 코너, D4는 남서쪽 코너의 화소를 제거하는데 사용된 마스크이며, E1은 북쪽 코너, E2는 동쪽 코너, E3는 남쪽 코너 및 E4는 서쪽 코너의 화소를 제거하기 위해 각각 사용되었다. 세선화의 성능중 중심축의 정도 및 세선화 시간이 중요한 요소이므로 한번의 단계로 세선화를 수행하는 것이 효율적이다. 따라서 본 논문에서는 D 마스크와 E 마스크를 동시에 적용하여 화소사이의 연결성을 유지하면서 최소의 화소 두께를 가지는 골격을 가지는 세선화 결과를 얻었다.

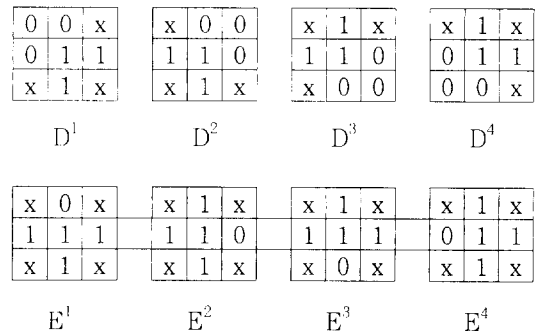


그림 3. 세선화에 사용된 마스크 집합
Fig. 3. Eight directional structuring masks.

필기체 숫자는 여러 가지 변형이 존재함에 따라, 숫자 영상의 흰 화소로부터 4 방향으로 교차 횟수를 구할때 많은 경우의 수가 발생할 수 있다. 이에따라 특징 벡터의 수가 많아지게 된다. 이 특징 벡터를 입력으로 사용한 신경회로망은 규모가 커지게 된다. 따라서 본 논문에서는 각 방향으로의 최대 교차 횟수를 3으로 제

한함에 따라 특징의 수를 81개로 고정하였다. 또한 숫자 영상의 모든 흰 화소에 대한 이 특징의 추출은 경우 시간이 오래 걸리므로 본 논문에서는 20개의 배경영역을 기준으로 특징을 추출하였다. 특성 궤적의 특징은 식(1)과 같이 표현된다. 여기서 i 는 각 4가지 방향을 나타내고 $i=1$ 은 북쪽 방향, $i=2$ 는 동쪽 방향, $i=3$ 은 남쪽 방향, $i=4$ 는 서쪽 방향일때의 가중치이며, C 는 각 방향에서 만나는 횟수를 나타낸다.

$$CL^k = \sum_{j=1}^{20} C \times 3^{(j-1)} \quad (1)$$

여기서 k 는 0에서 80까지의 정수이다. 식(1)에 의해 구해진 특징은 신경회로망의 입력으로 사용하기 위해 0과 1사이의 값으로 정규화하였다. 그림 4는 특성 궤적 특징을 추출하는 예이다.

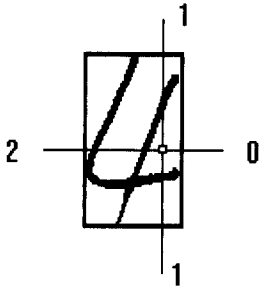


그림 4. 특성 궤적 특징의 추출 예 $(1x3^0 + 1x3^2 + 2x3^3 = 64)$

Fig. 4. Extraction of characteristic-loci feature.

3. 학습 알고리즘

인식기를 구성하는 신경회로망으로 본 논문에서는 GDR(generalized delta rule)을 이용한 역전파(backpropagation) 학습 알고리즘을 사용하였다. 이때 신경회로망의 출력층은 숫자의 종류인 10개로 구성되는데, 분류된 숫자를 가르키는 활성화 노드는 '1'로 나머지 노드는 '0'이 되도록 학습시켰다. 기존의 역전파 학습 방법은 학습 속도가 느린 단점이 있어, 이 문제점을 개선하기 위해 Fahlman^[13]이 제안한 학습 방법으로 학습을 수행하였으며, 학습 과정을 요약하면 다음과 같다.

- i) 아래층의 i 번째 뉴런과 윗층의 j 번째 뉴런사이의 연결 강도인 W_{ij} 와 각 뉴런의 임계치(threshold value) θ_j 를 랜덤하게 초기화한다.
- ii) 출력층에서 각 뉴런에 대한 실제 출력값을 계산한

다.

$$O_j = 1 / (1 + \exp(-(\sum W_{ji} O_i + \theta_j))) \quad (2)$$

- iii) 출력 에러 정정 과정에서 출력층과 중간층에서 에러값 δ_j 를 계산한다.

$$\delta_j = [\tanh^{-1}(t_j - O_j)] [O_j(1 - O_j) + 0.1] \quad (출력층) \quad (3)$$

$$\delta_i = O_i(1 - O_i) \sum_k \delta_k W_{ki} \quad (입력층) \quad (4)$$

t_j : 원하는 출력값

- iv) 연결 강도값을 조정한다.

$$\Delta W_{ji}(t) = \eta \delta_j O_i + \alpha \Delta W_{ji}(t-1) \quad (5)$$

여기서 η 는 학습률을 나타내며 α 는 관성항을 나타낸다. 수정형 역전파 학습 방법은 식(3)에서와 같이 시그모이드 함수를 미분한 항 $O_j(1 - O_j)$ 에 상수항 0.1을 더하고, $(t_j - O_j)$ 에 \tanh^{-1} 함수를 취하여 $(t_j - O_j)$ 가 클수록 그 값이 더 크게 δ_j 에 반영함으로써 학습 시간을 단축시켰다. 실제 실험에서는 \tanh^{-1} 함수가 $+\infty$ 와 $-\infty$ 가 되는 것을 방지하기 위해 $(t_j - O_j)$ 가 0.999보다 크면 +17로, -0.999보다 작으면 -17로 함수값을 고정시켰다.

4. 모듈화된 신경회로망

숫자는 크기와 모양이 다른 형태를 가지고 있고, 또한 숫자에 따라 다양한 변형이 발생할 수 있다. 따라서 하나의 특징을 입력으로 갖는 인식기로서 신뢰성있는 인식 시스템을 구성하기가 어렵다. 이를 해결하기 위한 하나의 방법은 숫자를 대표할 수 있는 여러 종류의 특징을 추출하여 복합 특징을 구성한후 신경회로망의 입력으로 사용하는 것이다. 그러나 특징의 종류 및 특징의 차원이 큰 복합 특징을 입력으로 사용한 단일 인식기는 기억 용량의 증대, 학습의 어려움 및 인식 시간이 오래 걸리는 문제점이 있다. 따라서 본 논문에서는 4개의 소규모 신경회로망을 부인식기로 구성하고 각 부인식기의 결과로부터 최종 인식 결과를 얻도록 하였다. 부인식기 1은 망특징, 교차특징 및 투영 특징을 포함한 80개의 필기체 숫자를 학습한 웨이트, 부인식기 2는 교차거리 특징 80개로서 필기체 숫자를 학습한 웨이트, 부인식기 3은 특성 궤적 특징 81개로서 필기체 숫자를 학습한 웨이트로 구성하였다. 또한 부인식기 4는 망특징, 교차특징 및 투영 특징을 포함한 80개의

특징으로 인쇄체 숫자를 학습한 웨이트로 구성되어 있다.

5. 신뢰도 지수 및 출력값 결정

신뢰도 지수는 신경회로망에 학습된 영상과 입력으로 사용된 인식 대상과의 유사도를 나타내는 값으로서 신경회로망의 출력값으로부터 구할 수 있다. 신뢰도 지수를 이용하면 인식 결과에 대해 인식 여부를 판단할 수 있다. 신뢰도 지수는 판별식 정의에 따라 인식률 및 오인식률에 영향을 미친다.¹⁹¹ 본 논문에서는 계산량이 적으면서 효율적인 판단이 가능한 Takahashi¹⁴¹가 제안한 아래의 식(6)을 이용하였다.

$$RF = M_1 + (1 - \frac{M_2}{M_1}) \quad (6)$$

여기서 M_1 은 출력층의 최고값, M_2 는 출력층값중 두 번째 큰값이다. 이식은 제대로 분류된 숫자의 출력 노드에는 '1'에 가까운 값이 나타나고, 나머지 9개의 출력 노드에는 '0'에 가까운 작은 값이 분포됨을 근거로 만든 식이다. 필기체 숫자와 같이 변형이 심한 숫자는 출력층 노드의 분포가 분산될 가능성이 있다. 따라서 이식은 출력층의 값으로부터 RF를 계산하여 임계값 이상이면 인식하고, 임계값 이하이면 인식 거절로 처리함으로써 인식 신뢰도를 향상시킬 수 있다.

모듈화된 각 신경회로망은 서로 다른 특징으로 학습되어 있음에 따라 출력값이 다르게 된다. 모듈화된 인식기의 최종 출력은 서로 다른 출력을 가지는 각 부인식기의 결과로부터 구할 수 있다. 최종 출력을 결정하는 방법으로 출력값의 최고치 선택 방법(winner-take-all)과 다수결 투표법(majority voting)등이 있다. 최고치 RF에 의한 방법은 각 부인식기의 출력 RF 중 최고의 RF를 가지는 부인식기의 출력값을 최종 결과로 선택하는 것이다. 다수결 투표법은 부인식기의 출력이 다수개인 결과를 선택하는 방법으로서, 부인식기의 출력값이 모두 다른 경우는 최고치 RF를 갖는 부인식기의 출력을 선택하는 방법이다. 본 논문에서는 최고치 RF 선택법에 의한 인식률이 다수결 투표법에 의한 인식률보다 높게 나타남에 따라 최고치 RF 선택법에 의해 최종 출력을 결정하였다.

III. 숫자 인식 실험 및 검토

1. 실험 환경

필기체 숫자의 인식은 실험에 사용하는 필기체 숫자

의 데이터 베이스에 따라 많은 영향을 미칠 수 있다. 기존에 알려져 있는 데이터 베이스로서 미국의 NIST (national institute of standards technology) 및 캐나다 Concordia 대학의 필기체 숫자 데이터 베이스 등이 있다. 그러나, 이들 데이터 베이스의 숫자 형태는 우리나라 필기자의 숫자체와는 많은 차이가 있으므로 이들 데이터 베이스로서 국내의 응용 분야에 적용하는 것은 문제점이 있다. 따라서 본 논문의 학습 및 실험에 사용된 필기체 숫자는 임의의 130명으로부터 4,000자를 수집하여, 이중 1,000자는 학습에 사용하고 나머지 3,000자로 인식 실험을 하였다. 수집한 필기체 숫자는 필기자의 평소 습관에 충실하도록 하였으며, 필기 도구 및 크기에는 제한을 두지 않았다. 인쇄체 숫자는 기존의 문서 편집기(호플 2.1, 윈도우 등)에서 크기 및 형태가 다양한 숫자 1,800자를 추출하여 이중 720자는 학습에 사용하고, 나머지 1,080자는 인식 실험을 하였다. 숫자는 스캐너를 이용하여 200dpi의 해상도를 갖는 이진 영상으로 변환하였으며, 학습 및 실험은 486DX/33MHz PC에서 수행하였다. 학습은 각 부인식기별로 역전파법을 사용한 신경회로망으로 수행하였다. 학습은 학습률과 관성항을 0.1과 0.7로 하고, 은닉층은 1개, 은닉층의 유니트 수는 30개, 학습 패턴의 전체 오차가 0.2이하가 되면 학습이 완료되는 것으로 간주하였다. 신경회로망에서 은닉층은 입력 특징에 대한 내부적인 표현으로 동작하게 되고, 또한 은닉층의 개수 및 각 은닉층에서 유니트수에 대한 정확한 설정 방법은 알려져 있지 않다. 본 논문에서의 은닉층 유니트 개수는 다양한 실험 결과로부터 정해진 것이며, 일반적으로 은닉층 유니트 개수가 너무 적으면 학습에 어려움이 있으며 은닉층이 너무 많을 경우 학습 시간이 많이 소요될 뿐만 아니라 학습된 특징에 비해 변형된 영상에 대해서 오인식될 수 있다. 또한 학습 완료의 기준으로서 전체 오차의 설정은 본 논문의 경우 전체 학습 패턴에 전체 오차가 1.0이하이면 학습된 패턴에 대해서는 100%의 인식률을 보이게 된다. 그러나 전체 오차의 값을 너무 적음으로 설정시 과도 학습에 따른 입력 패턴의 조그마한 변형에도 오인식이 발생할 수 있으므로 본 논문에서는 학습 시간 및 테스트 패턴에 대한 인식률을 고려하여 학습 완료를 결정하였다. 학습은 4개의 모든 인식기가 300회 이내의 반복 학습 횟수를 나타내었으며, 소요된 시간은 약 3시간이었다. 그림 5는 인식 실험에 사용된 다양한 크기 및 다양한

형태의 숫자 영상의 예이다.

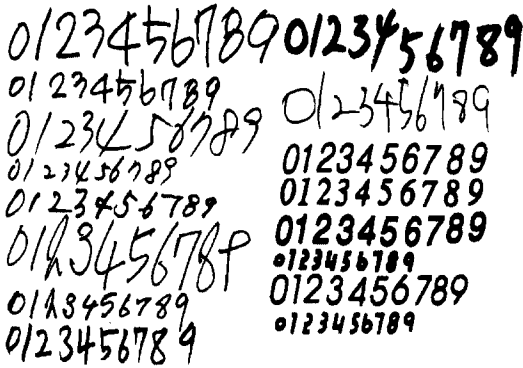


그림 5. 인식 실험에 사용된 숫자 영상의 예
Fig. 5. Example of numeral images for experiment.

2. 인식 실험 및 고찰

1) 각 부인식기의 인식 실험

인식 실험은 먼저 각 부인식기별로 수행하였다. 부인식기 1은 망특징 40개, 교차 특징 20개 및 투영 특징 20개를 복합한 80개의 특징을 입력으로 사용하였다. 부인식기 1의 인식 결과는 표 1에 나타낸것과 같다. 각 숫자별 300자씩 총 3,000자를 인식 실험시 부인식기 1의 인식률은 94.2%였으며, '2', '8' 및 '9'에서 오인식 분포가 많음을 보여준다. 특히 (1,7), (2,6), (8,5) 및 (9,7)의 관계에서 오인식된 경우가 많이 발생했다. 여기서 ()내의 앞의 숫자는 입력 숫자이며, 뒤의 숫자는 입력 숫자에 대한 오인식된 숫자를 나타낸다. 필기체 숫자는 다양한 변형 형태가 존재하므로 '1'의 망특징이 '7'과 유사해지는 경향이 있으며 교차 특징은 비슷한 형태를 가지는 숫자를 구분하기 어려운 단점이 있다. 또한, 투영 특징은 숫자의 두께에 많은 영향을 받는다. 따라서 이들 특징이 숫자의 구조적인 특성을 표현할 수 있는 한계에 의해 오인식이 발생하였다.

부인식기 2는 교차 거리 특징 80개를 입력으로 가지는 신경회로망으로서, 숫자의 형태에 따른 구조적인 특성을 표현할 수 있다. 표 2는 부인식기 2의 인식 결과를 나타내는데, 3,000자의 필기체 숫자에 대한 인식률은 93.4%였다. 인식 실험 결과로서 (0,6), (1,4), (2,3), (7,9), (8,5) 및 (9,7)의 사이에 오인식 분포가 커짐을 알 수 있다. 이것은 이들 입력/오인식 쌍사이의 숫자 형태가 비슷하기 때문이다.

표 1. 부인식기 1의 인식 결과

Table 1. Recognition result of subclassifier-1.

output input	0	1	2	3	4	5	6	7	8	9
0	293	0	0	0	0	3	2	1	1	0
1	0	291	1	0	0	3	0	5	0	0
2	0	1	266	4	1	7	14	2	5	0
3	0	0	0	288	0	5	0	0	4	3
4	0	0	7	0	288	0	0	1	2	2
5	1	0	1	9	0	281	3	3	2	0
6	0	0	0	1	0	2	296	0	1	0
7	1	0	0	0	0	0	0	291	0	8
8	0	0	1	4	0	19	1	0	275	0
9	0	0	0	4	4	1	0	33	1	257
Total recognition rate : 94.2% (2,826/3,000)										

표 2. 부인식기 2의 인식 결과

Table 2. Recognition result of subclassifier-2.

output input	0	1	2	3	4	5	6	7	8	9
0	284	0	0	0	0	1	13	2	0	0
1	0	270	1	0	24	3	0	2	0	0
2	0	2	278	13	2	0	1	1	3	0
3	0	4	1	286	0	2	2	2	3	0
4	0	1	3	0	287	3	1	1	2	2
5	0	0	0	0	0	297	3	0	0	0
6	0	0	0	0	0	2	297	0	1	0
7	1	1	0	1	0	1	0	288	0	8
8	3	4	9	6	1	19	9	0	248	1
9	1	0	1	1	0	2	0	26	0	269
Total recognition rate : 93.4% (2,804/3,000)										

표 3. 부인식기 3의 인식 결과

Table 3. Recognition result of subclassifier-3.

output input	0	1	2	3	4	5	6	7	8	9
0	285	0	0	0	0	0	0	1	13	1
1	0	287	4	2	1	4	0	1	0	1
2	0	1	288	0	3	0	0	6	2	0
3	0	0	10	281	0	2	0	2	1	4
4	0	0	0	1	295	1	0	0	0	3
5	0	1	2	3	3	275	14	0	2	0
6	0	0	0	0	0	1	294	0	5	0
7	0	1	0	3	1	0	0	286	0	9
8	3	0	0	0	0	1	3	0	293	0
9	0	1	0	5	1	4	0	3	0	286
Total recognition rate : 95.6% (2,870/3,000)										

부인식기 3은 특성 궤적 특징 81개로서 인식기 입력을 구성하고, 이 특징을 추출하기 전 숫자 영상에 세션

화 과정을 수행하였다. 특성 제적 특징이 영상의 배경 영역을 기준으로 4방향성을 고려한것이므로, 숫자의 형태를 표현할 수 있다. 표 3은 부인식기 3으로 인식 실험을 수행한 결과로서 인식률은 95.6%였다. 다른 부인식기의 결과와는 달리 숫자의 오인식 분포가 비교적 균일하지만, (0,8) 및 (5,6) 사이에는 오인식이 많이 발생하였다. 이것은 필기체 숫자의 변형이 심할 경우 비슷한 형태로 인식된 것이다.

부인식기 4는 부인식기 1에서 사용한 특징을 인쇄체 숫자에 대하여 인식 실험을 수행하기 위하여 구성된 것이다. 표 4는 크기와 두께가 다양한 인쇄체 숫자 1,080자에 대한 실험 결과로서, 99.9%의 인식률을 나타내었다. 이는 인쇄체 숫자에 대한 인식은 신경회로망을 이용한 단일 인식기로 구성 가능함을 보여준다.

표 4. 인식기 4의 인식 결과

Table 4. Recognition result of subclassifier-4.

output input	0	1	2	3	4	5	6	7	8	9
0	108	0	0	0	0	0	0	0	0	0
1	0	108	0	0	0	0	0	0	0	0
2	0	0	108	0	0	0	0	0	0	0
3	0	0	0	108	0	0	0	0	0	0
4	0	0	0	0	108	0	0	0	0	0
5	0	0	0	0	0	108	0	0	0	0
6	1	0	0	0	0	0	107	0	0	0
7	0	0	0	0	0	0	0	108	0	0
8	0	0	0	0	0	0	0	0	108	0
9	0	0	0	0	0	0	0	0	0	108
Total recognition rate : 99.9% (1,079/1,080)										

2) 모듈화 신경망에서 필기체 숫자 인식 실험

필기체 숫자를 학습한 3개의 부인식기로 필기체 숫자에 대한 인식은 최고치 RF 선택법으로 인식 실험을 수행하였으며, 인식 결과는 표 5와 같다. 인식 결과로서 인식률은 98.4%였다. 이것은 각 부인식기의 인식률보다 높은 인식률을 나타낸 것으로서 실험에 사용한 특징들 사이에는 상호 보완 관계가 있음을 나타낸다. 오인식의 분포는 (0,8), (5,6) 및 (9,7)에서 주로 나타났다. 실제 응용 분야에서 인식기의 성능이 오인식률의 최소화도 중요한 사항이다. 따라서, 오인식률의 최소화를 위한 실험에서는 인식기의 최종 출력값에 임계치를 설정하여 임계치 이상이면 인식하고 임계치 이하이면 인식 거절하는 방법을 사용하였다. 임계치에 의한 인식 실험 결과는 RF값이 1.5일때 인식률 96.2%, 인

식 거절률 3.0%, 오인식률 0.8%로 나타났으며, RF값의 변화에 따른 실험 결과는 표 6에 나타난 것과 같다.

표 5. 필기체 숫자 인식 결과

Table 5. Recognition result of handwritten numerals.

output input	0	1	2	3	4	5	6	7	8	9
0	295	0	0	0	0	0	1	0	4	0
1	0	295	2	0	2	1	0	0	0	0
2	0	1	295	2	1	0	0	0	1	0
3	0	0	2	296	0	0	0	0	2	0
4	0	0	0	0	297	1	0	0	0	2
5	0	0	0	0	0	292	6	0	0	0
6	0	0	0	0	0	0	298	0	2	0
7	0	0	0	0	0	0	0	298	0	2
8	1	1	0	1	0	2	2	0	293	0
9	0	0	0	3	0	1	0	4	0	292
Total recognition rate : 98.4% (2,951/3,000)										

표 6. RF의 변화에 따른 필기체 숫자의 인식 실험

Table 6. Recognition results of handwritten numerals with RF.

Threshold (RF)	Recognition	Rejection	Substitution
0	98.4 %	0 %	1.6 %
1.5	96.2 %	3.0 %	0.8 %
1.6	94.9 %	4.5 %	0.6 %
1.7	93.5 %	6.0 %	0.5 %
1.8	92.0 %	7.5 %	0.5 %

3) 필기체 숫자 및 인쇄체 숫자를 통합한 인식 실험

필기체와 인쇄체가 혼용된 숫자의 인식은 4개의 부인식기의 각 출력 RF를 이용하여 최고치 RF 선택법으로 수행하였다. 표 7은 인쇄체 1,080자 및 필기체 3,000자의 총 4,080자를 4개의 부인식기로써 인식한 결과를 나타낸다. 인식 결과로서 오인식의 분포는 (0,8), (3,9), (5,6) 및 (9,7)에서 주로 나타났고, 이는 필기체 숫자의 인식 결과와 비슷하다. 임계치에 의한 인식 실험 결과는 RF값이 1.5일때 인식률 97%, 인식 거절률 2%, 오인식률 1%로 나타났으며, 각 숫자별 인식 분포는 표 8에 나타난 것과 같다.

표 9는 RF값의 변화에 따른 실험 결과를 나타낸 것이다. 임계치를 크게 설정하면 인식 거절률은 증가함에 따라 인식률이 감소하지만, 오인식률에는 큰 영향을 미치지 못하였다. 또한 인쇄체와 필기체가 혼용된 숫자

인식을 다수결 투표법에 의해 수행한 결과, 최고치 RF에 의한 인식률보다 낮게 나타났다.

표 7. 필기체 및 인쇄체 혼용 숫자 인식 결과
Table 7. Recognition results of handwritten/printed numerals.

output input	0	1	2	3	4	5	6	7	8	9
0	402	0	0	0	0	0	2	0	4	0
1	0	405	1	0	2	0	0	0	0	0
2	0	0	401	2	1	0	0	3	1	0
3	0	0	1	393	0	9	0	1	4	0
4	0	0	0	0	403	0	0	0	0	5
5	0	0	0	1	0	397	9	1	0	0
6	0	0	0	0	0	0	406	0	2	0
7	0	0	0	0	0	0	0	406	0	2
8	1	1	0	1	0	4	3	0	397	1
9	0	0	0	3	0	2	0	8	0	395
Total recognition rate : 98.2%(4,005/4,080)										

표 8. 인식 거절을 적용한 필기체 및 인쇄체 혼용 숫자 인식 결과

Table 8. Recognition results of handwritten/printed numerals with rejection(RF = 1.5).

output input	0	1	2	3	4	5	6	7	8	9	rejection
0	401	0	0	0	0	0	1	0	0	0	6
1	0	396	1	0	0	0	0	0	0	0	11
2	0	0	391	0	1	0	0	1	1	0	14
3	0	0	1	391	0	9	0	0	2	0	5
4	0	0	0	0	399	0	0	0	0	3	6
5	0	0	0	1	0	395	5	0	0	0	7
6	0	0	0	0	0	0	404	0	2	0	2
7	0	0	0	0	0	0	0	404	0	0	4
8	0	0	0	0	0	3	1	0	388	0	16
9	0	0	0	3	0	1	0	5	0	389	10
Recognition: 97%, Substitution: 1%, Rejection : 2%											

표 9. RF의 변화에 따른 필기체/인쇄체 숫자의 인식 실험

Table 9. Recognition results of handwritten/printed numerals with RFs.

Threshold (RF)	Recognition	Rejection	Substitution
0	98.2%	0%	1.8%
1.5	97.0%	2.0%	1.0%
1.6	97.4%	2.8%	0.8%
1.7	95.5%	3.8%	0.7%
1.8	94.8%	4.6%	0.6%

3. 실험 결과 검토

필기체 숫자와 인쇄체 숫자는 글자의 형태가 서로

다르고, 특히 필기체 숫자는 사람에 따라 변형이 심하다. 본 실험에 따르면 두 가지 글자체 사이에 상관 관계는 크지 않은 것으로 나타났다. 즉, 필기체 숫자를 학습한 인식기로 인쇄체 숫자를 인식하는 경우 및 그 역과정에서 인식률은 약 60%였다. 그러나, 이들을 통합한 인식기는 필기체 숫자의 인식률에 약 0.5%의 영향을 미쳤으며 이때 인쇄체 숫자는 단일 인쇄체 인식기의 인식률을 나타내었다.

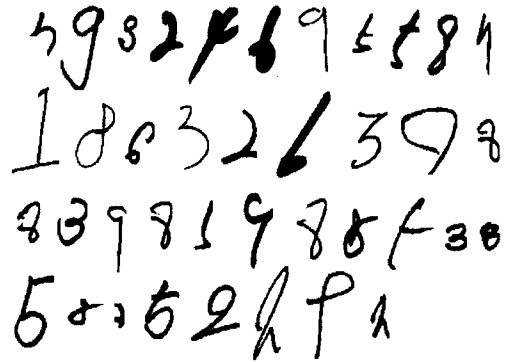


그림 6. 오인식된 숫자 영상의 예

Fig. 6. Example of substituted input images.

그림 6은 본 논문에 사용한 인식기에서 오인식된 숫자의 예이다. 오인식된 숫자는 변형이 아주 심한 경우가 대부분이었다. '1'의 경우 위쪽 가로선(serif) 또는 아래쪽 가로선이 있는 경우 '7' 또는 '2'로 오인식되었으며, 변형에 따라 '2'는 '3'으로, '3'은 '8', '4'는 '9'로 오인식되는 경우가 많았다. '5'는 완성되지 않은 변형의 경우 '9', 변형이 심한 경우 '6'으로 오인식되었으며, '6'은 '8'로 오인식되었다. '7'은 세로의 스트로크(stroke)가 긴 경우에 '9'로 오인식되었으며, '8'은 변형 형태에 따라 '5' 또는 '6'으로 오인식되는 것으로 나타났다. 또한 '9'는 세로의 스트로크가 아주 짧은 경우와 세로의 스트로크가 지나치게 원형일 경우에 '7' 또는 '3'으로 오인식되었으며, '0'은 '8'로 오인식되는 것으로 나타났다. 필기 도구에 의한 오인식은 아주 두꺼운 영상에서 '6', '8' 및 '9'와 같은 숫자의 내부 배경 영역이 모두 채워졌을 때 많이 나타났다.

IV. 결론

본 논문에서는 다수의 특징을 별개의 부인식기인 신경회로망의 입력으로 사용하여 인쇄체 및 필기체를 인

식할 수 있는 인식기를 구성하여 실험을 하였다. 인쇄체 숫자는 변형이 적은편이지만, 필기체의 경우 필자의 습관이나, 자세등의 개인적인 요인과 필기 도구, 입력 장치등에 의한 요인등으로 인해 숫자의 변형이 심하다. 인식기는 신경회로망의 특성이 결합 복구 능력과 변형에 적응력이 있으므로, 소규모의 신경회로망으로 모듈화하여 구성하였다. 인식기의 성능이 숫자의 형태를 가장 표현할 수 있는 특징의 적합성 여부에 따라 결정되지만, 실제의 경우 이상적인 특징으로 인식기를 구현하기가 어렵다. 따라서 다양한 특징을 사용하여 인식기를 구성하는 것이 높은 인식률을 얻을 수 있으므로 본 실험에서는 총 6개의 특징을 사용하였다. 많은 특징을 하나의 신경회로망의 입력으로 사용한 인식기는 시스템의 규모가 커진다. 따라서 학습의 어려움, 및 인식 시간의 증대등의 문제점이 있으므로, 사용한 특징별로 소규모의 인식기로 모듈화하여 사용하는것이 효율적이다. 본 논문에서는 필기체에 대해 3개의 부인식기, 인쇄체에 대해 하나의 부인식기로 모듈화하여 인쇄체 숫자 1,080자와 필기체 숫자 3,000자를 실험한 결과 98.2%의 인식률을 나타내었다. 이는 필기체 숫자 3,000자의 인식률과 거의 같았으며, 인쇄체와 필기체를 통합한 인식기의 구현 가능성을 보여준다. 필기체를 학습한 부인식기는 인쇄체의 인식에 영향을 미치지 않았으며, 인쇄체 인식기가 필기체 숫자의 오인식에 약 0.5%의 영향을 주었다. 각 부인식기의 출력값으로부터 신뢰도 지수를 구하여, 최고의 신뢰도 지수를 가지는 부인식기의 출력값을 인식기의 최종 출력으로 선택하므로써, 전체 인식률의 향상뿐만 아니라 오인식을 최소화하는데 상당한 효과가 있었다. 그러나, 모듈화된 신경회로망에서 신뢰성 있는 인식을 위해 각 부인식기에서 최대의 인식률을 나타내는 특징을 찾는 연구가 계속 수행되어야 하며, 인식기의 구성 방식에 관한 연구도 병행되어야 할 것이다.

참 고 문 헌

- [1] C. Y. Suen, M. Berthod, and S. Mori, "Automatic recognition of handprinted characters -The state of the art," Proc. *IEEE*, vol. 68, no.4, pp.469-487, Apr. 1980.
- [2] L. Xu, A. Krzyzak and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Systems, Man, and Cybernetics*, vol.22, no.3, pp.418-435, May/June, 1992.
- [3] H. Yan, "Design and implementation of optimized nearest neighbor classifiers for handwritten digit recognition," *ICDAR'93*, pp.10-13, 1993.
- [4] K. Tsirikolias and B. G. Mertzios, "Statistical pattern recognition using efficient two-dimensional moments with applications to character recognition," *Pattern Recognition*, vol.26, no.6, pp. 877-882, 1993.
- [5] I. S. I. Abuhaiba and P. Ahmed, "A fuzzy graph theoretic approach to recognize the totally unconstrained handwritten numerals," *Pattern Recognition*, vol.26, no.9, pp.1335-1350, 1993.
- [6] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no.10, Oct. 1993.
- [7] J. H. Paik, S. Y. Jung and Y. B. Lee, "Multiple combined recognition system for automatic processing of credit card slip applications," *ICDAR'93*, pp.520-523, 1993.
- [8] L. Heutte, J. V. Plessis, J. L. Plagnaud and Y. Lecourtier, "Handwritten numeral recognition based on multiple feature extractors," *ICDAR'93*, pp.167-170, 1993.
- [9] 김상수, 김우태, 진성일, "모듈화된 신경회로망을 이용한 필기체 숫자 인식," 제 3회 신경회로망 연구회 학술대회 논문집, 서울, pp. 219-224, 1992
- [10] A. L. Knoll, "Experiments with 'characteristic loci' for recognition of handprinted characters," *IEEE Trans. Computers*, vol.18, pp. 366-372, Apr. 1969.
- [11] M. Michael and W. C. Lin, "Experimental study of information measure and inter-intra class distance ratios on feature selection and orderings," *IEEE*

Trans. Systems Man and Cybernetics, vol. SMC-3, no.2, pp. 172-181, Mar. 1973.

- [12] B. K. Jang and R. T. Chin, "Analysis of thinning algorithms using mathematical morphology," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-12, no.6, pp.541-551, June, 1990.
- [13] S. E. Fahlman, "Faster-learning vari-

ation on backpropagation : An empirical study." *Proc. Connectionist Models Summer School*, Carnegie Mellon University, pp.38-51, 1988.

- [14] H. Takahashi, "A neural network OCR using geometrical and zonal pattern features." *ICDAR'91*, Paris, vol.2, pp.821-828, 1991.

저 자 소 개



柳 江 洙(正會員)

1982년 2월 경북대학교 전자공학과 졸업(공학사). 1984년 2월 경북대학교 대학원 전자공학과 졸업(공학석사). 1984년 3월 ~ 1992년 2월 한국전자통신연구소(ET-RI) 연구원 및 선임연구원. 1992-

년 3월 ~ 현재 경북대학교 대학원 전자공학과 박사과정 재학중 1992년 3월 ~ 현재 구미전문대학 전자통신과 조교수. 주관심분야는 패턴인식, 영상처리등



金 佑 泰(正會員)

1986년 8월 경북대학교 전자공학과 졸업(공학사). 1989년 2월 경북대학교 대학원 전자공학과 졸업(공학석사). 1993년 8월 경북대학교 대학원 전자공학과 졸업(공학박사). 1993년 3월 ~ 현재 창신

전문대학 전산정보처리과 조교수 주관심분야는 문자인식, 컴퓨터 비전등

秦 成 一(正會員) 第 32卷 B編 第 9號 參照

현재 경북대학교 전자공학과 부교수