

SVM 분류기를 이용한 필기체 숫자인식

Recognition of Handwritten Numerals using SVM Classifiers

박중조*, 김경민**

Joong-Jo Park*, Kyoung-Min Kim**

요약

최근의 인식 시스템 연구들에 의하면 SVM 분류기가 여러 다른 분류기에 비해 우수한 인식 성능을 나타내고 있다. 이에 본 논문에서는 SVM 분류기를 사용하여 필기체 숫자를 인식하는 알고리즘을 제시한다. 본 기법에서는 필기체 숫자의 특징으로서 망특징과 Kirsch 연산자에 의한 방향 특징 및 오목특징을 사용하는데, 이 중에서 처음 두 특징은 숫자를 이루는 선에 대한 전경 정보를 표현하며, 마지막 특징은 숫자의 배경 정보를 표현하여 상호 보완적인 역할을 수행한다. 본질적으로 SVM은 두 클래스 분류기이므로 이를 다중 클래스 분류기로 사용하기 위해서는 여러 개의 SVM들을 결합하여 사용해야 하는데, 본 논문에서는 "일대일" 방법과 "일대다" 방법을 사용하여 주어진 특징에 대한 인식을 수행하였다. 제시된 기법의 성능 평가를 위해 CENPARMI 필기체 숫자 데이터베이스를 사용하여 실험하였으며, 그 결과 98.45%의 인식률을 얻을 수 있었다.

Abstract

Recent researches in the recognition system have shown that SVM (Support Vector Machine) classifiers often have superior recognition rates in comparison to other classifiers. In this paper, we present the handwritten numeral recognition algorithm using SVM classifiers. The numeral features used in our algorithm are mesh features, directional features by Kirsch operators and concavity features, where first two features represent the foreground information of numerals and the last feature represents the background information of numerals. These features are complements each of the other. Since SVM is basically a binary classifier, it is required to construct and combine several binary SVMs to get the multi-class classifiers. We use two strategies for implementing multi-class SVM classifiers: "one against one" and "one against the rest", and examine their performances on the features used. The efficiency of our method is tested by the CENPARMI handwritten numeral database, and the recognition rate of 98.45% is achieved.

Keywords : handwritten numeral recognition, mesh features, directional features, concavity features, SVM classifier

I. 서론

필기체 숫자인식은 필기자들 고유의 필체에 따라 숫자의 패턴이 매우 다양하고 잡음이나 왜곡으로 변형되어 있기 때문에 까다로운 패턴 인식 분야이다. 필기체 숫자인식은 우편번호 인식, 문서 해독 및 생산 자동화 등 그 응용분야가 광범위하여 그 동안 이에 대한 많은 연구가 이루어져 왔으며, 앞으로도 인식률 향상을 위해 지속적인 연구가 요구된다.

필기체 숫자인식의 연구는 크게 두 가지 부분으로 나뉘

수 있는데, 이는 좋은 숫자 특징(good features)의 선정 및 개발과 효과적인 인식기(classifier)의 설계이다. 숫자의 특징이란 이종의 숫자 패턴의 차이점을 잘 표현하는 특성을 수치화한 것으로서, 이는 숫자 패턴의 국부적 또는 전역적 특성을 잘 표현하여야 한다. 이때 좋은 특징이란 높은 인식률을 달성하는 특징으로서 연구 방향에 따라 많은 종류가 보고되고 있다. 매우 효과적인 숫자특징으로서 방향특징(directional features)이 광범위하게 사용되고 있는데, 이는 숫자를 이루는 선의 국부적인 방향을 표현하는 특징으로서 비교적 높은 인식률을 나타내고 있다[1,2,3]. 또한 인식률의 개선을 위해 위의 방향특징과 더불어 구조적 특징(local structure features)[2], 곡률 특징(curvature features)[4], 오목특징(concavity features) [2,3,5] 및 프로젝션 특징(projection features) [6,7]들이 상호 결합 방식으로 사용되고 있다.

한편 필기체 숫자 인식을 위한 대표적인 인식기의 종류로는 숫자를 구성하는 화소의 통계적 특징에 근거한 인식

*경상대학교 제어계측공학과, **진남대학교 전자통신전기공학과
논문 번호 : 2007-2-7 접수 일자 : 2007. 5. 9

심사 완료 : 2007. 7. 11

*본 연구 수행시에 SVM 라이브러리 "HeroSVM"의 사용을 허락해 주신 캐나다 Concordia 대학의 Suen 교수께 감사함을 표합니다.

기(feature vector based classifier), 숫자의 형태 특성에 근거한 구조적인 인식기(syntactic and structural based classifier) 및 신경회로망(neural network based classifier)을 이용한 인식기 등이 있고[1][10-13], 최근에는 통계이론에 기반한 SVM(Support Vector Machine) 분류기가 많이 사용되고 있다[3,6,7,8,9,17].

본 연구에서는 3종의 숫자특징을 선정하여 복합적으로 사용하고, 인식기로서는 SVM 분류기를 사용하여 필기체 숫자를 인식하는 방안을 제시한다. 본 연구에서 숫자특징으로서 숫자의 전체형상을 특징으로 사용하는 전역적 특징인 망특징(mesh features)과 Kirsch 에지검출기에 의해 숫자선의 4-방향 선분을 분리 검출하여 특징으로 사용하는 국부적 방향특징(directional features) 및 숫자형태에서 오목한 부분을 검출하여 특징으로 사용하는 오목특징(concavity features)을 복합적으로 사용하였다. 이 중에서 망특징과 방향특징은 숫자의 전경특징(foreground features)에 해당하고 오목특징은 숫자의 배경특징(background features)에 해당하는 것으로서 상호 보완적인 특성을 가지게 된다. 인식기로서는 SVM 분류기를 사용하는데, 본질적으로 SVM은 2-class 분류기이므로 숫자인식과 같은 10-class 문제에 적용하기 위해서는 “일대다”(one against the rest) 및 “일대일”(one against one) 기법을 사용할 수 있다. 본 연구에서는 이 두 가지 방법을 모두 사용하여 인식 작업을 시행하였다. 제시된 기법의 성능 테스트를 위해 캐나다 Concordia 대학의 필기체숫자 데이터베이스를 사용하여 실험하였으며 실험 결과 98.45%의 인식률을 얻을 수 있었다.

본 논문의 구성은 I장의 서론에 이어, II장에서는 본 연구에서 선정하여 사용한 필기체 숫자의 특징추출기법에 대해서 설명하고, III장에서는 인식기로 사용한 SVM 분류기 대해 기술한다. IV장에서는 실험을 통해 제시된 기법의 인식 성능을 고찰하고, 마지막으로 V장에서 결론을 맺는다

II. 특징 추출

필기체 숫자인식에 있어서, 좋은 특징이란 서로 다른 종류의 숫자들 간에는 구별이 잘 되고 동일한 특성을 가지지 않으며 같은 종류의 숫자들 간에는 다소의 변형에도 둔감하여 같은 특성을 나타내어야 한다. 본 장에서는 필기체 숫자인식에 사용된 3종의 숫자특징인 망특징, 방향특징 및 오목특징의 추출기법을 기술한다.

첫 번째 특징은 망특징으로서, 이는 숫자영상의 크기를 일정한 크기로 정규화하여 숫자의 전체 형상을 특징으로 사용하는 것이다. 본 연구에서는 먼저 주어진 숫자영상을 2차보간을 이용한 축소 및 확대방법을 사용하여 크기를 [24x24]크기로 정규화하고 이를 화소값 '0', '1'로만 이루어진 이치영상으로 변환한 후, 이 [24x24]크기의 정규화된 이치영상을 [3x3]크기의 부영역들로 나누어 하나의 부영역에 하나의 특징값을 할당하는 조닝(zoning) 방법을 사용하여 [8x8]크기의 망특징을 생성한다. 이때 [3x3]크기의 부영역

에 하나의 특징값을 할당하기 위해 먼저 각각의 [3x3] 부영역내의 화소값을 합산하고 이 합산된 값을 다시 [0,1]사이의 값으로 정규화하여 특징값으로 사용한다. 이 망특징은 숫자의 전체형태를 내포하고 있으므로 전역특징으로 볼 수 있으며, 이 방식으로 구해진 망특징은 크기가 64 (8x8)인 특징벡터가 된다.

두 번째 특징은 방향특징으로서 이는 숫자영상내의 숫자를 이루는 선에서 각 방향의 선분을 분리 검출하여 방향특징영상을 만들어 이를 숫자특징으로 사용하는 것으로서, 4-방향 에지검출에 효과적인 kirsch 연산자를 사용하여 구한다. 여기서 4-방향이란 horizontal(H), vertical(V), right-diagonal(R), left-diagonal(L)을 의미한다. 이 방향특징을 구하는 과정은 먼저 주어진 숫자영상을 2차보간을 이용한 축소 및 확대방법을 사용하여 크기를 [32x32]크기로 정규화한 후 이치영상으로 만들고, 이에 kirsch연산을 적용하여 4개 방향의 에지영상들을 생성한 후 이를 각각 화소값 '0', '1'로만 이루어진 이치영상으로 변환하여 4종의 방향영상을 생성한다. 그림 1은 하나의 숫자영상에 대해 이러한 방식으로 구해진 4종의 방향영상들을 보인다. 이런 방식으로 구해진 [32x32]크기인 4개의 이치 방향영상을 조닝 방법에 의해 [8x8]크기의 부영역들로 나누어 하나의 부영역에 하나의 특징값을 할당함으로써 [4x4]크기의 방향특징을 생성한다. 이때 [8x8]크기의 부영역에 하나의 특징값을 할당하는 방법으로는 먼저 각각의 [8x8]부영역내의 화소값을 합산하고 이 합산된 값을 다시 [0,1]사이의 값으로 정규화하여 특징값으로 사용한다. 이 절차는 다음의 수식으로 간략히 표현될 수 있다.

$$F_i^k = \begin{cases} 1 & , \text{ if } s_i^k \geq T \\ s_i^k / T & , \text{ otherwise} \end{cases} \quad (1)$$

$$s_i^k = \sum_{y_i} \sum_{x_i} g^k(x, y)$$

여기서, F_i^k 는 k번째 방향영상의 i 번째 [8x8]부영역으로부터 얻어지는 특징값을 나타내며, s_i^k 는 k번째 방향영상의 i 번째 [8x8]부영역내에 있는 '1'값의 개수를 나타낸다. 그리고 $g^k(x, y)$ 는 k 번째 방향영상에서 위치 (x,y)인 화소의 화소값을 나타내며, T는 정규화를 위한 임계값으로서 이는 여러차례의 예비실험을 거쳐 조심스럽게 선정되어야 한다.

이 방식으로 구해진 방향특징은 크기가 64 (4x4, 4-방향)인 특징벡터가 된다.

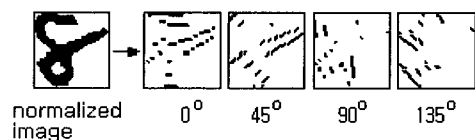


그림 1. 4-방향 영상

Fig. 1. 4-directional images

세 번째 특징인 오목특징은 숫자영상의 배경정보를 제공

하는 것으로서, 숫자형상의 볼록외피(convex hull)로부터 추출된다[3,8,9]. 이 특징을 추출하기 위해 먼저 주어진 숫자영상을 2차보간을 이용한 축소 및 확대방법을 사용하여 크기를 [40x40]크기로 정규화한 후 이치영상으로 만들고, 이로부터 숫자형상의 볼록외피를 구한다. 볼록외피영상에 원래의 숫자형상을 겹쳐 엮은 영상을 수정된 볼록외피영상이라 하면, 이 영상으로부터 5종의 오목특징영상(left-opening, right-opening, top-opening, bottom-opening 및 closing 영상)이 구해진다. 여기서, left-opening 영상은 수정된 볼록외피영상에서 좌측에서 볼 때 숫자형상에 의해 막히지 않고 보이는 볼록외피 화소들만으로 구성되는 영상이고, 마찬가지로 방법으로 right-opening, top-opening, bottom-opening 영상은 각각 수정된 볼록외피영상에서 우측, 상측, 하측에서 볼 때 숫자형상에 의해 막히지 않고 보이는 볼록외피 화소들만으로 구성되는 영상이다. 또한 closing 영상은 수정된 볼록외피영상에서 위의 4개의 오목특징영상에 속하지 않는 볼록외피 화소들만으로 구성되는 영상이다. 이때, 5종의 오목특징영상의 화소값으로는 볼록외피화소에 대응하는 화소는 '1'값을 가지며, 그 외의 화소는 모두 '0'값을 갖는다. 그림 2는 하나의 숫자영상에 대해 이러한 방식으로 구해진 5종의 오목특징영상들을 보이는데, 이 그림에서 검은 화소 부분만이 특징추출에 사용된다. 이런 방식으로 구해진 [40x40]크기의 이치 오목특징영상을 조닝 방법에 의해 [8x8]크기의 부영역들로 나누어 하나의 부영역에 하나의 특징값을 할당함으로써 [5x5]크기의 오목특징을 생성한다. 이때 [8x8]크기의 부영역에 하나의 특징값을 할당하는 방법으로는 먼저 각각의 [8x8]부영역내의 화소값을 합산하고 이 합산된 값을 다시 [0,1]사이의 값으로 정규화하여 특징값으로 사용한다. 이를 위해 역시 식 (1)의 방식이 사용될 수 있다. 이러한 방법으로 구해진 오목특징은 크기가 125 (5x5, 5-종)인 특징벡터가 된다.

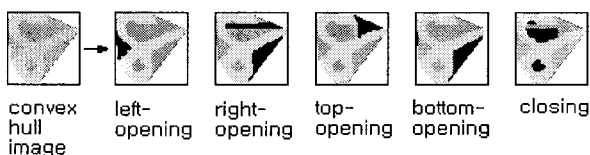


그림 2. 5-오목특징영상
Fig. 2. 5-concavity feature images

상기의 각 특징들은 각기 다음과 같은 특성을 가진다. 먼저 첫 번째 특징인 망특징은 숫자의 전역적인 형태 정보를 제공하나, 국부적인 특성의 표현이 취약하다. 그리고 두 번째 특징인 방향특징은 숫자를 이루는 각 방향선분의 정보를 표현하여 숫자의 국부적인 특성을 제공한다. 마지막으로 세 번째 오목특징은 숫자를 이루는 선이 아닌 숫자의 배경정보를 효과적으로 제공한다. 이들 각각의 특징들을 상호 보완적인 특성을 가지므로 이들을 결합하여 사용하면 인식률의 향상에 기여할 수 있을 것이다. 본 연구에서는 인식률의 향상을 위해, 상기의 3가지 특징들을 결합하여 사용

하였다. 이때 각 특징을 결합하여 복합특징을 제작하는 방식으로는, 미리 각 특징에 대해 실험하여 인식률을 구해보고 각 특징이 가져 오는 인식률에 따라 특징 종류별로 하중(weight)을 주어 결합하는 방식을 생각할 수 있다. 예비 실험에서 특징 종류별로 하중을 주는 방식을 사용하여 인식률을 구해 보았으나 오히려 인식률이 저하되는 현상이 나타나, 본 연구에서는 하중을 사용하지 않고 단지 모든 특징값을 대등하게 취급하여 모음으로써 복합특징을 구성하여 사용하였다.

III. SVM 분류기

SVM은 통계이론에 기반한 강력한 학습기법이다[14,15]. 본 인식기의 우수한 학습기능과 일반화성능에 힘입어 최근 수년간 SVM은 패턴인식분야에 있어서 중요한 역할을 수행하고 있다. 본 장에서는 SVM의 원리를 간략히 고찰하고 그 이용 방안을 기술한다[9,16,17].

SVM은 기본적으로 두 개의 클래스를 분류하는 분류기로서, 특징공간에서 학습 오차를 어느정도 허용하면서 두 클래스사이의 마진을 최대화하는 결정경계를 찾는다. 최적 초평면이라고도 하는 이 결정 경계는 SV (Support vectors)라고 불리는 학습용 샘플들중의 일부분 샘플들에 의해 정의된다.

x_i 를 학습용 벡터라 하고 y_i 를 양수 및 음수인 클래스 라벨이라고 할 때, 모든 i 에 대해 (x_i, y_i) , $x_i \in X$, $y_i \in \{+1, -1\}$ 로 주어지는 분리가능하며 선형인 두 클래스 문제에서, SVM은 분리용 초평면인 결정경계 $\langle w, x \rangle + b = 0$ 에서 입력공간을 두 개의 클래스로 나눈다. 이때 만일 이 초평면에 의해 두 클래스사이에 최대의 마진이 나타나게 되면, 이 초평면을 최적 초평면이라고 하고 이것은 식 (2)로 주어지는 결정함수에 대응한다.

$$f(x) = \langle w, x \rangle + b \quad (2)$$

여기서, $\langle \cdot \rangle$ 는 두 벡터의 내적을 나타내며, 이 결정함수의 부호에 의해 입력벡터 x 가 어느 클래스에 속하는지가 결정된다.

이때 최대의 마진은 다음의 방법으로 구해 질 수 있다. 만일 초평면에 가장 가까운 샘플점이

$|\langle w, x_i \rangle + b| = 1$ 를 만족하도록 두 파라미터 w 와 b 가 조정된다면 다음의 두 구속조건이 유효하다.

$$\langle w, x_i \rangle + b \geq +1, \text{ for } y_i = +1 \quad (3.1)$$

$$\langle w, x_i \rangle + b \leq -1, \text{ for } y_i = -1 \quad (3.2)$$

이들 두 부등식은 다음과 같이 결합될 수 있다.

$$y_i \cdot (\langle w, x_i \rangle + b) \geq 1, \text{ for all } i \quad (4)$$

이 경우, 마진은 $2/\|w\|$ 이 된다. 여기서 $\|\cdot\|$ 는 유클리드 노름(norm)을 나타낸다.

따라서, 두 클래스사이에서 최대의 마진을 갖는 최적 초평면은 다음의 2차 최적화 문제를 풀면 얻어질 수 있다.

$$\min \frac{1}{2} \|w\|^2 \quad (5)$$

$$\text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1, \text{ for all } i$$

Lagrange 승수 a_i 를 도입하면 이 최적화 문제는 다음의 등가 문제로 변환될 수 있다.

$$\max \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_i a_j y_i y_j \langle x_i, x_j \rangle \quad (6)$$

$$\text{subject to } \sum_{i=1}^N a_i y_i = 0, \quad a_i \geq 0$$

이 등가 최적화 문제를 풀면 Lagrange 승수 a_i 가 결정되며, 이때 $a_i > 0$ 에 대응하는 샘플데이터 x_i 를 SV (Support vector)라고 한다.

선형 분리를 위한 위의 최적화 문제는 커널 함수 $k(x, x')$ 를 도입함으로써 비선형 분리를 위한 최적화 문제로 확장될 수 있다. 커널 함수를 사용함으로써 식 (6)으로 주어지는 최적화 문제는 다음과 같이 변형된다.

$$\max \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_i a_j y_i y_j k(x_i, x_j) \quad (7)$$

$$\text{subject to } \sum_{i=1}^N a_i y_i = 0, \quad a_i \geq 0$$

이때 사용될 수 있는 커널 함수로는 Gaussian 커널 함수, polynomial 커널 함수 및 linear 커널 함수등이 있다. 이들 중에 본 연구에서는 분류성능이 가장 뛰어난 것으로 알려진 RBF (Gaussian radial basis function)을 사용한다. RBF는 다음의 식으로 주어진다.

$$k(x, x') = \exp(-\|x - x'\|^2 / 2\sigma^2), \quad \sigma > 0 \quad (8)$$

여기서, σ 는 본 함수의 radius 파라미터이다.

위의 원리에 더하여 특징공간에서 학습오차를 어느 정도 허용하는 SVM을 설계하기 위해, slack variable ξ_i 가 도입되었다[15]. Slack variable은 식 (5)의 최적화 문제에 서 구속조건을 다음과 같이 완화하는데 사용된다.

$$y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \text{ for all } i \quad (9)$$

그리고 목적함수에 slack variable들의 합으로 주어지는 추가의 비용을 할당함으로써, 식 (5)의 최적화 문제는 다음과 같이 변형된다.

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (10)$$

$$\text{subject to } y_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

여기서, 파라미터 C는 학습오차를 조정하는 상수로서 설계자가 임의로 선정하는데, C값이 클수록 학습과정에서 더 많은 오차가 허용된다.

Lagrange 승수 a_i 와 커널 함수를 사용하면 식 (10)은 다음의 등가 최적화문제로 변환된다.

$$\max \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_i a_j y_i y_j k(x_i, x_j) \quad (11)$$

$$\text{subject to } \sum_{i=1}^N a_i y_i = 0, \quad 0 \leq a_i \leq C$$

이제 주어진 학습용 샘플데이터를 사용하여 식 (11)의 최적화문제를 풀어 일단 Lagrange 승수 a_i 가 구해지면, 미지 데이터 x 에 대한 분류작업은 다음의 결정함수 $f(x)$ 에 의해 수행된다.

$$f(x) = \sum_{i=1}^N a_i y_i k(x_i, x) - b \quad (12)$$

여기서, x_i 와 y_i 는 학습용 샘플 데이터이다. 이 결정함수 $f(x)$ 의 부호와 크기가 미지 데이터 x 의 클래스 결정에 사용된다.

원리상 SVM은 2-class 클래스 분류기이므로 이를 multi-class 분류에 사용하기 위해서는 전략이 필요하며, 일반적으로 “일대다” (one against the rest) 방법과 “일대일” (one against one) 방법이 고려될 수 있다. 먼저, “일대다” 방법에서는, 클래스의 개수가 k 일 경우 k 개의 SVM 분류기들로 구성된다. 이 경우에는, 각각의 클래스에 대해, 양수와 음수의 클래스 라벨을 갖는 두 클래스 분류기들로 이루어지는데, 이때 목표 클래스에 대응하는 샘플들은 양의 라벨을 가지며, 그 외의 모든 클래스에 대응하는 샘플들은 음의 라벨을 가진다. 결국 이 방법에서는 k 개의 최적화 문제를 풀어야 하며, k 개의 결정함수를 사용하게 된다. 미지 데이터 x 가 주어지면 이를 k 개의 결정함수 각각에 적용하고 가장 큰 양수값을 출력하는 결정함수에 대응하는 클래스가 이 미지 데이터의 클래스가 된다. 한편, “일대일” 방법에서는, 클래스의 개수가 k 일 경우 $k(k-1)/2$ 개의 SVM 분류기들로 구성된다. 이 경우에는, 각각의 클래스 쌍에 대해, 양수와 음수의 클래스 라벨을 갖는 두 클래스 분류기들로 이루어지는데, 이때 목표 클래스에 대응하는 샘플들은 양의 라벨을 가지며, 또 다른 클래스에 대응하는 샘플들은 음의 라벨을 가진다. 결국 이 방법에서는 $k(k-1)/2$ 개의 최적화 문제를 풀어야 하며, $k(k-1)/2$ 개의 결정함수를 사용하게 된다. 미지 데이터 x 가 주어지면 이를 $k(k-1)/2$ 개의 결정함수 각각에 적용하고 “Max Wins” 같은 투표방법을 사용하여 이 미지 데이터의 클래스를 결정한다.

IV. 실험 및 고찰

본 연구에서는 필기체 숫자 데이터베이스로서 잘 알려져 있는 캐나다 Concordia대학의 CENPARMI 데이터베이스를 사용하여 실험하였다. 이 숫자 데이터베이스는 USPS의 봉투영상들로부터 취득된 6000개의 다양한 크기인 숫자영상들로 구성되는데, 이중 4000개(각 숫자당 400개씩)의 숫자영상은 학습용으로 지정되어 있고, 2000개(각 숫자당 200개씩)의 숫자영상은 테스트용으로 지정되어 있다. 그림 3은 CENPARMI 데이터베이스에 있는 필기체숫자들의 일

부를 보인다.

영상 전처리와 특징추출 및 인식알고리즘은 C 언어로 구현하였으며, Visual C++ 6.0 컴파일러를 사용하여 Windows XP에서 수행되었다.

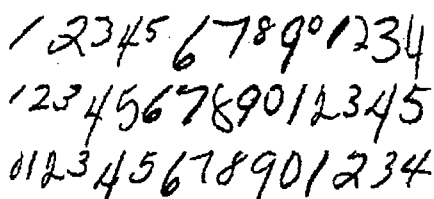


그림 3. 실험에 사용한 필기체 숫자들의 예
Fig. 3. Some examples of Numeral Data

필기체 숫자인식을 위해, 3종의 특징(망특징, 방향특징, 오목특징)이 사용되었으며, CENPARMI 데이터베이스내의 숫자영상들은 그 크기가 다양하므로 먼저 입력영상에 크기 정규화를 수행하였다. 이때 정규화 크기는 망특징, 방향특징 및 오목특징 각각에 대해 [24x24], [32x32], [40x40]로서, 이들 크기는 해당 특징이 효과적으로 검출되도록 여러 예비실험을 거쳐 결정된 것이며, 2차보간을 이용한 영상의 축소, 확대기법을 사용하여 크기 정규화를 수행하였다. 표 1은 예비실험에서 얻어진 각종 정규화 크기에 대한 인식률의 관계를 보인다. 본 연구의 알고리즘 구현상 정규화 크기를 최종 특징 크기의 정수배가 되도록 정하였다. 이 예비실험에서는 단독 특징에 대해 SVM 분류기를 "일대일" 방식으로 구현하여 분류처리를 수행하였으며, 각각의 특징에 대해 가장 높은 인식률을 나타내는 정규화 크기를 채택하였다.

표 1. 입력영상의 정규화크기와 인식률의 관계
Table 1. Relationship between normalization sizes and recognition rates

특징 종류	정규화 크기	인식률 (%)
망특징	[16x16]	95.75
	[24x24]	96.65
	[32x32]	96.50
방향특징	[24x24]	96.90
	[32x32]	97.90
	[40x40]	97.90
오목특징	[20x20]	96.35
	[40x40]	97.25
	[60x60]	96.60

이 정규화된 영상으로부터 2장에서 기술된 특징추출 방법을 사용하여 최종적으로 구해진 망특징, 방향특징 및 오목특징의 크기와 개수는 각각 [8x8]x1종, [4x4]x4방향, [5x5]x5종이며, 특징값을 구간 [0,1]사이의 값으로 정규화하는데 사용된 임계값은 각각 12, 16, 54로서 이 임계값들 역시 여러차례의 예비실험을 거쳐 결정하였다.

인식기로는 SVM 분류기를 사용하였으며, SVM 분류기의 입력으로는 특징값을 스케일링하여 사용하였다. 본 연구

에서 사용한 특징값의 스케일링의 방법으로는 특징공간상에서 상호간에 가장 먼 특징값들 사이의 유클리드거리가 1이 되도록 모든 특징값을 조정하는 방법을 사용하였다.

필기체 숫자인식을 위해서는 10-class 분류기가 필요하다. 그런데 SVM 분류기는 본질적으로 2-class 분류기이므로 이를 10-class 분류기로 사용하기 위해서는 전략이 필요한데, 일반적으로 사용되는 방법으로서 "일대다" (one against the rest) 방법과 "일대일" (one against one) 방법이 있다. 본 연구에서는 이 두 가지 방법을 모두 사용하여 실험하였으며, 각각의 방법에서 최종 판정을 위해서는 "Max Wins" 방법을 사용하였다. CENPARMI 숫자 데이터베이스중에 4000개의 학습용 데이터를 사용하여 SVM 분류기를 훈련시켰으며, 학습된 SVM 분류기에 2000개의 테스트 데이터를 적용하여 필기체 숫자의 인식률을 측정하였다. SVM에 사용된 커널로는 최상의 정확도를 제공하는 것으로 알려진 RBF 커널을 사용하였다[3][16]. 또한 SVM 파라미터인 C 및 σ^2 값으로는 10과 0.1을 사용하였는데, 이 두 파라미터 값은 최상의 인식률이 얻어지도록 예비실험을 통하여 결정된 것이다.

본 실험에서는 개개의 단일특징과 다수의 특징들을 여러 경우로 결합한 복합특징에 대해 인식률을 구해보았다. 표 2는 SVM 분류기를 "일대일" 방법으로 구현한 인식기에 의한 인식률을 보이며, 표 3은 SVM 분류기를 "일대다" 방법으로 구현한 인식기에 의한 인식률을 보인다. 이들 표에서 보면 단일 특징을 사용한 경우보다 여러 특징을 복합적으로 사용한 경우가 대체적으로 더 높은 인식률을 나타낼 수 있다. 또한 전반적으로는 SVM 분류기를 "일대일"로 사용하는 경우가 "일대다"로 사용하는 경우보다 더 높은 인식률을 나타내고 있다. 표 2의 "일대일"로 구성된 SVM 분류기를 사용한 결과에서 특이한 점은, 망특징과 방향특징으로 이루어진 복합특징의 결과와 3개의 특징을 모두 사용한 복합특징의 결과가 동일한 인식률을 나타내는데, 이는 "일대일"로 구성된 SVM 분류기에서는 복합특징의 견지에서 오목특징이 분류작업에 큰 효과를 발휘하지 못했음을 나타낸다. 또한, 표 3의 "일대다"로 구성된 SVM 분류기를 사용한 결과에서 특이한 점은, 방향특징과 오목특징으로 이루어진 복합특징의 결과가 3개의 특징을 모두 사용한 복합특징의 결과보다 더 높은 인식률을 나타내는데, 이는 "일대다"로 구성된 SVM 분류기에서는 복합특징의 견지에서 망특징이 분류작업에 큰 효과를 발휘하지 못했음을 나타낸다.

한편, 표 4는 파라미터 C가 SVM 분류기의 분류성능에 미치는 영향을 실험한 결과이다. 이 실험에서는 최대의 인식률을 보인 방향특징 및 오목특징과 "일대다" 방법의 SVM 분류기를 사용하였다. 원리상 파라미터 C값이 클수록 학습과정에서 더 많은 오차가 허용되는데, 이 결과에서 보면 작은 C값의 경우 인식률의 저하를 초래하고 있다. 이는 작은 C값에 의해 학습시에 분류기의 일반화 성능이 저하되기 때문이라고 생각된다.

표 2. SVM(일대일)에 의한 인식률

Table 1. Recognition rates by SVM (one against one)

특징 개수	사용된 특징	인식률 (%)
단일 특징	(1) 망특징	96.65
	(2) 방향특징	97.90
	(3) 오목특징	97.25
2개의 특징	(1) + (2)	98.40
	(1) + (3)	97.65
	(2) + (3)	98.30
3개의 특징	(1) + (2) + (3)	98.40

표 3. SVM(일대다)에 의한 인식률

Table 1. Recognition rates by SVM (one against the rest)

특징 개수	사용된 특징	인식률 (%)
단일 특징	(1) 망특징	92.85
	(2) 방향특징	96.10
	(3) 오목특징	95.70
2개의 특징	(1) + (2)	97.75
	(1) + (3)	96.75
	(2) + (3)	98.45
3개의 특징	(1) + (2) + (3)	98.05

표 4. SVM 파라미터 C의 인식률에 대한 영향

Table 4. Effects of a parameter C on recognition rates

C 값	3	5	10	15	20
인식률 (%)	98.20	98.40	98.45	98.45	98.45

일반적으로 다중 클래스 분류기의 구현 방안으로서 SVM 분류기를 “일대일”로 사용하는 경우가 “일대다”로 사용하는 경우보다 더 많은 계산량을 요구하는 대신 더 높은 인식률을 나타내는 것으로 알려져 있다. 그러나, 본 실험의 결과, 사용되는 특징에 따라 분류기의 성능이 달라짐을 알 수 있었다. 본 실험에서는 2개의 특징(방향특징과 오목특징)을 사용하고 SVM 분류기를 “일대다”로 사용한 경우가 가장 높은 인식률인 98.45%를 성취함을 나타내었는데, 이는 특징공간상에서 방향특징과 오목특징으로 이루어진 복합특징의 분포에 대해 “일대다”로 구성된 SVM 분류기가 최적으로 분류성능을 발휘한 것으로 생각된다.

주어진 특징에 대해 다층퍼셉트론 신경회로망으로 분류를 수행하여 SVM 분류기와의 분류성능을 비교해 보았다. 본 신경회로망은 오류역전과 학습 알고리즘을 사용하며, 입력층의 노드수는 입력 특징에 의해 정해지고, 출력층의 노드수는 클래스 개수로서 10이며, 은닉층은 1개로서 40개의 노드로 구성되었다. 표 5는 신경회로망 분류기에 의한 인식률을 보인다. 이를 SVM 분류기의 결과와 비교해 보면 SVM 분류기가 신경회로망 분류기에 비해 뛰어난 분류성능을 보이고 있다. 두 분류기의 결과를 볼 때 방향특징과 오목특징의 조합이 효과적인 복합특징으로 작용하는 것

로 생각된다. 한편, 사용된 SVM 분류기와 신경회로망 분류기의 수행 속도를 비교해 보면, SVM 분류기는 신경회로망 분류기에 비해 학습시간이 현저히 적게 소요되었으나 분류시간은 다소 더 길게 소요되었다. 분류작업시에 신경회로망 분류기는 주어진 입력에 대해 한번의 처리로 분류가 완료되지만, SVM 분류기는 각각의 클래스에 대해 별도로 분류를 수행해야 하고 그 후 얻어진 결과값들 중에서 최대값을 찾아야 하므로 분류시간이 더 길어진다고 생각된다.

표 5. 신경회로망 분류기에 의한 인식률

Table 5. Recognition rates by Neural net classifier

사용된 특징	인식률 (%)
방향특징 + 오목특징	97.80
망특징 + 방향특징 + 오목특징	96.95

V. 결론

본 논문에서는 3종이 숫자특징과 SVM 분류기를 사용한 필기체 숫자인식 알고리즘을 제시하였다. 사용된 특징은 망특징, 방향특징 및 오목특징으로서, 이중에 망특징과 방향특징은 숫자영상의 전경특징으로서 각각 숫자를 이루는 선의 전역적 특성과 국부적 특성을 표현하며, 오목특징은 배경특징으로서 전경특징에 대해 보완적인 특성을 제공한다. 사용된 인식기는 RBF 커널을 이용한 SVM 분류기로서 이는 본질적으로 2-class 분류기이므로 10-class의 분류를 위해 SVM 분류기를 “일대일” 및 “일대다” 방식으로 사용하여 인식기를 구현하였다. CENPARMI 필기체 숫자 데이터베이스를 사용하여 제시된 알고리즘을 실험해 본 결과 다수의 특징들이 상호 보완작용을 하여 인식률 향상에 기여함을 알 수 있었으며, 최고 인식률로서 “일대다” 방식의 SVM 분류기에 의해 98.45%의 인식률을 얻을 수 있었다.

향후 연구과제로는 기존의 숫자특징에 보완적으로 작용할 수 있는 새로운 특징추출 기법을 고안하고, 보다 효율적인 인식기의 개발을 위해 특징값들의 클러스터링과 SVM 분류기를 접목하는 방안을 연구하고자 한다.

참고 문헌

- [1] S. H. Lee, "Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.18, No. 6, pp.648-652, 1996.
- [2] J. T. Favata and G. Srikanth, "A multiple feature/resolution approach to handprinted digit and character recognition," *International Journal of Imaging Systems and Technology*, Vol.7, pp.304-311, 1996.
- [3] C. L. Liu, K. Nakashima, H. Sako and H. Fujisawa,

"Handwritten digit recognition: benchmarking of state-of-the-art techniques," *Pattern Recognition*, Vol.36, pp.2271-2285, 2003.

[4] M. Shi, Y. Fujisawa, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition using gradient and curvature of gray scale image," *Pattern Recognition*, Vol.35, pp.2051-2059, 2002.

[5] C. L. Liu, K. Nakashima, H. Sako and H. Fujisawa, "Handwritten digit recognition: investigation of normalization and feature extraction techniques," *Pattern Recognition*, Vol.37, pp.265-279, 2004.

[6] D. Gorgevik and D. Cakmakov, "An efficient three-stage classifier for handwritten digit recognition," *Proceedings of 17th Int. Conference on Pattern Recognition, ICPR2004*, Vol. 4, pp. 507-510, IEEE Computer Society, Cambridge, UK, 23-26 August 2004.

[7] D. Gorgevik, D. Cakmakov, "Handwritten digit recognition by combining SVM classifiers," *Proceedings of EUROCON 2005*, pp.1393-1396, Serbia & Montenegro, Belgrade, November 22-24, 2005.

[8] Chun Lei, *Error analysis of a hybrid multiple classifier system for recognizing unconstrained handwritten numerals*, A Master Thesis in the department of Computer Science & Software Engineering, Concordia University, May. 2005.

[9] Sapargali Kambar, *Generating synthetic data by morphing transformation for handwritten numeral recognition (with v-SVM)*, A Master Thesis in the department of Computer Science & Software Engineering, Concordia University, Aug. 2005.

[10] L. Lam and C. Y. Suen, "Structural classification and relaxation matching of totally unconstrained handwritten zip code numbers," *Pattern Recognition*, Vol. 21, No. 1, pp. 19-31, 1988.

[11] G. L. Martin and J. A. Pittman, "Recognizing hand-printed letters and digits using backpropagation learning," *Neural Computation*, Vol. 3, pp.258-267, 1991.

[12] K. M. Mohiuddin and J. Mao, "A comparative study of different classifiers for handprinted character recognition," *Pattern Recognition in practice IV*, pp. 437-448, 1994.

[13] S. Knerr and L. Personnaz and G. Dreyfus, "Handwritten digit recognition by neural networks with single-layer training," *IEEE Trans. Neural Networks*, Vol. 3, No. 6, pp. 962-968, 1992.

[14] B. Boser, I. Guyon, V. Vapnik, "A training algorithm for optimal margin classifiers," *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144-152, 1992.

[15] C. Cortes, V. Vapnik, "Support vector networks," *Machine learning*, v.20, pp.273-297, 1995.

[16] H. Guo and H. Zhang, "Chinese classifier assignment using SVMs," *Fourth SIGHAN Workshop on Chinese Language Processing*, pp.25-31, 2005.

[17] J. X. Dong, C. Y. Suen and A. Krzyzak "A fast SVM training algorithm," *International Journal of Pattern recognition and Artificial Intelligence*, vol. 17, no. 3, pp. 367-384, 2003.



박 중 조(Joong-Jo Park)

1981.2 고려대학교 전기공학과 (공학사)
 1983.2 고려대학교 전기공학과 (공학석사)
 1995.2 고려대학교 전기공학과 (공학박사)
 2005.9~2006.8 캐나다 Concordia대학 방문
 교수

1996.3.~현재 경상대학교 제어계측공학과 교수.
 관심분야 : 컴퓨터비전, 생체인식.



김 경 민(Kyoung-Min Kim)

1988.2. 고려대학교 전기공학과 (공학사)
 1993.2. 고려대학교 전기공학과 (공학석사)
 1996.2. 고려대학교 전기공학과 (공학박사)
 2002.2.~2003.2. 캐나다 CENPARMI 방문
 교수

현재 전남대학교 전자통신전기공학과 부교수.
 관심분야 : 컴퓨터비전, 퍼지 및 신경회로망 응용,
 생체인식, 시스템자동화