

접촉된 필기체 숫자에 대한 효과적인 분할 기반 인식 방법

김 계 경[†] · 김 진 호^{††} · 박 희 주^{†††} · 부 기 동^{††††}

요 약

본 논문에서는 접촉된 숫자들에 대한 효과적인 분할 기반 인식 방법을 제안하였다. 접촉 숫자들을 연결획 정보와 분할 후보점을 기반으로 여섯 개의 접촉 유형으로 구분하였다. 전체 후보 분할점을 해석하여 네 개의 최종 후보 분할점을 도출함으로써 과 분할로 인한 오인식을 줄일 수 있도록 하였다. 이 방법에서는 다수의 분할 후보점으로부터 신뢰성이 높은 소규모의 분할 후보점들에 대해 우선권을 부여하는 방식으로 최종 분할 후보점들을 찾고 인식을 시도하기 때문에 전통적으로 분할기반 방식의 인식에서 초래되는 오분할에 의한 치명적인 오인식률을 줄일 수 있도록 하였다. NIST 접촉숫자 데이터 베이스에 대한 실험 결과 92.5%의 비교적 높은 인식 성능을 얻을 수 있었다.

An Efficient Segmentation Based Recognition of Unconstrained Handwritten Touching Digits

Kye-Kyung Kim[†] · Jin-Ho Kim^{††} · Hee-Joo Park^{†††} · Ki-Dong Bu^{††††}

ABSTRACT

This paper proposes an efficient segmentation-based recognition algorithm for unconstrained handwritten touching digits. We classify touching into six types based on an analysis of the ligature component and the characteristic of candidate segmentation points. Four kinds of candidate segmentation points are obtained from contour profiles and the final segmentation point of touching digits is determined by verifying the candidate segment combinations. The main advantages of this method are that reliable segment combinations are used in the multiple hypothesis recognition, and segmentation errors of traditional segmentation-based approach is reduced by introducing prioritized segmentation points. To evaluate the proposed method, we have experimented with 3,500 touching digits of the NIST database. An encouraging recognition rate of 92.5% has been obtained.

1. 서 론

무제약 (unconstrained) 필기체 숫자 열에 대한 인식[1-16]은 우편 번호, 양식 서류 및 수표 인식 등 다양한 응용 분야를 가지므로 최근까지 활발히 연구되고 있다. 일반적으로 숫자 열에는 독립된 개별 숫자, 접촉된 숫자, 겹친 숫자 및 여러 획으로 조각(broken) 난 숫자들이 포함되어 있다. 숫자 열에서 개별 숫자들을 분할하는 방법은 숫자 열을 이루는 개수에 대한 정보를 사전에 알고 있는 경우와 그렇지 않은 경우에 대해 구분되어 연구되고 있다. 그러나 인접하는 두 개의 숫자가 서로 접촉되어 있을 경우 이들을 개별 숫자로 분할하는 것이 매우 어렵기 때문에 이를 해결하기 위해 많은 분할 방법들 [1-10]이 별도로 연구되어 왔다.

숫자 열로부터 개별 숫자를 분할하는 방법은 분할 기반(segmentation-based), 인식 기반(segmentation-free) 및 전역(holistic) 방법 들로 분류될 수 있다. 분할 기반 방법 [11, 12]은 분할 특징 점들을 찾아 분할 법칙에 의해 인식 결과를 이용하지 않고 숫자 열에서 개별 숫자를 분할하여 인식하는 것이다. 수직 투영, 윤곽선 및 세션화된 영상의 특징 점들이 분할 점(segmentation point)을 찾는 데 이용되어 왔다. 수직 투영 방법은 처리 과정이 간단하기 때문에 많이 이용되어 왔으나 개별 숫자의 일부분을 분할시키는 문제점이 나타났다. 분할 기반 방법의 이러한 문제점을 개선하기 위하여 인식기반(segmentation-free) 분할 방법 [10, 14]이 제시되었다. 여기서는 여러 개의 후보 분할 점들에 의해 분리된 숫자들의 인식 결과에 따라 최종 후보 분할 점을 결정하도록 하였다. 이 방법은 분할 기반 방법의 오 분할에 의한 인식 오차는 줄일 수 있지만 최종 분할 점을 결정하기 위한 계산 과정이 매우 복잡하다. 위의 두 가지 방법은 숫자 열에

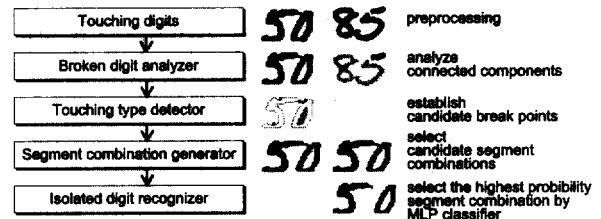
† 정 회 원 : 한국전자통신연구원 영상처리연구부 컴퓨터비전팀 선임연구원
 †† 정 회 원 : 경일대학교 전자정보공학과 교수
 ††† 정 회 원 : 경일대학교 컴퓨터공학과 교수
 †††† 종신회원 : 경일대학교 컴퓨터공학과 교수
 논문접수 : 2001년 3월 15일, 심사완료 : 2001년 4월 6일

서 개별 숫자로의 분할 과정을 거쳐 인식하는 방법들이다. 전역(holistic) 방법 [2]은 분할을 기반으로 하여 인식하는 위의 두 가지 방법들이 가질 수 있는 오 분할에 의한 인식 오차를 줄이기 위해서 접촉된 숫자들을 분할하지 않고 인식하는 것이다. 그러나, 인식하기 위해 추출해야 하는 특징의 수나 메모리가 많이 필요하며 복잡한 인식이 요구된다. 최근에는 각 분할 방법의 장점을 이용하여 서로의 문제점을 보완시켜 줄 수 있는 결합된 분할 방법이 연구되고 있다[10].

본 논문에서는 무제약 접촉된 필기체 숫자에 대한 효과적인 분할 기반 인식 방법을 제안하였다. 접촉된 숫자들의 윤곽선(contour profile)으로부터 네 가지 서로 다른 종류의 초기 후보 분할 점을 얻고 연결선(ligature), 후보 분할 점의 위치 및 특성에 따라 접촉 숫자들을 여섯 가지 접촉 유형으로 구분하였다. 최종 분할 점은 접촉 유형 분석 결과 얻어진 후보 분할 점들에 의해 분리된 숫자 조합들(segment combinations)의 인식 결과 가운데 최고의 신뢰도 값을 가지는 조합으로 결정 되도록 하였다. 이때 다수의 후보 분할점들 중에서 가능성이 높은 후보 분할점들에게 우선권을 부여하여 최종 후보 분할 점들을 찾음으로써 기존의 분할 기반 방법이 가지는 오 분할에 의한 치명적인 오 인식 결과를 줄이도록 하였다. 제안된 방법은 초기 후보 분할 점들에 대한 해석을 통해 인식 단계에서 최소화된 후보 분할 점들을 이용하여 접촉된 숫자를 분할함으로써 다수의 초기 후보 분할 점들에 의해 나뉘어진 숫자에 대한 오인식 결과를 줄일 수 있도록 하였다. 제안된 방법의 타당성을 검증하기 위해서 NIST 데이터 베이스의 접촉된 숫자들을 이용하여 인식 실험을 하였다.

2. 접촉된 숫자들의 분할

전체 인식 시스템은 조각난 숫자 해석기(broken digit analyzer), 접촉 유형 분류기(touching type analyzer), 후보 숫자 조합 생성기(candidate segments generator) 및 인식기로 구성되어 있다. (그림 1)은 접촉된 숫자들에 대한 전체 인식 시스템을 도시하였다. 조각난 숫자 해석기는 심하게 굴곡진(jagged) 윤곽선을 가진 입력 영상, 작은 영상 조각들에 대한 전처리 및 여러 획들로 나뉘어진 숫자를 찾아내는 과정들을 포함한다. 접촉 유형 분류기는 입력 영상의 윤곽선으로부터 후보 분할 점들을 찾아내고 연결선(ligature) 요소를 가지는 후보 분할 점들의 해석을 통해 접촉된 숫자들을 여섯 가지 접촉 유형으로 분류해낸다. 후보 숫자 조합 생성기는 접촉 유형 분류기로부터 이미 해석되어진 후보 분할 점들에 대해서 후보 숫자 조합들을 생성시킨다. 최종 분할 점은 후보 숫자 조합들의 인식 결과들로부터 최고의 신뢰도를 가지는 후보 숫자 조합에 따라 결정되도록 하였다.



(그림 1) 접촉된 숫자들에 대한 전체 인식 시스템 구성도

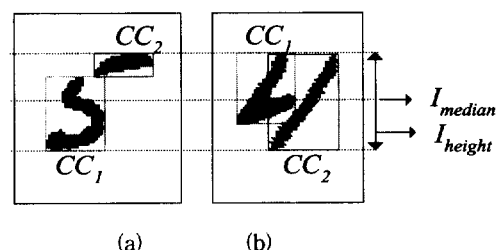
2.1 조각난 숫자 해석기

조각난 숫자 해석기는 입력 영상에서 작은 영상 조각들을 제거하고 여러 획으로 나누어진 숫자들을 찾아낸다. 개인적인 필기 습관이나 저해상도에서 영상을 입력할 경우 영상 획득 과정 중에 잡음이 섞인 작은 조각들 및 조각난 획들이 영상 내에 존재하게 된다.

입력 영상의 윤곽선은 후보 분할 점들을 찾는 데 이용되므로 굴곡이 심한 부분들을 제거하는 과정이 진행된다. 입력 영상의 레이블링은 입력 영상에서 연결된 요소들(connected components)을 찾기 위해 수행된다. 만약 입력 영상이 하나의 연결요소로 되어 있는 경우 입력 영상은 접촉된 숫자인지를 해석할 필요가 있다. 한편 여러 개의 연결된 요소들이 입력 영상 내에 있는 경우 잡음 화소들, 여러 획으로 나뉘어진 숫자 그리고 접촉된 숫자가 포함되어 있는 경우로 볼 수 있다.

잡음 화소들은 연결된 요소들과의 상대적인 크기를 비교함으로써 제거된다[17]. 조각난 획을 가지는 숫자는 입력 영상의 절반 높이에 해당되는 수평선과의 교차 여부 및 획의 높이를 조사하여 찾을 수 있다. 조각난 획을 가지는 숫자가 영상 내에 존재하는 경우 조각난 획은 후보 숫자 조합 생성기에서 후보 분할 점에 의해 나뉘어진 인접 영상과 합쳐져서 하나의 부 영상을 생성하게 된다. 조각 획에 대한 정보를 이용하면 전체 후보 숫자 조합들의 생성 수와 그에 대한 처리 시간을 줄일 수 있다.

조각난 숫자 해석을 위하여 조각난 획을 가지는 숫자 영상의 예를 (그림 2)에 나타내었다. 대부분의 조각난 획은 두 획으로 이루어진 숫자들 "4"와 "5"에서 볼 수 있다. 숫자 "4"와 "5"는 연결된 요소들이 수직 및 수평 방향으로 각각 나뉘어져 있다. 조각난 획을 가지는 숫자는 인식 단계 이전에 하나의 연결된 요소로 합친 다음 인식하도록 하였다.



(그림 2) 조각난 획을 가지는 (a) 숫자 "5" 및 (b) 숫자 "4"의 예

연결된 요소가 영상의 절반 높이에 해당하는 수평선, I_{median} 과 교차되지 않는 경우 조각난 획으로 해석된다. 연결 요소가 I_{median} 과 교차되지만 연결 요소의 높이가 영상의 절반 높이보다 작은 경우에도 조각난 획으로 고려된다. (그림 2)에서 숫자 “5”의 조각난 획 CC_2 또는 저해상도 영상의 조각난 획은 위의 조건들을 이용하여 찾을 수 있다. 그러나 조각난 숫자 “4”의 경우는 위의 조건을 만족하지 않으므로 인접하는 연결 요소간의 높이 비[10]를 조사하여 찾을 수 있다.

2.2 접촉 유형 분류기

접촉 유형 분류기는 네 가지 서로 다른 후보 분할 점들을 찾아내고 이들로부터 연결선, 분할 점들의 특성에 따른 접촉 유형 해석 및 조각난 숫자 해석기를 도입함으로써 적은 개수의 후보 분할 점들을 찾아낸다.

2.2.1 연결선 해석

연결선은 인접하는 두 숫자를 이어 주는 획으로서 숫자의 일부분이 인접하는 숫자에 접촉되어 있는 경우와 인접하는 숫자가 다른 획에 의해서 연결되어지는 경우에 나타난다. 전자는 인접하는 두 숫자가 허위 연결선(fake ligature)에 의해 접촉되어 있는 형태를 말하며 후자는 인접하는 숫자가 실제 연결선(real ligature)에 의해 접촉된 경우를 말한다. 연결선에 대한 해석은 연결선 특성을 가지는 후보 분할 점들에 대해서 행해진다. 후보 분할 점들은 입력 영상의 구조적인 특징 점들로부터 구해진다. 네 가지 서로 다른 특성을 가지는 후보 분할 점 가운데 두 종류의 분할 점들은 상위(upper) 윤곽선과 하위(lower) 윤곽선 사이의 수직 성분의 차이 $D(x, y)$ 로부터 아래의 식과 같이 구해지며 연결선 해석에 이용되는 후보 점들이다.

$$B_{up}(x) = \max\{y \in \{f(x, y) \in C_{up}\}\} \quad (1)$$

$$B_{dn}(x) = \min\{y \in \{f(x, y) \in C_{dn}\}\} \quad (2)$$

$$D(x) = B_{up}(x) - B_{dn}(x) \quad (3)$$

여기서 $B_{up}(x)$ 와 $B_{dn}(x)$ 는 상·하위 윤곽선 C_{up}, C_{dn} 상에서 $f(x, y)$ 화소들의 각 x 축 좌표에서 최대 y 값 및 최소 y 값들로 구성된다. 나머지 두 종류의 후보 분할 점들은 각 윤곽선에서의 국소 최대 및 최저 점들로부터 구해진다. 네 가지 종류의 후보 분할 점들은 아래의 식들과 같이 정의되었다.

$$P_1(x, y) = \left\{ f(x, y) \in C_{up} \mid \frac{D(x)}{D(x+1)} > T_1, D(x) > 0 \right\}, \quad T_1 > 2.0 \quad (4)$$

$$P_2(x, y) = \left\{ f(x, y) \in C_{up} \mid \frac{D(x+1)}{D(x)} > T_2, D(x) > 0 \right\}, \quad T_2 > 1.7 \quad (5)$$

$$P_3(x, y) = \left\{ f(x, y) \in C_{up} \mid f(x, y) \in \min_x [C_{up}(x, y) \mid y] \wedge f(x, y) \in \min_x [C_{up}(x, y) \mid y] \right\} \quad (6)$$

$$P_4(x, y) = \left\{ f(x, y) \in C_{dn} \mid f(x, y) \in \min_x [C_{dn}(x, y) \mid y] \wedge f(x, y) \in \min_x [C_{dn}(x, y) \mid y] \right\} \quad (7)$$

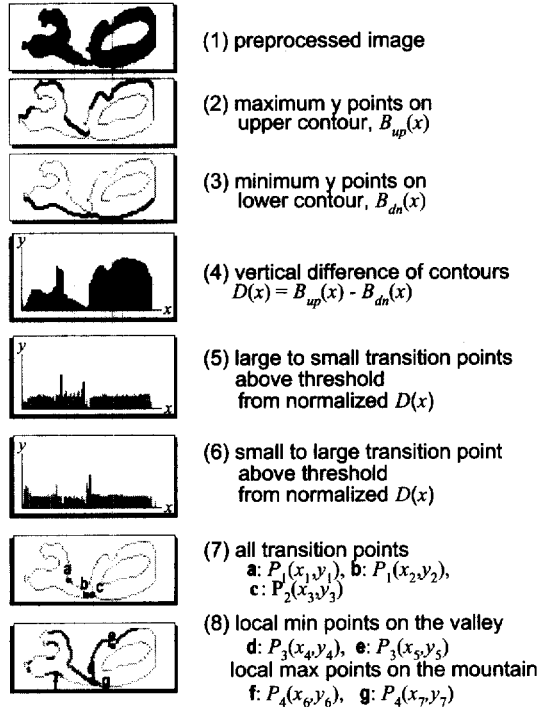
위 식에서 임계치는 실험적으로 구하였으며 정의된 $P_1(x, y)$ 는 인접하는 x 값들 사이에서 수직 윤곽선의 변화가 큰 값에서 작은 값으로 변할 때 그 값이 임계치 이상이 되는 점들로서 인접하는 두 숫자를 잇고 있는 연결선의 시작점으로도 판단된다. $P_2(x, y)$ 는 $P_1(x, y)$ 와 반대되는 특성을 가지는 점들로서 수직 윤곽선의 변화가 작은 값에서 큰 값으로 바뀌어지고 그 값이 임계치 이상이 될 때 선택되며 연결선의 끝점으로도 판단된다. $P_3(x, y)$ 와 $P_4(x, y)$ 는 각 윤곽선의 계곡 및 언덕 영역 위의 점들로 이루어진다. 입력 영상의 윤곽선들과 수직 윤곽선의 차 및 각 윤곽선의 특징으로부터 네 가지 종류의 후보 분할 점들을 찾아내는 과정을 (그림 3)에 나타내었다

연결선 해석은 연결선으로 인한 분할 오차가 인식에 미치는 영향을 줄이기 위한 것이다. 연결선 해석은 연결선의 시작점과 끝점에 해당되는 후보 분할 점들을 가지는 획에 대해서 행해진다. 여기서 획이 접촉된 숫자들 사이에 존재하는 실제 연결선인지 숫자의 일부분이 이웃하는 숫자에 접촉되어 있는 허위 연결선인지를 조사하게 된다. 대부분 연결선은 접촉된 영상의 상위 영역에서 볼 수 있으므로 여기서는 상위 연결선 해석에 국한시켰다. 하위 영역에 존재하는 연결선은 숫자들 “2” 및 “3”이 다른 숫자에 연결된 경우와 유사하므로 인식 결과가 거절되었을 경우에만 다시 해석하도록 하였다. 연결선은 다음의 조건들에 의해서 찾을 수 있다.

$$\left\{ \text{foreground pixels between } f(x_1, y_1') \text{ and } f(x_2, y_2') \text{ on } C_{dn} \text{ are above } I_{median} \right\} \\ \frac{1}{x_2 - x_1 + 1} \sum_{i=x_1}^{x_2} D(i) < T_{diff}$$

여기서 x_1 과 x_2 는 $P_1(x_1, y_1)$ 과 $P_2(x_2, y_2)$ 의 x 값이고, y_1' 과 y_2' 는 C_{dn} 위의 y 값이다. 첫 번째 조건으로부터 후보 분할점 $P_1(x_1, y_1)$ 과 $P_2(x_2, y_2)$ 에 따라 찾은 획이 I_{median} 위의 상위 영역에 존재하면 상위 연결선으로 고려되어진다. 두 번째 조건은 후보 연결선으로 고려된 획이 평균 수직 윤곽선의 차, T_{diff} 보다 작으면 수평 획으로 간주될 수 있음을 의미한다. 이상의 두 조건이 모두 만족되면 획은 연결선으로 간주되어 연결선을 제거 시킨 숫자 영상도 인식 단계에

서 고려될 수 있도록 하였다. 이상의 두 조건이 모두 만족되지 않는 경우는 숫자의 일부분이 이웃하는 숫자에 접촉되어진 형태로 간주하였다.



(그림 3) 접촉된 숫자의 상 하위 윤곽선, 두 윤곽선의 수직 성분차 및 후보 분할 점들의 예

2.2.2 접촉 유형 해석

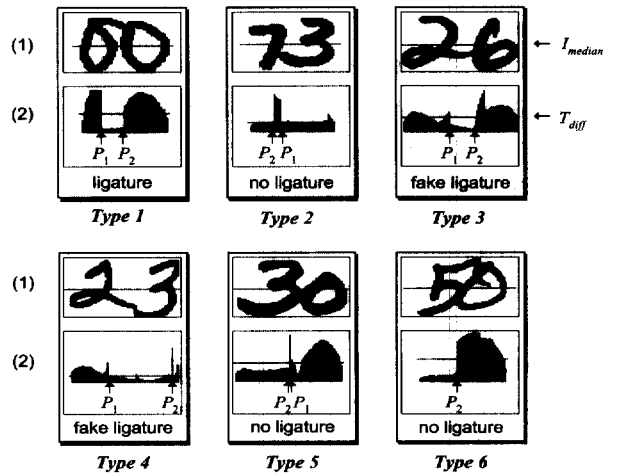
접촉된 숫자들은 각각 구별되는 후보 분할 점들을 가지므로 이러한 특성을 반영하여 여섯 종류의 접촉 유형을 정의하였다. 여섯 종류의 접촉 유형은 연결선 해석, 각 후보 분할 점의 분포 및 위치 정보를 이용하여 정의되었다. (그림 4)는 여섯 종류의 접촉 유형을 가지는 접촉된 숫자들의 예를 보여 준다.

유형1(type 1)은 이웃하는 숫자들 사이에 연결선이 존재하며, 유형 2, 3 및 4 는 숫자의 일부분이 이웃하는 숫자들에 접촉된 형태로서 마치 연결선에 의해 서로 접촉된 것처럼 보인다. 유형 5와 6은 접촉된 숫자 사이에 연결선이나 허위 연결선이 존재하지 않으며 각각 숫자의 한 부분 및 다수 부분이 접촉된 숫자가 이에 해당 된다. 여섯 가지 접촉 유형으로 구분하기 위한 수식은 아래와 같다.

- (1) Type 1 : $\{ \exists P_1(x_1, y_1) \text{ and } \exists P_2(x_2, y_2) \} \wedge \{ y_1 \text{ and } y_2 \text{ above } I_{median} \}$
- (2) Type 2 : $\{ \exists P_1(x_1, y_1) \} \wedge \{ y_1 \text{ below } I_{median} \}$
- (3) Type 3 : $\{ \exists P_2(x_2, y_2) \} \wedge \{ y_2 \text{ below } I_{median} \}$
- (4) Type 4 : $\{ \exists P_3(x_1, y_1) \text{ or } \exists P_4(x_2, y_2) \} \wedge \{ y_1 \text{ and } y_2 \text{ above } I_{median} \}$

- (5) Type 5 : $\{ \exists P_3(x_1, y_1) \text{ and } \exists P_4(x_2, y_2) \} \wedge \{ dist(P_3(x_1, y_1), P_4(x_2, y_2)) < T_3 \}$
- (6) Type 6 : $\{ \exists P_3(x_1, y_1) \text{ and } \exists P_4(x_2, y_2) \} \wedge \{ dist(P_3(x_1, y_1), P_4(x_2, y_2)) < T_3 \} \wedge \{ \text{line between } P_3(x_1, y_1) \text{ and } P_4(x_2, y_2) \text{ cross the inner contour} \}$

T_3 와 $dist(\cdot)$ 는 평균 획의 폭 및 두 점 사이의 거리를 나타낸다.



(1) six touching types
(2) vertical difference of contours $D(x)$ with T_{diff} and $P_1(\cdot), P_2(\cdot)$

(그림 4) 여섯 종류의 접촉 유형을 가지는 숫자들의 예

2.2.3 비수직 후보 분할 점(non-vertical candidate segmentation points)

유형 5와 6에 속하는 접촉된 숫자들이 수직 후보 분할선에 의해 나뉘어질 경우 숫자의 일부분이 잘려나가는 경우가 생기므로 비수직 후보 분할 점의 선택은 불가피하다. 이는 각각 상·하위 윤곽선 위의 점들 $P_3(x_1, y_1)$ 와 $P_4(x_2, y_2)$ 로 구성된다. 식 (8)의 비 수직 후보 분할 점 $P_5((x_1, y_1), (x_2, y_2))$ 은 식 (9)의 조건을 만족하는 점들이다.

$$P_5((x_1, y_1), (x_2, y_2)) = \{ f(x, y) \in C_{up} \vee f(x, y) \in C_{dn} \mid f(x_1, y_1) \in P_3(x, y) \wedge f(x_2, y_2) \in P_4(x, y) \} \quad (8)$$

$$\theta_{th} = \tan^{-1} \left[\frac{y_2 - y_1}{x_2 - x_1} \right]$$

$$\angle(f(x_1, y_1), f(x_2, y_2)) < \theta_{th}$$

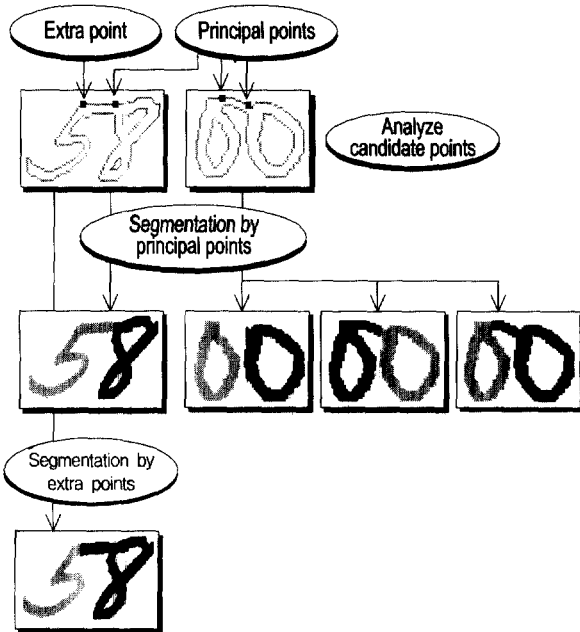
where $\theta_{th} = \pm(\theta_s \pm 0.1 \times \theta_s)$ (9)

식 (9)에서 θ_{th} 는 숫자의 기울어진 각도 θ_s 에서 구해진 결정치 각이다. 이상과 같은 후보 분할점들의 해석과 조각난 숫자 해석을 바탕으로 찾아낸 초기 후보 분할 점들 가운데 불필요한 후보 분할점을 제거하여 신뢰성 있는 최소의 후보 분할점들을 얻을 수 있도록 하였다.

3. 후보 숫자 조합 생성기

후보 숫자 조합 생성기는 조각난 숫자, 연결선 및 접촉 유형에 의해 해석된 후보 분할 점들을 이용하여 후보 숫자 조합(candidate segment combinations)들을 생성시킨다. 후보 숫자 조합들의 생성 예를 (그림 5)에 나타내었다.

후보 숫자 조합들은 우선권이 부여된 점들, 즉 핵심(principal) 후보 분할 점과 외적(extra) 후보 분할 점들에 의해 나누어진 두개의 후보 숫자 부 영상으로 구성되어 있다. 핵심 후보 분할 점은 접촉 유형 분류기에서 찾아낸 후보 분할 점 가운데 분할 가능성이 가장 높은 점으로 결정되며, 나머지 후보 분할 점들은 외적 후보 분할 점으로 고려된다. 해석되지 않은 다수의 후보 분할 점에 의해 나누어진 숫자 조합들의 인식 결과는 그렇지 않은 경우에 비해 다른 숫자로 오인식 할 확률이 더 크다. 또한 기존의 분할 기반 방법의 문제점인 핵심 분할 점에 의한 오분할 결과는 오인식의 원인이 되므로 이상의 문제점들을 줄이고자 우선권이 부여된 후보 분할 점들에 의한 분할 방법이 제안되었다.



(그림 5) 해석된 후보 분할 점들에 의해 생성된 후보 숫자 조합들의 예

분할 과정은 하나의 연결 요소를 가지는 경우와 그렇지 않은 경우에 대해 진행된다. 하나의 연결 요소를 가진 경우, 이미 해석되어진 후보 분할 점들을 이용하여 접촉된 숫자를 분할하여 숫자 조합들을 구성한다. 반면, 여러 개의 연결된 요소를 가진 경우는 조각난 숫자의 획 성분을 같이 고려하여 후보 숫자 조합들을 구성한다. 위 두 가지 경우 모두 핵심 후보 분할 n_1 점 및 외적 후보 분할 n_2 점 들을 가질 수 있다. 만약 $n_1=1$ 이라면, 연결선 없이 두 개의 숫

자가 서로 접촉하고 있는 형태로 $n_1=2$ 이면 접촉된 숫자가 연결선을 가지는 것으로 해석한다. 전자의 경우는 후보 분할점에 의해 나뉘어진 두 개의 숫자 부 영상으로 후보 숫자 조합을 생성하고, 후자의 경우는 연결선이 제거된 숫자의 조합 및 첫번째, 두 번째 핵심 후보 분할점에 의해 나뉘어진 부 영상 숫자 조합들을 구성한다. 외적 후보 분할 점의 경우는 전자와 같은 방법으로 숫자 조합들을 구성한다. 최종 분할 점은 후보 숫자 조합들의 인식 결과 가장 높은 신뢰도를 가지는 숫자 조합이 선택되도록 하였다. 분할된 부 영상 숫자 조합들을 인식하기 위해 다층 퍼셉트론이 사용되었다. 인식률의 오차를 줄이기 위해 분할된 숫자 부 영상 조합의 인식 확률 값을 반영하여 숫자 조합의 인식 확률 값을 합한 값이 임계치 이상에서 최대값을 가지면 최종 인식 결과로 받아 들이도록 하였다. 임계치는 독립된 숫자들의 인식 실험 과정에서 구한 인식 확률값들을 평균 한 값 으로부터 다음과 같이 구한다.

$$R_1 = \frac{1}{N_d} \frac{1}{N_s} \sum_{i=0}^{N_d} \sum_{j=0}^{N_s} r_{ij} \quad (10)$$

$$R_2 = R_1 - 0.1 \times R_1 \quad (11)$$

여기서 r_{ij} 는 i 번째 숫자 클래스에서 j 번째 인식 실험 숫자의 인식 확률값을 나타낸다. N_d 와 N_s 는 열 개의 숫자 클래스 및 개별 숫자 클래스 내에 있는 인식 실험 숫자들의 개수를 나타낸다. 식 (11)에서 임계치 R_2 를 R_1 보다 10% 낮은 값으로 결정한 것은 분할된 부 영역들에 대한 실험이 독립된 숫자들에 의해 훈련된 가중치를 사용하여 행해졌기 때문이다. R_2 는 인식 실험 과정에서 사용된 영상이 얼마나 잘 분할되었는지를 나타내는 기준치로 사용되었다.

4. 다층 퍼셉트론

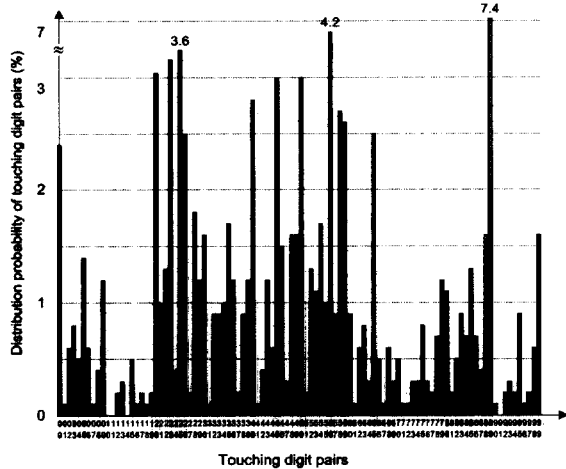
인식을 위한 여섯 가지의 피쳐를 추출하기 위해서 체인 윤곽선과 평활화 된 이진 영상이 이용되었다. 사용된 피쳐는 추출 방법이 간단하면서도 인식 성능이 뛰어난 메쉬 피쳐와 체인 코드, 내부 구멍의 개수, 영상의 높이에 대한 넓이의 비, 전이(transition) 및 거리 피쳐들이 사용되었다. 메쉬 피쳐는 분할된 이진 영상을 부 영역 $n_1 \times m_1$ 으로 나누어 각 부 영역에 존재하는 화소 값들을 합하여 정규화 시킨 값으로 구하였다. 체인 코드 피쳐는 부 영역 $n_2 \times m_2$ 에서 각각 8방향 성분 값을 합하여 정규화 시킨 것이다. 전이 피쳐는 수직 및 수평 방향 각각에 대해서 n_3 및 m_3 간격마다 배경 화소에서 문자 화소로 바뀌는 점의 횟수를 합하여 정규화 한 것이다. 거리 피쳐는 n_4 및 m_4 간격마다 수평 및 수직 방향으로 최소 사각 영역의 테두리에서 첫번째 문자 화소까지의 거리를 정규화 시킨 것이다. 또한 분할된 영상

내에 존재하는 구멍의 개수를 피쳐 벡터에 추가하였다.

분할된 영상을 인식하기 위해 한 개의 은닉층을 가지는 MLP(multi-layer perceptron)를 이용하였고 개선된 *backpropagation* 알고리즘[18]으로 구현하였다.

5. 실험

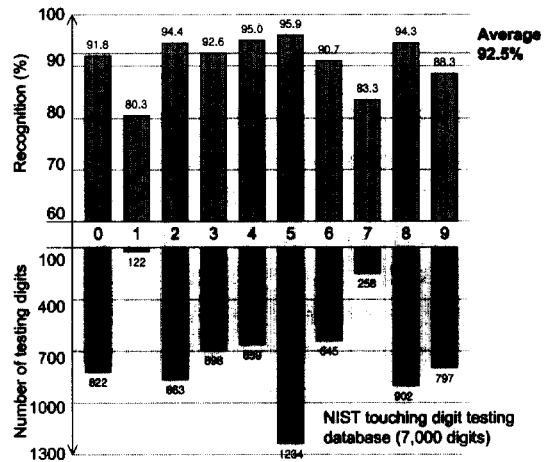
제안된 방법의 성능을 평가하기 위하여 NIST데이터 베이스의 접촉된 숫자들을 가지고 실험 하였다. 반복적인 실험에 따라 입력층, 은닉층 및 출력층의 뉴런 개수를 각각 176, 120 및 10개로 정하고 이를 이용하여 다층 퍼셉트론을 구현한 다음 50,000개의 독립된 숫자들을 훈련시키고 3,500개의 접촉된 숫자들에 대해 평가해 보았다. (그림 6)은 시물레이션에 이용된 NIST 접촉 숫자들의 분포를 나타낸 것이다.



(그림 6) 인식 실험에 사용된 접촉 숫자들의 분포

그림으로부터 숫자 “2”와 “5”는 다른 클래스의 숫자들에 비하여 이웃하는 숫자와 많이 접촉하고 있음을 알 수 있다. 이는 연결선 해석을 상위 연결선에 국한 시킨 타당성을 뒷받침해 주고 있다.

(그림 7)은 분할된 숫자들에 대한 인식 결과를 10개 클래스 각각에 대하여 나타낸 것이다. 대부분의 숫자들은 비교적 높은 인식 결과를 보여 주는 반면 숫자 “1”과 “7”은 많이 접촉된 숫자가 아님에도 불구하고 낮은 인식 결과를 보여 준다. 분할된 숫자 “1”과 “7”은 단순한 수직 및 수평 획 성분으로만 구성되어 있으므로 분할된 영상 내에 이웃하는 숫자의 획 일부분이 포함되어 있을 경우 다른 숫자로 인식되기 쉽다. 이는 독립된 숫자 데이터 베이스를 이용하여 훈련시킨 가중치를 분할된 숫자의 인식 실험에 이용한 결과이기도 하다. <표 1>은 각 클래스에 속하는 숫자들의 인식 결과를 나타낸 것이다.



(그림 7) 각 클래스에 따른 접촉된 숫자들의 인식 결과

<표 1> 10개의 숫자 클래스 각각에 대한 인식 행렬

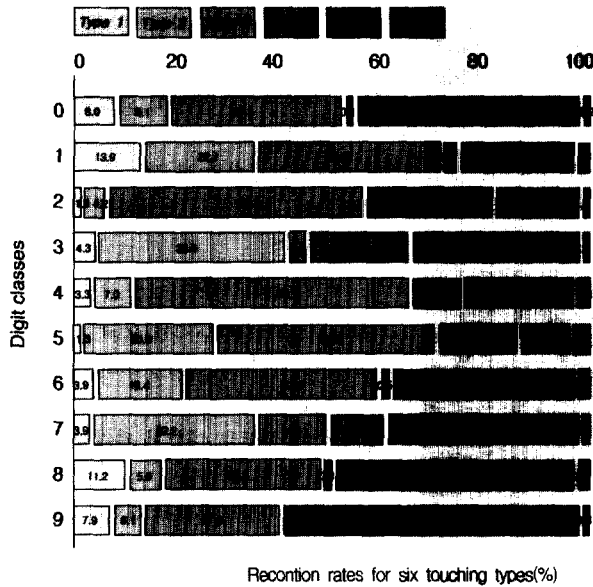
Classes[%]	0	1	2	3	4	5	6	7	8	9
0 [91.8]	755	1	16	10	6	8	10	6	9	
1 [80.3]	2	98	8	2	4	0	3	5	0	0
2 [94.4]	17	1	815	2	8	0	5	12	3	0
3 [92.6]	6	0	14	646	6	7	3	9	1	6
4 [95.0]	2	11	10	0	626	1	1	3	1	4
5 [95.9]	6	1	7	7	13	1183	3	3	7	4
6 [90.7]	17	9	15	2	9	3	585	1	2	2
7 [83.3]	2	6	23	3	3	0	0	215	2	4
8 [94.3]	5	1	11	3	8	8	6	3	851	6
9 [88.3]	8	21	7	6	21	5	2	14	9	704

<표 2>는 여섯 가지 접촉 유형에 따른 인식 결과를 각 클래스별로 나타낸 것이다. 유형 1부터 5에 속한 숫자들을 분할 하여 인식해 본 결과 비교적 높은 인식률을 얻을 수 있었다. 그러나 유형 6, 즉 다수의 접촉 점에 의해 연결된 숫자들을 분할 한 후의 인식 결과는 다른 유형에 비해 낮은 인식률을 나타낸다. 유형 6에 속한 숫자들은 대부분 비수직 분할 선에 의하여 분할되는데 이러한 분할 선을 구성하는 후보 분할 점 쌍 가운데 한 점이라도 후보 분할 점으로 발견되지 않는 경우 수직 분할 선에 의해 분할을 시도하게 되므로 숫자의 일부분이 잘려 나가는 경우가 생기게 된다. 이러한 오 분할이 오 인식의 원인이 되었다. 이러한 문제점은 내부 윤곽선에 의한 후보 분할 점의 추가 및 분할 없는 인식 방법에 의해 해결될 수 있다.

<표 2> 여섯 가지 접촉 유형에 따른 인식률

Type	# of testing samples	Recognition	Rate (%)
1	366	325	88.8
2	1058	980	92.6
3	2494	2344	94.0
4	672	646	96.1
5	2262	2084	92.1
6	148	99	66.9

(그림 8) 은 여섯 가지 접촉 유형에 속하는 숫자들의 10 클래스에 따른 인식 결과를 나타낸 것이다. <표 3>은 다른 연구 결과들과 인식 성능을 비교한 것이다.



Recognition rates for six touching types(%)

(그림 8) 각 클래스의 여섯 가지 접촉 유형에 따른 인식 분포

<표 3> 접촉 숫자들의 인식 성능 비교

Authors	Database	# of samples in training set	# of samples in testing set	Rejection rate(%)	Recognition rate(%)
Shi, 1997	CEDAR	2,819	1,089	-	80.8
Wang, 2000	CEDAR	2,819	2,778	-	85.1
Lu, 1999	NIST	53,449	3,355	4.7	92.5
This study	NIST	50,000	3,500	0	92.5

<표 3>에서 접촉된 숫자들에 대한 인식 결과는 서로 다른 데이터 베이스, 훈련 패턴 및 실험 패턴을 이용하여 수행되었으므로 직접적인 비교를 하는 것은 어렵다. 일반적으로 CEDAR 데이터 베이스보다 NIST데이터 베이스를 이용하여 인식한 결과 더 높은 인식률을 얻었음을 볼 수 있다.

6. 결 론

본 논문은 무제약 필기체 접촉 숫자에 대하여 우선권이 부여된 분할 방법을 이용한 인식 알고리즘을 제안하였다. 네 종류의 후보 분할점 및 여섯 종류의 접촉 유형이 정의되었고 연결선, 후보 분할점의 구조적인 특성 및 접촉 유형을 이용하여 후보 분할점들을 분석하였다. 우선권이 부여된 후보 숫자 조합을 구성하여 분할을 시도함으로써 기존의 분할 오차를 줄일 수 있도록 하였다. NIST의 접촉된 숫자 3,500개의 표본 영상을 이용하여 실험한 결과 92.5%의 인식률을 얻을 수 있었다. 보다 개선된 인식 결과를 얻기 위하여 조합된 분할 방법의 이용, 후처리 및 개선된 분할 알고리즘 개발에 대한 연구가 계속 될 예정이다.

참 고 문 헌

- [1] Z. Shi and V. Govindaraju, "Segmentation and recognition of connected handwritten numeral strings," *Pattern Recognition*, Vol.30, No.9, pp.1501-1504, 1997.
- [2] X. Wang, V. Govindaraju, and S. Srihari, "Holistic recognition of handwritten character pairs," *Pattern recognition*, pp.1967-1973, 33, 2000.
- [3] R. Fenrich, "Segmentation of Words," *Pre-Proc. of IWFHR*, Bonas, France, 1991.
- [4] H. Nishida and S. Mori, "A model-based split-and-merge method for character string recognition," *Intl. Journal of Pattern recognition and Artificial Intelligence*, Vol.8, No.5, pp.1205-1222, 1994.
- [5] H. Fujisawa et al, "Segmentation models for character recognition : Form segmentation to document structure analysis," *Proc. of IEEE*, Vol.8, No.7, pp.1079-1082, 1992.
- [6] B. Zhao, H. Su and S. Xia, "A new method for segmenting unconstrained handwritten numeral string," *ICDAR*, Aug. 18-20, Vol.2, pp.524-527, 1997.
- [7] J. Keeler and D. E. Rumelhart, "A self-organizing integrated segmentation and recognition neural network," *Advances in Neural Information Processing Systems*, Vol.4, J. E. Moody, S. J. Hanson, and R. P. Lippmann, eds., pp.496-503, Morgan Kaufmann, San Mateo, CA, 1992.
- [8] A.Filatov, A.Gitis and I.Kil, "Graph-based handwritten digit string recognition," *Proc. 3rd ICDAR*, 14-16 Aug., pp.845-848, Montreal Canada, 1995.
- [9] Nishiwaki et al, "A new recognition method of Black ruled line forms including connected handwritten numerals," *Proc. of 2nd IAPR Workshop on Graphics Recognition*, pp.152-159, 1997.
- [10] T. M. Ha, M. Zimmermann and H. Bunke, "Off-line handwritten numeral string recognition by combining segmentation-based and segmentation-free methods," *Pattern Recognition*, Vol.31, No.3, pp.257-272, 1998.
- [11] Y. Lu and M. shridhar, "Character segmentation in handwritten words An overview," *Pattern Recognition*, Vol.29, No.1, pp.77-96, 1995.
- [12] N. W. Strathy, C. Y. Suen et al, "Segmentation of handwritten Digits using contour features," *Proc. of ICDAR*, pp.577-580, 1993.
- [13] R. G. Casey and E. Lecoliner, "Strategies in character segmentation : A survey," *Proc. of the ICDAR*, pp.1028-1033, Montreal, Canada, 1995.
- [14] J. Hu, D. Yu, and H. Yan, "Construction partitioning paths for touching handwritten characters," *Pattern Recognition Letters*, Vol.20, pp.293-303, 1999.
- [15] Z. Lu, Z. Chi, W. Siu, and P. Shi, "A background-thinning

based approach for separating and recognizing connected handwritten digit strings," *Pattern Recognition*, Vol.32, No.6, pp.921-933, 1999.

- [16] G. Congedo, G. Dimauro, S. Impedovo, G. Pirlo, "Segmentation of numeric strings," *Proc. of the 3rd ICDAR*, Aug. 14-16, pp.1038-1041, 1995.
- [17] S. Madhvanath, G. Kim, and V. Govindaraju, "Chaincode contour processing for handwritten word recognition," *IEEE Trans.*, on PAMI 21, pp.928-932, 1999.
- [18] B. Kosko, "Neural networks and fuzzy system," *Prentice-Hall*, Englewood Cliffs, NJ, 1992.



김 계 경

e-mail : kyekyung@etri.re.kr
 1989년 경북대학교 전자공학과 졸업
 (공학사)
 1992년 경북대학교 대학원 전자공학과
 졸업(공학석사)
 1997년 경북대학교 대학원 전자공학과
 졸업(공학박사)

1998년~2001년 캐나다 콘코디아대학 CENPARMI연구소 방문 연구원

2001년~현재 한국전자통신연구원 컴퓨터비전팀 선임연구원
관심분야 : 패턴인식, 문자인식, 영상처리, 신경회로망



김 진 호

e-mail : kjinho@bear.kyungil.ac.kr
 1986년 경북대학교 전자공학과 졸업
 (공학사)
 1988년 경북대학교 대학원 전자공학과
 졸업(공학석사)
 1992년 경북대학교 대학원 전자공학과
 졸업(공학박사)

1992년~현재 경일대학교 전자정보공학과 부교수
관심분야 : 패턴인식, HMM, NN, 필기인식, 영상처리



박 회 주

e-mail : hjpark@bear.kyungil.ac.kr
 1978년 영남대학교 전자공학과 졸업
 (공학사)
 1981년 영남대학교 대학원 전자공학과
 졸업(공학석사)
 1995년 대구가톨릭대학교 대학원 전산통
 계학과 졸업(이학박사)

1982년 현재 경일대학교 컴퓨터공학과 교수
관심분야 : 패턴인식, 신경회로망



부 기 동

e-mail : kdbu@bear.kyungil.ac.kr
 1984년 경북대학교 전자공학과 졸업
 (공학사)
 1988년 경북대학교 대학원 전자공학과
 전산공학 전공(공학석사)
 1996년 경북대학교 대학원 전자공학과
 전산공학전공(공학박사)

1983년~1985년 포항종합제철 시스템개발실
1988년~현재 경일대학교 컴퓨터공학과 부교수
관심분야 : 데이터베이스, GIS