2001 년 한국정보처리학회 추계 학술발표논문집 제 8 권 제 2 호

고속 필기 한글 주소 인식을 위한 낱자 인식

정선화, 임길택, 송재관, 남윤석
한국전자통신연구원, 우정기술연구부, 자동구분처리연구팀
e-mail : sh-jeong@etri.re.kr

Character Recognition for Fast Handwritten Korean Address Reading

Seon-Hwa Jeong, Kil-Taek Lim, Jae-Gwan Song, Yun-Seok Nam
Postal Technology Development Department, ETRI

요 약

본 논문에서는 고속 필기 한글 주소 인식을 위한 낱자 인식기를 제안한다. 인식 대상은 우편번호 여섯 자리에 할당된 주소에 출현 빈도가 높은 필기 한글 469 자이다. 제안된 방법은 낱자 인식 기법을 채택하고 있으며, 인식률과 처리속도를 향상시키기 위하여 2 단계 인식 전략을 채택하였다. 인식기로는 다층퍼셉트론, 최소거리분류기, Subspace 방법을 고려한다. 다층퍼셉트론은 비교적 높은 인식률과 처리속도를 보유하지만 출력값이 확률이 아님으로써 후처리를 필요로 하는 시스템에서 사용하기 어렵다. 최소거리분류기는 간단한 알고리즘으로 처리속도가 높고 확률을 출력하는 장점을 갖지만 인식률이 낮아 활용되기 어렵다. 또한 Subspace 방법은 인식률이 높고 확률을 출력하지만 처리속도가 매우 느리다는 단점이 있다. 따라서 제안방법에서는 처리속도가 높은 인식기 - 다층퍼셉트론, 최소거리분류기 - 를 사용하여 선인식을 수행한 후, 이 결과를 활용하여 인식 대상을 제한한 후 Subspace 방법을 사용하여 정확하게 인식하는 전략을 도입함으로써, 높은 인식결과를 유지하면서 처리속도를 높이고 후처리에 적합하도록 하였다. PE92 데이터베이스를 사용하여 실험한 결과 제안방법이 한글 469 자에 대하여 비교적 높은 인식률과 처리속도를 갖음을 알 수 있었다.

1. 서론

필기 한글 낱자의 오프라인 인식이란, 스캐닝된 문서상에 존재하는 필기 한글 문자를 자동으로 인지하는 기술을 의미하며, 전표처리 자동화, 우편물 자동분류, 팩스배달, 전자도서관 구축 등에 응용할 수 있다.

한글이나 한자처럼 인식 대상이 방대한 경우 인식 문제는 매우 어려워진다. 이는 동일한 낱자에 대한 다양한 필체들 간의 변화를 흡수해야 하고, 또한 유사한 모양을 갖는 서로 다른 낱자들 간의 변별력을 높여야 하는 두 가지의 상충된 문제를 동시에 해결해야 하기 때문이다. 한글의 경우, 24 개의 기본 자소가 2 차원 평면상에 조합되어 최대 11,172 개의 낱자를 형성하기 때문에 유사한 모양을 갖는 패턴의 종류가 많다. 이러한 문제점으로 인해 지난 수 십년 동안 당 분야에 대한 많은 연구[1, 2]에도 불구하고 현재까지 실제 응용분야에 실용화될 수 있을 만큼 우수한 성능을 보유한 인식 기술은 없는 실정이다.

필기 한글 인식과 관련된 기존 연구를 살펴보면 크게 두 가지로 분류할 수 있다. 즉, 한글의 제작 원리에 입각하여 낱자를 자소 단위로 분리하여 인식하는 자소 단위 인식 기법과 낱자 자체를 모델링하여 인식하는 낱자 단위 인식 기법으로 나눌 수 있다. 자소 단위 인식 기법[3, 4]은 자소 인식을 수행하기 때문에 낱자보다 인식 대상의 수가 줄어들어 그 만큼 인식이 수월해지는 장점을 가지고 있다. 반면, 자소의 위치도 일정하지 않고 자소 간의 접촉도 많은 필기 한글 낱자를 자소 단위로 분리해야 한다는 큰 어려움을 갖는다. 자소 분리 문제는 또 다른 연구 과제이다. 이와는 반대로 낱자 단위 인식 기법은 자소 분리의 어려움을 갖지 않는 대신, 인식 대상이 되는 모든 낱자에 대하여 모델을 만들어야 하기 때문에 11,172 자 대상의 인식은 사실 불가능하다. 무엇보다도 11,172 자를 모델링하기 위한 데이터가 아직까지 마련되어 있지 않다. 따라서 대부분의 낱자 단위 인식은 응용분야의 정보를 활용하여 인식 대상을 줄임으로써 적용된다[5, 6].

본 논문에서는 고속 주소 인식에 적합한 필기 한글 낱자 인식 방법을 제안한다. 주소 인식을 목표로 하기 때문에 인식 대상을 주소에 빈번히 출현하는 469 자로 제한할 수 있었으며, 이에 따라 낱자 단위 인식 기법을 채택하여 자소 분리의 어려움을 피하였다. 제안된 방법에 대한 자세한 설명은 2 절에서 기술할 것이다.

2. 제안 방법

제안된 낱자 인식 방법은 우편봉투에 필기된 주소를 고속으로 인식하기 위한 시스템의 일부분으로 개발되었다. 따라서 낱자 인식기는 정확한 인식 성능을 갖추어야 될 뿐만 아니라, 신뢰할 만한 인식 후보를 출력할 수 있어야 하고 처리 속도가 우수해야 한다. 본 논문은 위의 세 조건을 만족하는 인식 방법에 대한 연구 결과이다.

제안 방법은 자소 분리의 어려움을 피하기 위하여 낱자 단위 인식 기법을 채택하였으며, 이는 응용분야의 정보를 활용하여 인식 대상을 대폭 줄일 수 있었기 때문이다. 즉, 본 연구에서는 우편번호 여섯 자리와 대응되는 주소열을 인식하는데 주안점을 두고 있으므로, 낱자 인식기의 인식 대상을 모든 한글 낱자에서 주소열에 빈번히 출현하는 낱자들로 제한하였다. 그 결과 인식 대상의 수는 11,172 자에서 469 자(그림 1 참조)로 줄어들었다.

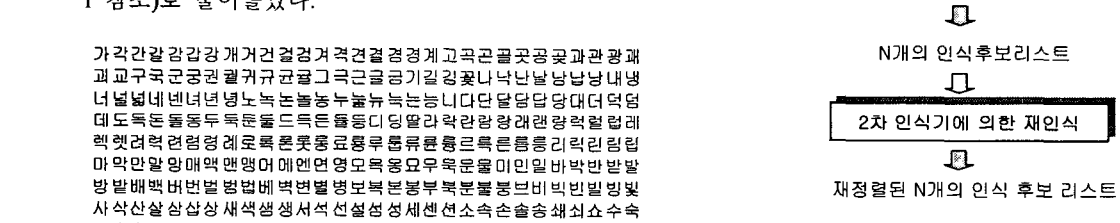

그림 1. 제한 방법의 필기 한글 인식 대상: 469 자

469 자에 대한 낱자 단위 인식을 수행하기 위하여 다층퍼셉트론, 최소거리분류기, Subspace 방법을 고려하였다. 다층퍼셉트론은 비교적 높은 인식률과 빠른 인식 속도를 자랑하지만, 신뢰할 만한 인식 후보 및 점수를 출력하지 않으므로 문맥정보 및 사전정보를 활용하는 주소열의 후처리가 불가능하다는 단점을 갖는다. 반면 최소거리분류기는 통계적 이론에 기반한 단순한 알고리즘으로 낱자에 대한 출력값이 확률로 주어지므로 후처리에 효과적이고 빠른 처리 속도를 가지므로 고속 주소 인식 시스템에 적합하지만, 다른 인식기에 비해 인식률이 낮다는 단점을 갖는다. 마지막으로 Subspace 방법은 최소거리분류기와 마찬가지로 통계적 이론에 기반한 모델링 기법으로 낱자에 대한 출력값이 확률로 주어지고 본 연구에서 PE92 데이터베이스를 사용하여 실험한 결과 다층퍼셉트론보다 더 높은 인식률을 보유함이 관측되었지만, 인식 속도가 매우 늦어 고속 주소 인식에 부적합하다.

본 연구에서는 언급한 인식기들의 장단점을 상호 보완할 수 있는 2 단계 인식 방법을 구상하였다. 즉, 빠른 인식 속도를 갖는 인식기를 사용하여 1 차 인식을 수행한 후, 1 차 인식 결과를 기반으로 상위 N 개의 인식 후보를 2 차 인식기의 인식 대상으로 제한하였다. 따라서 2 차 인식기는 인식 대상이 변화하는 것에 능동적으로 적응할 수 있어야 하며, 빠른 처리 보다는 정확한 인식을 수행해야 한다. 이러한 조건을 감안하여 세 인식기 중 2 차 인식기로 Subspace 방법을 선택하였으며, 1 차 인식기는 다층퍼셉트론과 최소거리분류기를 모두 고려하여 각각의 방법을 비교하였다. 이에 대한 실험 결과는 3 절에서 자세히 설명될 것이다.

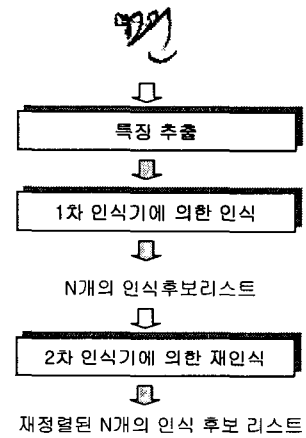

그림 2. 필기 한글 낱자 인식 구조

마지막으로, 세 인식기에서 공통으로 사용하는 낱자 영상에 대한 특징은 비선형 방향 성분 특징[5] 이다. 본 연구에서는 필기 낱자 영상의 가로 대 세로 비율이 7 대 9 라는 통계치를 반영하여 252(=7*9*4)차원의 특징벡터를 사용하였다.

2.1 다층퍼셉트론

채택된 다층퍼셉트론(Multi Layer Perceptrons: MLP)은 3 개의 층으로 구성되었으며, 입력층은 입력 낱자 영상으로부터 추출된 252 차원의 방향 성분 특징이 입력되며, 은닉층의 노드의 개수는 200 개, 출력층의 노드의 개수는 인식 대상 낱자 개수와 같은 469 개이다. 즉, 출력층의 각 노드는 하나의 낱자와 대응된다. 또한 입력층과 은닉층에는 편기항(bias) 노드가 하나 추가되었으며, 은닉층과 출력층의 활성화 함수로는 시그모이드 함수를 사용하였다. 각 노드 사이의 연결강도를 훈련하기 위해 목표값과 출력값의 차이에 대한 평균 제곱을 최소화하도록 조정하였으며, 이때 평균 제곱에 대한 최소값을 찾기 위해 오류역전파 학습 알고리즘[7]을 사용하였다.

2.2 최소거리분류기

최소거리분류기는(MDC: Minimum Distance Classifier)는 동일한 특징 공간상에 표현되는 각 낱자의 대표벡터와 입력 낱자의 특징벡터 간의 거리를 계산한 후, 가장 작은 거리를 갖는 대표벡터가 나타내는 낱자로 입력 낱자의 소속 클래스를 결정하는 통계적인 인식 방법론이다[8]. 본 연구에서는 각 낱자의 대표벡터를 훈련벡터집합의 평균벡터로 정의하여 사용하였고, 두 벡터 간의 거리를 측정하기 위해서 맨하탄 거리(Manhattan distance) 함수를 사용하였다. 즉, 입력 낱자의 특징벡터가 주어지면, 각 낱자의 평균벡터와 맨하탄 거리를 계산한 후, 가장 작은 거리를 갖는 낱자로 입력의 소속 클래스를 결정한다.

2.3 Subspace 방법

Subspace 방법은 입력 낱자를 인식 대상의 각 낱자에 대한 subspace 모델로 복원했을 때 발생하는 복원 오차를 서로 비교하여 최소의 복원 오차를 갖는 낱자로 입력 낱자의 소속 클래스를 결정하는 통계적인 인식 방법론이다[9]. 이때 각 낱자의 subspace 모델은 각 낱자의 훈련벡터집합의 평균벡터 μ와 공분산 행렬의 고유벡터로 정의되는 변환행렬 $W_{N\times M}= [\varphi_1,\cdots, \varphi_M]$로 정의된다. 이때 φ_i는 i 번째 크기의 고유값 λ_i에 대응하는 고유벡터이다. 복원 오차는 N 차원 특징벡터 x를 이보다 작은 M 차원 subspace 상의 y로 변환한 후, 다시 y를 N 차원 특징공간으로 복원한 후 얻어지는 x'와 x의 차이이다.

M 차원 subspace 공간으로 변환: $y = W^T(x-\mu)$
N 차원 특징공간으로 복원: $x' = Wy + \mu$

Subspace 방법에서 변환행렬을 몇 개(M)의 고유벡터로 구성하는가에 따라 복원 오차가 달라지며, 이에 따라 전체 성능에서 차이가 발생한다. M 의 값이 커지면 인식 속도가 낮아지며, 작은 M 의 값은 인식률에 영향을 미친다. 따라서 M 은 실험에 의하여 최적의 값으로 학습되어야 한다. 본 연구에서는 실험을 통하여 M 을 30 으로 결정하였다.

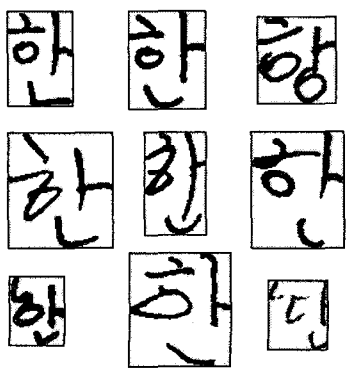

그림 3. PE92 데이터베이스에 있는 낱자의 예

3. 실험 및 결과

본 연구에서 채택한 필기 낱자 인식 알고리즘은 펜티엄 1.3GHz PC 상에서 PE92 데이터베이스[10]를 사용하여 평가되었다. PE92 데이터베이스는 완성형 한글 2,350 종에 대하여 사람들이 필기한 100 세트의 낱자집합으로 되어있다. 이 중 그림 1 에 제시한 469 종의 낱자에 대하여 상위 68 세트를 학습용으로 사용하였으며, 학습용으로 사용되지 않은 하위 29 세트를 성능분석을 위하여 사용하였다. 그림 3 에는 PE92 데이터베이스에 있는 낱자 '한'에 대한 몇 가지 예를 보여주고 있는데, 여기서 볼 수 있듯이 잘못 분류된 낱자뿐만 아니라 영상 자체가 매우 희미하여 사람도 판별하기 어려운 낱자가 다수 존재한다.

PE92 데이터베이스 상위 68 세트를 사용하여 학습한 내용은 다음과 같다. 다층퍼셉트론의 경우 입력층과 은닉층 그리고 은닉층과 출력층을 연결하는 연결강도를 오류 역전파 학습 알고리즘을 사용하여 학습하였다. 이때 학습률 α 를 0.7, 모멘텀항 ε 를 0.1 로 설정하고 실험한 결과 반복회수가 194 일때 최적의 인식 결과를 출력하였다. 최소거리분류기의 경우 각 낱자의 대표벡터를 구하기 위하여 68 개의 학습패턴으로부터 평균벡터를 구하였으며, Subspace 방법의 경우는 최소거리분류기에서 구한 평균벡터뿐만 아니라 변환행렬도 계산하였다. 이렇게 학습된 각 방법에 대하여 하위 29 세트를 사용하여 평가한 결과가 표 1, 2, 3 에 각각 제시되어 있다.

표 1. 다층퍼셉트론 인식 성능

누적순위	1-순위	2-순위	5-순위	10-순위
인식률(%)	72.5	83.4	91.8	95.2

처리속도: 1.22ms/낱자

표 2. 최소거리분류기 인식 성능

누적순위	1-순위	2-순위	5-순위	10-순위
인식률(%)	64.2	78.0	88.8	93.8

처리속도: 1.65ms/낱자

표 3. Subspace 방법 인식 성능

누적순위	1-순위	2-순위	5-순위	10-순위
인식률(%)	74.7	85.6	93.5	96.3

처리속도: 45ms/낱자

각 낱자 인식기의 성능 테스트 결과에서 알 수 있듯이, 다층퍼셉트론은 인식률과 처리속도가 우수하지만 후처리에 적합한 값을 출력하지 않는다는 단점을 가지고 있다. 또한 최소거리분류기와 Subspace 방법은 통계적 인식 기법으로 모두 후처리에 적합한 인식기이지만, 최소거리분류기는 인식률이 낮고 Subspace 방법은 처리속도가 늦다는 단점을 보유하고 있다. 따라서 제안방법에서 언급했듯이, 다층퍼셉트론 또는 최소거리분류기를 사용하여 인식 대상을 줄인 후, Subspace 방법을 사용하여 정확하게 인식하는 전략을 세웠다. 그 결과가 그림 4 와 5 에 제시되어 있다.

그림 4 는 다층퍼셉트론를 사용하여 469 자에 대하여 인식한 결과 얻어지는 상위 10 개의 낱자들에 대하여 다시 Subspace 방법을 적용하여 재인식하여 얻어진 결과를 보여준다. 이러한 방법의 적용결과 다층퍼셉트론만을 사용하였을 때보다 인식률이 향상되었고, 처리속도는 약 2.18ms = 1.22ms + 0.96ms(=45ms/469*10)정도 소요되었다. 그림 5 는 최소거리분류기를 사용하여 469 자에 대하여 인식한 결과 얻어지는 상위 10 개의 낱자들에 대하여 다시 Subspace 방법을 적용하여 재인식하여 얻어진 결과를 보여준다. 이렇게 적용한 결과 최소거리분류기만을 사용하였을 때보다 1 순위 인식률이 64.2%에서 76.5%로 대폭 상승 하였으며, 처리속도 약 2.61ms = 1.65ms + 0.96ms(=45ms/469*10)정도 소요되었다.

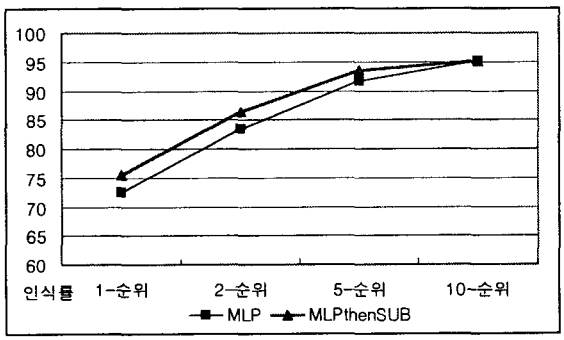

그림 4. 2-단계 인식 성능: MLP, Subspace 방법

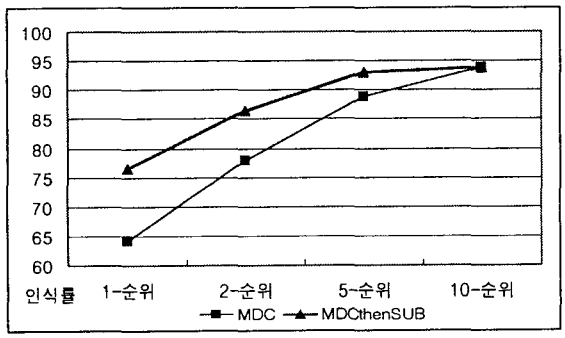

그림 5. 2-단계 인식 성능: MDC, Subspace 방법

4. 결론

본 논문은 고속 필기 한글 주소 인식을 위한 낱자 인식기의 연구 결과이다. 주소 인식을 위한 낱자 인식기 개발이 목적이므로, 주소에 빈번히 출현하는 469 자로 인식 대상을 줄임으로써 문제를 축소하였다. 또한 높은 인식률, 빠른 처리속도, 그리고 후처리에 적합한 인식 결과의 출력을 낱자 인식기가 필히 갖추어야 할 요건으로 설정하였다.

본 연구에서는 다층퍼셉트론, 최소거리분류기, Subspace 방법을 낱자 인식기로 선정하여 각 방법의 성능을 비교 분석하였다. 그 결과 각 방법은 앞에서 언급한 세 요건 중 두 가지는 만족하였으나, 하나씩은 만족하지 못함을 실험을 통하여 알 수 있었다. 따라서 서로의 장점을 유지하면서, 서로의 단점을 보완할 수 있도록 인식기들의 결합을 시도하였다. 제안 방법은 빠른 낱자 인식기 - 다층퍼셉트론, 최소거리분류기 - 를 사용하여 선 인식을 수행한 후, 처리속도는 늦으나 비교적 정확하게 인식하는 Subspace 방법을 사용하여 빠른 낱자 인식기의 상위 인식 후보들에 대하여 재인식을 시도하는 것이다. 실험결과 이러한 인식기 결합 방법이 위의 세 요건을 만족시킴을 알 수 있었다.

참고문헌

[1] 이성환, 박희선, "한글 인식의 사례 연구: 최근 5 년간의 연구 결과를 중심으로," 제 1 회 문자인식워크샵 발표논문집, pp.3-46, 1993.

[2] 김수형, 정선화, 오일석, "필기 한글 문자의 오프라인 인식에 관한 사례 연구," 1998 년도 추계 정보과학회 학술발표논문집, 제 25 권 2 호, pp.396-398, 1998.

[3] H.Y. Kim and J.H. Kim, "Handwritten Korean character recognition based on hierarchical random graph modeling," Proc. 6th IWFHR, pp.577-586, 1998.

[4] Y.S. Hwang and S.Y. Bang, "Recognition of a handwritten Korean character by combining segments using constraint satisfying graph," Proc. 6th IWFHR, pp.515-526, 1998.

[5] 김수형, "대용량 필기 문자인식을 위한 최소 거리 분류법의 성능 개선 전략," 정보처리논문지, 제 5 권 제 9 호, pp.2600-2608, 1998.

[6] H.S. Park and S.W. Lee, "Off-line recognition of large-set handwritten characters with multiple hidden Markov models," Pattern Recognition, vol.29, no.2, pp.231-244, 1996.

[7] D.E. Rumelhart, G.E. Hitton and R.J. Williams, "Learning representations by back-propagating error," Nature, vol.332, pp.533-536, 1986.

[8] R.O. Duda and P.E. Hart, Pattern classification and scene analysis, Wiley-Interscience Pub., 1973.

[9] E. Oja, Subspace methods of pattern recognition, Research Studies Press LTD., 1983.

[10] 김대환, 방승양, "한글 필기체 영상 데이터베이스 PE92 의 소개," 제 4 회 한글 및 한국어 정보처리 학술발표논문집, pp.567-575, 1992.