

전표 금액란에 나타나는 필기 한글의 신경망-기반 인식 (Neural Network-based Recognition of Handwritten Hangul Characters in Form's Monetary Fields)

이진선* 오일석**

(Jin-Seon Lee) (Il-Seok Oh)

요약 한글은 부류수의 방대성과 글자간의 유사성으로 인해 인식이 어려운 문자 집합으로 간주되고 있다. 기존 연구 대부분은 일반적으로 사용되는 2,350 글자를 대상으로 인식을 시도하였는데, 이는 일반성을 제공하는 대신 낮은 성능 문제를 안고 있다. 이에 반해, 우편 영상이나 전표 영상 등의 특정 필드에 나타나는 한글만을 대상으로 하는 접근 방법이 보다 현실적이라 할 수 있다. 본 논문은 금액란에 나타나는 필기 한글을 인식하는 연구를 기술한다. 인식을 위해 모듈러 신경망 인식기를 사용하였으며, 세 종류의 특징을 사용하였다. 표준 한글 데이터베이스 PE92에 대해 실험한 결과 정인식률 97.56%를 얻었다.

Abstract Hangul is regarded as one of the difficult character set due to the large number of classes and the shape similarity among different characters. Most of the conventional researches attempted to recognize the 2,350 characters which are popularly used, but this approach has a problem of low recognition performance while it provides a generality. On the contrary, recognition of a small character set appearing in specific fields like postal address or bank checks is more practical approach. This paper describes a research for recognizing the handwritten Hangul characters appearing in monetary fields. The modular neural network is adopted for the classification and three kinds of feature are tested. The experiment performed using standard Hangul database PE92 showed the correct recognition rate 97.56%

1. 서론

한글 인식은 우편물 분류, 전표 인식, 각종 양식 인식 등 여러 분야에서 필수적으로 요구되는 분야로서, 그동안 많은 연구가 수행되어 왔다[1-11]. 이들에 대한 사례 조사는 [12, 13]에서 찾아볼 수 있다. 지금까지 수행된 이들 연구들은 한글 전체를 대상으로 한 연구로서, 512 또는 2,350 글자를 대상으로 하는 경우가 대부분이다.

한글은 최대 11,172 글자, 일상 생활에서 사용되는 2,350 글자를 가지므로 부류 수가 방대한 특성을 가지고 있다. 또한 작은 획 하나의 존재 여부에 따라 서로 다른 글자가 되는 글자 간 높은 유사성이라는 특성을 가지고 있다. 이 두 가지 특성으로 인해 인식이 어려운 문자 집합으로 간주되고 있다. 따라서 전체 한글을 대상으로 수행한

연구들은 성능 면에서 한계를 가질 수밖에 없다. 결국 특정 분야에 적합한 인식기에 대한 연구가 성능 면이나 비용 면에서 현실적으로 매우 필요하다. 예를 들어 한글 금액란에는 18글자만 나타나는데, 2,350 글자를 대상으로 하는 인식기를 한글 금액 인식기로 채택할 이유가 없다.

본 논문은 우리나라에서 사용하는 각종 양식에 포함된 금액란에 나타나는 필기 한글만을 대상으로 인식하는 연구로서, 연구 결과는 은행 전표 인식 등에서 유용하게 사용할 수 있다. 유사한 연구는 [14]에서 찾아볼 수 있는데, 이 논문은 성능 평가를 위해 자체 수집한 데이터베이스를 사용하였다.

본 논문은 [15]에서 제안한 모듈러 신경망 인식기를 사용하였다. 모듈러 신경망은 기존 신경망에 비해 필기 숫자와 영문자[16], 그리고 우편 주소에 나타나는 352개 한글 글자[15]에 대해 우수한 성능이 입증되었다. 특징 벡터 추출 방법으로는 AGD(Area-based Gradient Distribution,

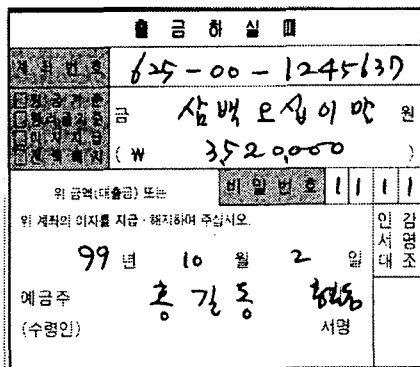
* 우석대학교 정보통신컴퓨터공학부
** 전북대학교 컴퓨터과학과

영역-기반 경사 분포)[17], DDD(Directional Distance Distribution, 방향 거리 분포)[16], 그리고 DSM(Directional-Segment Mesh, 방향-성분 메쉬)[5]의 세 종류를 사용하고 그들 간의 성능을 비교하였다. 실험 데이터로는 PE2 데이터베이스에서 금액란에 나타나는 글자 샘플들을 추출하여 사용하였다. 실험 결과 DSM 특징이 세 가지 특징 중에서 가장 우수한 정인식률 97.56%를 보였다.

2장에서는 본 논문의 문제에 대해 자세히 기술한다. 3장에서는 인식에 사용한 세 종류의 특징과 모듈러 신경망 인식기에 대해 기술한다. 4장에서는 인식 실험을 위한 환경과 결과를 기술하고, 5장에서 결론을 제시한다.

2. 문제 기술

우리 나라에서 사용하고 있는 각종 양식에는 금액란이 있다. <그림 1>은 은행 전표의 한 예제 영상이다. 한글 금액란에 나타나는 글자들은 '일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십, 백, 천, 만, 억, 조, 원, 정, 금'의 총 18 글자이다 [14]. 현실적으로 각종 은행에서 사용하는 전표, 수표 등의 양식에서 금액란에 필기되는 글자에는 '원, 정, 금'이 포함되어 있지 않다. 본 논문에서는 이러한 사실을 고려하여 18 글자와 15 글자에 대한 두 가지 인식 실험을 수행한다.



<그림 1> 은행 전표 영상.

금액란에 있는 정보는 돈과 관련되어 있으므로 높은 인식률과 높은 신뢰도를 요구한다. 따라서 기존 서양 수표 인식기들은 숫자로 기입된 란과 영어로 기입된 란을 모두 인식하여 서로를 검증하는 방법을 사용하여 높은 신뢰도를 얻고 있다[18]. 은행 전표를 자동 인식하는 시스템을 구축

하기 위해서는 한글로 기입된 금액란과 숫자로 기입된 금액란을 추출하고, 각 란에 있는 글자들을 낱자 단위로 분할하는 모듈을 만들어야하는데, 이러한 작업을 위한 논문은 [19, 20] 등에서 찾아볼 수 있다.

본 논문이 해결하고자 하는 문제는 이미 금액란 분할과 낱자 분할이 된 상태에서, 필기된 18 글자 또는 15 글자를 대상으로 고성능 인식하는 것이다. 문자 인식 연구에서는 인식 성능을 다른 연구와 객관적으로 비교하는 것이 중요하므로, 표준 필기 한글 데이터베이스 PE92를 사용한다.

3. 인식 방법

3.1 특징 추출

이 절에서는 실험에 사용할 세 가지 종류의 특징에 대해 기술한다. 보다 자세한 알고리즘은 표시된 참고 문헌을 참고하기 바란다.

(1) AGD (Area-based Gradient Distribution, 영역-기반 경사 분포) 특징[17]

입력 패턴 $P_{m \times n}$ 은 먼저 16×16 메쉬, $R_{16 \times 16}$ 로 크기 정규화된다. 크기-정규화된 메쉬 R 에 Sobel 에지 연산자를 적용하여, 경사 강도 맵과 경사 방향 맵을 얻는다. 0° 와 360° 사이 값을 갖는 경사 방향은 16단계 (22.5° /단계)로 구간화한다. 경사 방향 맵 $D_{16 \times 16}$ 을 4×4 블록 메쉬로 나눈다. 즉, D 의 4×4 부 영역이 하나의 블록이 된다. 각 블록에 대해 16개 구간을 갖는 하나의 히스토그램을 계산한다. 히스토그램 누적 시 경사 크기를 가중치로 사용한다. 16개 블록 모두에 대해 히스토그램 계산을 마친 후, 히스토그램 값을 0.0과 1.0사이의 값으로 정규화한다.

(2) DDD (Directional Distance Distribution, 방향 거리 분포) 특징[16]

이 특징은 크기-정규화된 메쉬 R 로부터 거리정보에 기반하여 계산한다. R 의 흑 (백) 화소로부터 8 방향으로 광선을 쏘아 백 (흑) 화소까지의 거리를 계산한다. R 의 모든 화소에 대해 방향-거리 정보를 계산한 후, 방향-거리 분포를 계산한다. 이를 위해서는 16×16 메쉬를 4×4 블록 메쉬로 나눈다. 즉, 4×4 부 영역이 하나의 블록이 된다. 각 블록에 대해 방향별로 평균값을 계산하여 특징 값을 추출한다. 모든 값을 0.0과 1.0사이의 값으로 정규화한다.

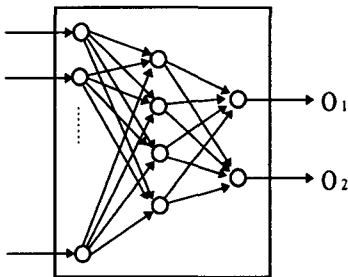
(3) DSM (Directional-Segment Mesh, 방향-성분 메쉬) 특징[5]

필기 한글의 세로 대 가로 비율(aspect ratio)의 통계치

를 고려하여 $9 \times 7 \times 4$ 차원의 벡터를 채택한다. 즉, 입력 문자에 대한 9×7 차원의 비선형 메쉬(nonlinear mesh)와 수평(H), 수직(V), 사선(L), 역사선(R) 방향의 성분을 구한 후, 9×7 차원 메쉬 상의 63개 셀(cell) 각각으로부터 이틀네 가지 방향 성분에 대한 특징 값을 계산하여 $9 \times 7 \times 4 (=252)$ 차원의 벡터를 형성한다.

3.2 모듈러 신경망 인식기

인식기로는 [15]에서 제안한 모듈러 신경망 아키텍처를 사용한다. 각각 ω_k 로 부르는 g 개의 부류가 있다고 가정하자. 금액란에 대한 한글 인식 문제의 경우는 g 는 18 또는 15이다. 모듈러 신경망은 g 개의 부 망으로 구성되는데, 각 부 망은 g 개 부류중 하나씩을 책임진다. 부류 ω_k 를 위한 부 망을 <그림 2>가 보여 준다. 부 망 하나의 입력층, 하나의 은닉층, 하나의 출력층을 갖는다. 이들 세 층은 완전 연결되어 있다. 이 부 망의 기능은 두 개 부류 군, $\Omega_1 = \{\omega_k\}$ 과 $\Omega_2 = \{\omega_k | 1 \leq k \leq g, k \neq i\}$ 를 분류하는 것이다. 입력층은 n -차원 특징 벡터를 받기 위해 n 개의 노드를 갖는다. 은닉층은 4개의 노드를 갖는다. 출력층은 Ω_1 과 Ω_2 를 위한 두 개의 노드를 갖고 이들을 각각 O_1 과 O_2 라 부른다.



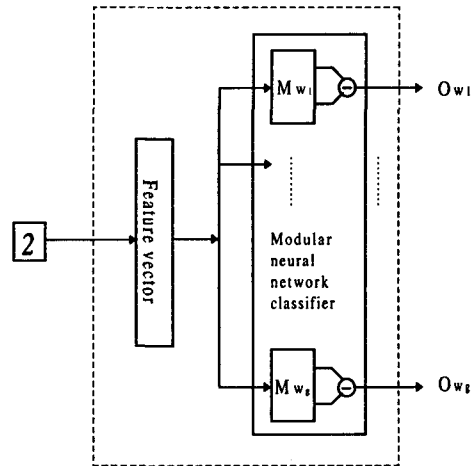
<그림 2> 모듈러 신경망을 위한 부망.

g 개의 부 망의 훈련은 각각 독립적으로 에러 후향(error backpropagation) 알고리즘으로 수행한다. 부류 ω_k 를 훈련하기 위해, 훈련 집합을 두 개의 군, Z_{positive} 와 Z_{negative} 로 나눈다. Z_{positive} 는 Ω_1 에 속한 샘플, Z_{negative} 는 Ω_2 에 속한 샘플을 갖는다.

Z_{positive} 의 샘플은 기대 출력값으로 $(O_1, O_2) = (1.0, 0.0)$ 와 함께 입력되며, Z_{negative} 샘플은 $(O_1, O_2) = (0.0, 1.0)$ 와 함께 입력된다. 초기 가중치는 0.0과 0.2 사이의 임의수로 설정한다. 훈련 샘플의 순서는 새로운 세대를 시작하기 전에 임의로 섞는다. 훈련 계수는 세대가 증가할수록 0.5에서

0.001값으로 선형적으로 줄여 나간다.

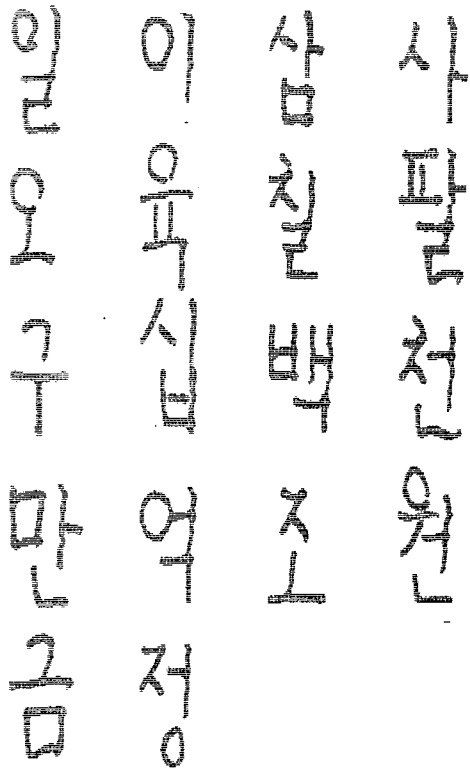
부류 ω_k 는 자신의 부 망으로서 M_{ω_k} 를 가지며, 전체 분류기의 아키텍처는 <그림 3>과 같다. 입력 패턴으로부터 특징 벡터를 추출하고, 이를 모든 부 망에 입력한다. 부 망 M_{ω_k} 는 O_1 과 O_2 의 값을 뺄으로써 단일 값을 출력한다. 최종적으로 입력 패턴은 최고 출력값을 갖는 부류로 분류한다.



<그림 3> 전체 분류기 아키텍처.

4. 인식 실험

인식 실험을 위해 PE92 데이터베이스를 사용하였다[21]. PE92는 2,350 글자를 포함하고 있으며 글자 별로 100개의 필기 샘플을 갖는다. 그림 4는 PE92에 들어 있는 필기 한글의 예를 보여준다. 이 데이터베이스로부터 금액란에 나타나는 18 글자에 해당하는 샘플을 추출하여 실험하였다. 100개 샘플 중 70개는 신경망 훈련(training)에 사용하고 나머지 30개는 검사(testing)에 사용하였다.



<그림 4> PE92 데이터베이스에서 추출한 한글 샘플

<표 1>과 <표 2>는 각각 18글자와 15 글자에 대한 정인식률을 보여준다. 정인식률은 기각을 허용하지 않은 상황에서 (옳게 인식한 샘플 수/전체 샘플 수)로 계산하였다. <표 3>은 15 글자를 DSM 특징으로 인식한 경우 출력된 혼돈행렬(confusio

<표 1> 18 글자에 대한 정인식률

	훈련 집합	검사 집합
AGD	97.50%	82.80%
DDD	97.02%	78.32%
DSM	99.84%	95.37%

<표 2> 15 글자에 대한 정인식률

	훈련 집합	검사 집합
AGD	99.61%	87.11%
DDD	97.78%	84.22%
DSM	99.90%	97.56%

실험 결과 DSM 특징이 18 글자와 15 글자 각각에 대해 95.37%와 97.56%로서, AGD와 DDD에 비해 월등히 우수한 인식률을 보였다. 이 성능은 실용화에 미치지 못하지만 근접했다고 볼 수 있다. 실용화하기 위해서는 99%이상의 정인식률이 필요할 것으로 보인다. AGD와 DDD는 80%대에 머물어 실용화에는 큰 거리가 있음을 알 수 있다.

현재 실험은 PE92 데이터베이스로 수행하였는데, 한가지 문제는 글자 별로 샘플이 100개 밖에 없다는 것이다. 이 중 70개를 훈련에 사용하였는데, 이는 아주 부족한 양이다. 보다 다양한 환경에서 보다 다양한 필기자로부터 획득한 충분한 양의 훈련 집합을 사용하면 인식률을 개선할 수 있을 것이다. 또한 현재까지는 특징 추출, 신경망 인식 기 구조와 훈련에 관련된 파라미터들에 대한 최적 값을 결정하지 못한 상태에서 실험을 수행하였다. 은닉 노드 개수 등과 같은 신경망 구조에 관련된 파라미터, 가중치 초기

<표 3> 15 글자를 DSM 특징으로 인식한 결과 혼돈 행렬

	일	이	삼	사	오	육	칠	팔	구	십	백	천	만	억	조	(정인식,오인식,기각)
일	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(30, 0, 0)
이	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	(30, 0, 0)
삼	0	0	29	0	0	1	0	0	0	0	0	0	0	0	0	(29, 1, 0)
사	0	0	0	28	0	0	0	0	0	2	0	0	0	0	0	(28, 2, 0)
오	0	0	0	0	28	2	0	0	0	0	0	0	0	0	0	(28, 2, 0)
육	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	(30, 0, 0)
칠	0	0	0	0	0	0	29	0	1	0	0	0	0	0	0	(29, 1, 0)
팔	0	0	0	0	0	0	0	29	0	0	0	0	0	0	1	(29, 1, 0)
구	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	(30, 0, 0)
십	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	(30, 0, 0)
백	0	0	0	0	0	0	1	0	0	0	28	0	0	0	1	(28, 2, 0)
천	1	0	0	0	0	0	0	0	0	0	0	28	1	0	0	(28, 2, 0)
만	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	(30, 0, 0)
억	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	(30, 0, 0)
조	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	(30, 0, 0)

(439, 11, 0) = (97.56%, 2.44%, 0%)

값이나 훈련 계수와 같은 신경망 훈련에 관련한 계수, 그리고 대상 한글 글자에 적합한 특징 개발 등의 작업을 통해 인식률을 실용화 수준까지 올릴 수 있을 것이다.

5. 결 론

본 논문은 한글 금액란에 나타나는 18 또는 15 글자를 대상으로 하는 인식기에 대한 연구 결과를 기술하였다. 인식기로는 우수한 성능이 입증된 모듈러 신경망을 사용하였으며, 특징으로는 세 종류를 사용하여 그들을 비교하였다. DSM 특징의 경우 15 글자에 대해 97.56%의 정인식률을 보여 실용화에 근접함을 보였다. 실용화 시스템을 위한 향후 연구로서 신경망 구조 및 훈련 계수 최적 튜닝, 보다 적절한 특징 개발 등을 수행할 계획이다. 이러한 인식률 개선을 통해 현재 국내 금융 기관을 중심으로 개발하려는 자동 전표 입력 시스템에 효과적으로 사용할 수 있다.

참고 문헌

- [1] H.Y. Kim and J.H. Kim, "Handwritten Korean Character Recognition Based on Hierarchical Random Graph Modeling," Proceedings of 6th IWFHR, pp.577-586, 1998.
- [2] 황영섭, "제약조건 만족 그래프를 이용한 필기 한글 인식," 박사학위논문, 포항공과대학, 1997.
- [3] 김은진, "인간의 한글 지각 처리에 근거한 한글 인식 방법에 관한 연구," 박사학위논문, 연세대학교, 1997.
- [4] 김수형, 도정인, 김준호, 이상규, "통계적 군집화 및 신경망 기법을 이용한 필기체 문자의 오프라인 인식," 제2회 문자인식워크샵 발표논문집, 서울, pp.133-142, 1994.
- [5] 김수형, "대용량 필기 문자인식을 위한 최소거리 분류법의 성능 개선 전략," 한국정보처리학회 논문지, 제5권 제10호, 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, 29(2), pp.231-244, 1996.
- [7] 박희선, "은닉 마르코프 메쉬 랜덤 필드: 이론 및 필기체 문자 인식에의 응용," 박사학위논문, 충북대학교, 1996.
- [8] W.S. Kim and R.H. Park, "Off-line Recognition of Handwritten Korean and Alphanumeric Characters using Hidden Markov Models, Pattern Recognition, 29(5), pp.845-858, 1996.
- [9] H. Kwak and K. Chung, "Recovery of Temporal Information from Static Images of Hangeul Handwriting Based on Its Structural Information," Proc. 6th IWFHR, pp.527-536, 1998.
- [10] H.J. Kim and P.K. Kim, "Recognition of Off-line Handwritten Korean Characters," Pattern Recognition, 29(2), pp.245-254, 1996.
- [11] 감민오, 조형제, "속성 그래프를 이용한 오프라인 필기체 한글 인식," 한국정보과학회 1997년 봄 학술 발표논문집(B), pp. 525-528, 1997.
- [12] 이성환, 박희선, "한글 인식의 사례 연구: 최근 5년간의 연구 결과를 중심으로," 제1회 문자인식워크샵 발표논문집, pp.3-46, 1993.
- [13] 김수형, 정선화, 오일석, "필기 한글 문자의 오프라인 인식에 관한 사례 연구," 한국정보과학회 1998년 봄 학술발표논문집 (B), pp.396-398, 1998.
- [14] Y. Baek, K. Lim, S. Chien and J. Park, "Off-line Handwritten Hangeul Recognition Based on Multiple Features and Modular Partially Connected Multi-layer Perceptron," Proceedings of the Sixth IWFHR'98, pp.269-278, 1998.
- [15] Il-Seok Oh and C.Y. Suen, "A Class-modular Feedforward Neural Network for Handwriting Recognition," IEEE Tr. on Neural Networks (Submitted).
- [16] 오일석, Ching Y. Suen, "광학 문자 인식을 위한 거리 특징," 한국정보과학회 논문지 (B), 제25권 제7호, pp.1028-1043, 1998.
- [17] G. Srikanth, S.W. Lam, and S.N. Srihari,

"Gradient-based Contour Encoding for Character Recognition," Pattern Recognition, Vol.29, No.7, pp.1147-1160, 1996.

[18] C.Y. Suen, K. Liu, and N. W. Strathy. "Sorting and Recognizing Cheques and Financial Documents," Proceedings of 3rd IAPR Workshop on Document Analysis Systems, pp. 1-18, Nagano, Japan, November 1998.

[19] 변영철, 윤성수, 김경환, 최영우, 이일병, "부분 매칭을 이용한 서식 이해에 관한 연구," 한국정보과학회 1999년 가을 학술 발표 논문집(III), pp. 443-445, 1999.

[20] 황순자, 김문현, "접촉 유형에 기반한 오프라인 필기체 한글 문자 분할," 한국인지과학회 논문지, 제6권 제4호, pp.51-71, 1995.

[21] 김대환, 방승양, "한글 필기체 영상 데이터베이스 PE92의 소개," 제4회 한글 및 한국어 정보처리 학술발표 논문집, pp.567-575, 1992.



이진선

1985년 전북대학교 전산통계학과 (이학사)
1988년 전북대학교 전산통계학과 (이학석사)
1995년 전북대학교 전자계산기공학과 (공학박사)
1988년-1992년 한국전자통신연구원 근무

1995년-현재 우석대학교 정보통신컴퓨터공학부 조교수
관심분야 영상처리, 패턴인식, 멀티미디어, GIS



오일석

1984년 서울대학교 컴퓨터공학과 학사
1992년 한국과학기술원 전산학과 박사
1992년 - 현재 전북대학교 컴퓨터공학과 부교수

관심분야 패턴인식, 컴퓨터비전, 멀티미디어