

다중 신경망을 이용한 인식단위 결합 기반의 인쇄체 문자인식

임 길 택[†]·김 호 연[†]·남 윤 석^{††}

요 약

본 논문에서는 다중 신경망을 이용한 인식단위 결합 기반의 인쇄체 문자인식 방법을 제안한다. 입력 문자영상은 한글 문자 형식 6가지와 한글 이외의 기타 문자 형식의 전체 7가지 형식으로 분류되어 인식된다. 한글 문자는 2단계의 MLP 신경망 인식기에 의해 인식된다. 첫째 단계에서는 한글 문자를 자소의 조합 형태에 따라 2개 또는 3개의 인식단위로 나누고, 각 인식단위에서 추출된 방향각도 특징 벡터를 입력으로 하는 MLP 신경망으로 1차 인식한다. 둘째 단계에서는 첫째 단계의 인식단위별 MLP 신경망 인식기의 인식양상 특징을 추출하고 다른 MLP 신경망에 입력하여 최종 한글 문자인식을 한다. 한글 이외의 기타 문자의 인식을 위해서는 단일 MLP 신경망을 사용한다. 인식 실험에서는 실제 우편물 50,000통 영상으로부터 추출한 문자영상 데이터베이스를 이용하였는데, 실험 결과 본 논문에서 제안한 방법이 매우 우수함을 알 수 있었다.

Machine Printed Character Recognition Based on the Combination of Recognition Units Using Multiple Neural Networks

Kil-Taek Lim[†] · Ho-Yon Kim[†] · Yun-Seok Nam^{††}

ABSTRACT

In this paper, we propose a recognition method of machine printed characters based on the combination of recognition units using multiple neural networks. In our recognition method, the input character is classified into one of 7 character types among which the first 6 types are for Hangeul character and the last type is for non-Hangeul characters. Hangeul characters are recognized by several MLP (multilayer perceptron) neural networks through two stages. In the first stage, we divide Hangeul character image into two or three recognition units (HRU : Hangeul recognition unit) according to the combination fashion of graphemes. Each recognition unit composed of one or two graphemes is recognized by an MLP neural network with an input feature vector of pixel direction angles. In the second stage, the recognition aspect features of the HRU MLP recognizers in the first stage are extracted and forwarded to a subsequent MLP by which final recognition result is obtained. For the recognition of non-Hangeul characters, a single MLP is employed. The recognition experiments had been performed on the character image database collected from 50,000 real letter envelope images. The experimental results have demonstrated the superiority of the proposed method.

키워드 : 문자인식(Character Recognition), 신경망(Neural Networks), 인식단위(Recognition Unit), 형식분류(Type Recognition)

1. 서 론

지난 20여 년간 한글 문자인식에 관한 많은 연구가 있었다[1-12]. 초기의 대부분의 연구는 제한된 크기와 글꼴의 인쇄체 문자인식에 관한 연구였으며, 1990년대 초반에는 다중 활자체에 관한 문자인식 연구가 진행되었고, 1990년대 중반부터는 상용 인쇄체 문자인식 시스템이 나오기 시작했다. 최근의 문자인식 연구는 사람에 의해 쓰여진 필기체 문자인식[11-14]이 주류를 이루고 있고 인쇄체 문자인식에 관한 연구는 관심에서 많이 멀어져 있다.

그러나 기존의 인쇄체 한글 문자인식 방법들 및 현재의 상용 문자인식 시스템은 문자영상 획득 조건과 문자의 종

류에 민감하여 저품질 및 저해상도의 문자에서는 인식률이 안정적이지 못하다. 이러한 문제점은 한글이 가진 특성에서 비롯된다. 한글은 19개의 초성과 21개의 중성 및 27개의 종성들이 2차원적 조합으로 되어 문자가 구성되기 때문에 인식대상이 11,172개가 될 정도로 매우 많다. 사용빈도가 높은 문자만을 대상으로 인식하더라도, 1,000~2,000자가 되므로 인식시스템의 부담이 매우 커진다. 또한 한글 문자는 문자를 구성하는 자소 간의 구분이 작은 획 하나에 의해 결정될 정도로 유사성이 높아 영문이나 숫자에 비해 인식이 어렵다. 이러한 이유로 대부분의 한글 문자인식 시스템이 300~400dpi(dots per inch)이상의 해상도를 기준으로 구현되고 있다. 그러나, 높은 해상도에도 불구하고 문자 글꼴이 다양하거나 문자영상의 품질이 좋지 않은 경우와 여러 가지 제약조건으로 인해 고해상도의 영상입력이 불가능한 우편물 자동인식, FAX 문서인식의 경우 등에서는 인식률이 현저히

[†] 정 회 원 : 한국전자통신연구원 우정기술연구센터 선임연구원

^{††} 정 회 원 : 한국전자통신연구원 우정기술연구센터 자동구분처리연구팀장
논문접수 : 2003년 5월 20일, 심사완료 : 2003년 10월 13일

떨어지게 되어 새로운 인식방법에 대한 연구가 필요하다.

본 논문에서는 다중 신경망을 이용한 인식단위 결합 기반의 인쇄체 문자인식 방법을 제안한다. 인식대상 입력 문자영상은 우편물 영상에서 추출된 저해상도 문자영상으로 한글 문자와 영어 알파벳, 숫자, 기호 등이며, 이들은 문자의 자소조합 정보에 따라 각각 다른 신경망 인식기에 의해 독립적으로 인식된다. 한글의 경우 문자영상은 몇 개의 인식단위로 나뉘어져 인식되고, 한글 이외의 문자는 전체 문자영상이 하나의 인식단위로 인식된다. 본 논문에서는 인식 실험을 위하여 실제 우편물에서 수집된 문자영상 데이터베이스를 사용하였는데 실험 결과, 본 논문의 방법이 인쇄체 문자인식에 유용함을 알 수 있었다.

본 논문의 나머지 구성은 다음과 같다. 먼저, 2장에서 인쇄체 문자인식과 관련한 기존의 연구를 기술하고, 3장에서는 제안하는 문자인식 방법을 기술 하는데 인식단위의 정의와 방향각도 특징 및 인식기의 구성에 대해서 상세히 설명한다. 4장에서는 실험에 사용된 문자영상 데이터와 이를 이용한 인식 실험 결과를 기술하고, 5장에서 결론을 맺는다.

2. 관련 연구

한글 문자인식 방법은 인식단위에 따라 음절단위 인식[9]과 자소단위 인식 방법[1-5]으로 나눌 수 있다. 음절단위 인식이 인식대상 음절의 수에 따라 인식률, 속도 및 메모리 요구량이 비교적 크게 달라지는 반면, 자소단위 인식은 비교적 일정한 수의 자소만을 인식하면 되므로 성능의 변화가 상대적으로 적다. 그러나, 두 접근법 중 어떠한 방법이 우수한가를 단언하는 것은 쉽지 않은 일이다. 왜냐하면 각 방법의 구현에 있어서 사용되는 특징과 적용되는 인식기의 종류, 분류대상 문자의 종류 등에 따라 성능의 차이가 나타나기 때문이며, 주어진 문제에 보다 적절한 인식방법이 적용되어야 한다. 하지만, 자소단위 인식 방법은 음절단위 인식 방법에 비해 근본적으로 한글의 고유한 자소조합 특성을 이용하고 인식대상의 확장 가능성이 크므로 자소조합 특성을 용이하게 이용할 수 있는 인쇄체 문자인식과 온라인 문자인식에 많이 이용된다.

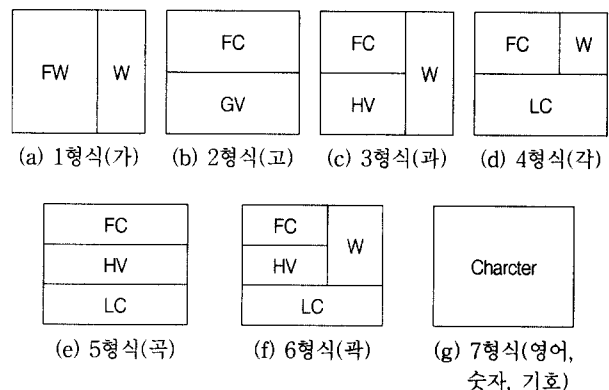
기존의 자소단위 인식에 관한 연구[1-5]는 대부분 한글 문자를 6가지 형식중 하나로 분류한 후, 형식정보를 이용하여 자소분리를 하고 분리된 자소들을 각각 인식하고 조합함으로써 최종 인식을 하는 과정으로 구성되어 있다. 이러한 일련의 과정 중 자소의 분리 과정은 인식 결과에 직접적인 영향을 미치게 되므로 성공적인 자소의 분리는 인식 성능을 향상하는 데 있어 매우 중요하다. 그러나, 완벽한 자소의 분리는 매우 어려우며, 인식대상 자소의 일부분이 소실되는 가능성과 자소 이외의 잡영이 포함될 가능성이 매우 높다. 전자의 경우는 어떠한 인식방법으로도 인식이 어렵고, 후자의 경우는 좋은 인식특징과 인식기를 사용하면 인식이 가능할 수 있다. 몇몇 연구[1-3]에서는 자소의 완전한 분리가 어렵다고 보고 자소 이외의 잡영이 포함되더라도 자소의 소실이 없도록 자소 영역을 충분히 크게 두어 인식을 수행하였

다. 이러한 자소분리 과정에서 포함되는 잡영을 해결하기 위해서 신경망 인식기를 적용하였으며 인식 실험 결과 비교적 우수한 인식률을 얻을 수 있었다. 이관호 등의 연구[4]에서는 자소 인식기로는 앞서 서술한 방법들과 마찬가지로 신경망을 이용하였으나, 자소의 분리는 앞의 연구들과 달리 한글 자소의 구조적인 특성을 적극적으로 이용하는 동적인 분리 방법을 사용하였다. 이들의 방법들은 각각 장단점이 있겠으나, 실험에 사용한 문자 폰트의 종류, 문자영상 품질 등이 서로 상이하므로 어떠한 방법이 보다 우수하다고 단언하기는 매우 어렵다. 기존의 연구에서 사용되는 특징을 살펴보면, 교차거리 및 회교차 특징[6, 10], 방향기여도 특징[4, 8] 등이 많이 사용되었는데, 특히 흑화소 밀도 특징[1-3, 6]이 가장 많이 사용되었다. 이 특징은 인식대상 영상을 몇 개의 그물망으로 나누고, 각 그물망 내에 존재하는 흑화소의 밀도 계산함으로써 간단히 추출된다. 이 특징은 추출의 용이성과 더불어 그물망 내의 회 변화에 비교적 둔감한 특성이 있어 많이 사용되었지만 'o'와 'o' 등 유사한 형태의 자소 구분에는 취약성이 있다고 알려져 있다.

3. 제안하는 인쇄체 문자인식 방법

3.1 문자 형식분류 및 인식단위

본 논문의 연구에서는 인쇄체 한글 문자의 인식을 위해서 자소의 조합 방식에 따라 한글을 6가지의 형식으로 분류하며, 영어 알파벳, 숫자 기호 등의 문자를 하나의 형식으로 분류한다. 즉, 인식대상 형식은 전체 7가지로 이를 (그림 1)에 나타내었다.

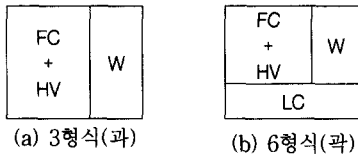


FC : First Consonant, VV : Vertical Vowel, HV : Horizontal Vowel, LC : Last Consonant

(그림 1) 7개의 문자 형식

(그림 1)과 같이 한글 문자를 6개의 형식으로 나누어 인식하는 기존 방법에서는 인식대상 자소의 수는 최소 5개 (2, 5형식 문자의 횡모음)에서 최대 27개(4, 5, 6형식 문자의 종성자음)로, 각 인식대상 자소는 인식 전에 반드시 분리되어야 한다. 대부분의 자소분리는 문자 형식별 자소의 배열과 글꼴에 따라 매우 경험적인 방법을 이용하게 되는데 저해상도 및 저품질 문자영상에 대해서는 분리 난이도가 매우 높아지게 된다. 특히, 3형식 및 6형식과 같이 이중 모음으로

구성된 문자의 경우에는 자소분리가 더욱 어려워진다. 본 논문에서는 기존의 자소단위 인식 방법과 다르게 인식단위를 자소 또는 결합된 자소로 정의함으로써 인식단위의 분리를 보다 용이하게 하고자 하였다. 자소분리 과정에서 자소의 조합이 가장 복잡한 3과 6형식의 문자에서 자소분리가 어려우므로 (그림 2)에서와 같이 3형식과 6형식의 초성자음과 횡모음을 하나의 인식단위로 하여 분리한다. 1, 2, 4, 5형식의 인식단위는 (그림 1)에서 나타낸 기본 자소들이다. 본 논문에서는 한글의 인식단위를 HRU(Hangul Recognition Unit)로 정의한다. 즉, 1, 2, 3형식 문자의 경우 2개의 HRU가 있고 4, 5, 6형식 문자에는 3개의 HRU가 있으며 이를 <표 1>에 나타내었다. 그리고, 7형식 문자는 문자영상 전체를 인식단위로 한다. <표 1>에서 HRU(T, n)은 T형식 문자의 n번째 인식단위를 의미한다.



(그림 2) 3형식 및 6형식 문자의 인식단위

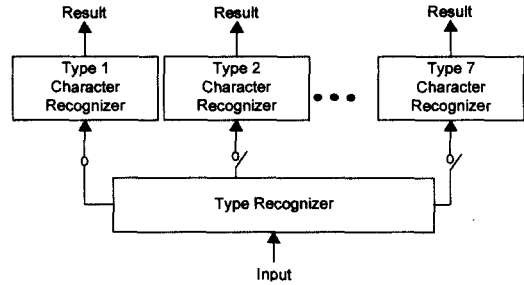
<표 1> 형식별 한글 인식단위 HRU

HRU(Type, index)	Recognition Unit	Example
HRU(1, 1)	FC	ㄱ of 기
HRU(1, 2)	VV	ㅣ of 기
HRU(2, 1)	FC	ㄴ of 노
HRU(2, 2)	HV	ㅏ of 노
HRU(3, 1)	FC + HV	구 of 귀
HRU(3, 2)	VV	ㅣ of 귀
HRU(4, 1)	FC	ㅎ of 한
HRU(4, 2)	VV	ㅏ of 한
HRU(4, 3)	LC	ㄴ of 한
HRU(5, 1)	FC	ㅇ of 울
HRU(5, 2)	HV	ㅡ of 울
HRU(5, 3)	LC	ㄹ of 울
HRU(6, 1)	FC + HV	교 of 관
HRU(6, 2)	VV	ㅏ of 관
HRU(6, 3)	LC	ㄴ of 관

3.2 문자인식기의 구성

본 논문의 인쇄체 문자인식 방법의 대략적인 구성은 (그림 3)과 같다. 입력 문자영상은 형식분류기에 의해 형식이 인식되면, 해당 형식별 인식기로 전송되어 최종 인식이 수행되게 된다.

이러한 인식기의 인식 과정은 비교적 우수한 성능으로 보고되었던 기존의 자소단위 인식 방법들의 인식 과정과 유사한 것이다. 그러나, 인식단위에서 추출되는 특징과 각 형식별 인식기의 세부 구성은 매우 상이한 것으로 이에 대해서는 3.3절과 3.4절에서 상세히 설명하고자 한다.



(그림 3) 인쇄체 문자인식기 구성도

3.3 방향각도 특징

일반적으로 문자인식기의 입력으로 문자영상을 그대로 사용하지 않고 인식에 보다 용이한 특징 벡터로 변환하여 사용하는 것이 효율적이다. 문자인식에서 가장 많이 사용되는 특징은 흑화소 밀도 특징[1, 6]으로, 입력 문자영상을 $N \times M$ 의 그물망으로 분할하여 각 그물망 내에서의 흑화소 밀도를 특징 값으로 사용한다. 이 특징은 추출이 용이하고 그물망 내의 문자획의 변형에 어느 정도 둔감하다는 장점이 있지만, 'ㅇ', 'ㅁ' 등 유사한 문자간의 분별력이 떨어지는 단점이 있다. 본 논문에서는 인식기의 입력 특징을 흑화소 밀도 특징을 보완한 방향각도 특징[15]을 사용한다. 흑화소 밀도 특징은 각 그물망 내에서 한 개의 값으로 표현되는데 반해 방향각도 특징은 각 그물망 내에서 다수 개의 방향각도로 표현되어, 유사 문자간의 분별력이 보다 향상되는 장점이 있다.

영상을 구성하고 있는 각 화소의 위치를 기준으로 인접하는 흑화소들을 벡터로 표현하고 이 벡터들을 합하여 얻어지는 방향각도를 해당 위치에서의 화소 방향각도로 한다. (x, y) 위치에서의 화소 값을 $f(x, y)$ 로 표현하면, 화소의 방향각도 (x, y) 는 식 (1)과 같이 정의된다.

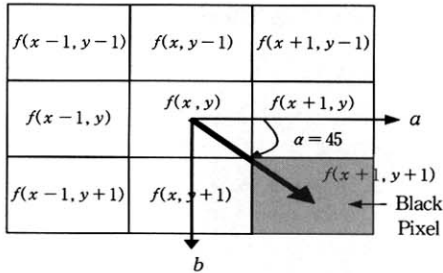
$$a(x, y) = \tan^{-1} \left(\frac{\sum_{a,b} a \cdot f(x+a, y+b)}{\sum_{a,b} b \cdot f(x+a, y+b)} \right), \tag{1}$$

$$\text{where } -\frac{w-1}{2} \leq a, b \leq \frac{w-1}{2}$$

여기서, $f(x, y)$ 는 0 또는 1의 값을 갖는 이진 화소값이며, $f(x+a, y+b)$ 는 $f(x, y)$ 에서 x 축 방향으로 a 만큼, y 축 방향으로 b 만큼의 거리에 위치한 화소의 값을 나타낸다. 또한 w 는 화소 방향각도 계산에 포함되는 인접화소들을 제한하는 창 크기이며, a 와 b 는 정수이다. w 를 3으로 두었을 때 8개의 인접하는 화소가 계산에 포함되며 5로 두었을 때는 24개의 인접화소가 포함된다. (그림 4)에서는 w 를 3으로 두었을 때 $f(x, y)$ 경우의 예를 도시하였다. 창에서 한 개의 검은 화소가 존재하여 화소 벡터는 $[1, 1]^T$ 이며 방향각도는 45° 이다. 여기서, T 는 벡터 전치(transpose)를 의미한다.

$N_x \times N_y$ 크기의 영상에 대해서는 $B_x \times B_y$ 개의 그물망으로 나누는 다음, 각 그물망의 영역별로 bw (bin-width)에 의해 결정되는 L 개의 빈을 할당하여 방향각도 (x, y) 를 식 (2), 식 (3)에 의해 계산되는 빈에 누적한다((그림 5) 참조). 여기서, bw 는 빈의 폭을 나타내는 각도로써 0보다 크고 360도 보

다 작은 값이고, L 은 각 그물망의 빈의 개수이다. S_{bi} 는 bi 번째 그물망 영역의 화소 집합이며, $Fr(bi, n)$ 은 (x, y) 가 저장되는 빈의 누적치로써 bi 번째 그물망의 n 번째 빈에 누적된 (x, y) 들의 총 수를 나타낸다. 입력 영상의 특징벡터는 $Fr(bi, n)$ 을 해당 그물망의 크기인 $m_x \times m_y$ 로 나눈 값을 성분으로 한다. 따라서, $B_x \times B_y$ 개의 그물망별로 L 개의 빈이 있으므로 특징벡터의 차원은 $B_x \times B_y \times L$ 이 된다.

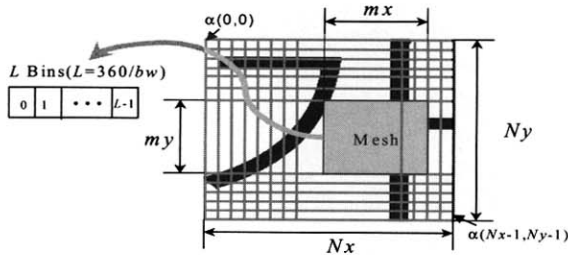


(그림 4) 화소 $f(x, y)$ 의 방향각도 예

본 논문에서는 w 를 3으로 제한하며, 빈의 폭 bw 를 45°로 하여 각 그물망마다 8개의 빈을 둔다. 예를 들어, 방향각도가 0°보다 크고 45°보다 작은 화소의 경우 0번째 빈의 값을 증가시키고, 270°보다 크고 360°보다 작은 화소의 경우 7번째 빈의 값을 증가시킨다. 그물망내의 모든 화소에 대해서 방향각도를 구한 후 8개의 각 빈의 누적값을 그물망 내에 존재하는 화소의 수로 나눈 값을 방향각도 특징으로 한다.

$$L = \frac{360}{bw}, f(x, y) \in S_{bi} \quad (2)$$

$$Fr(bi, n) = \left| \left\{ (x, y) \mid n = \left\lfloor \frac{\alpha(x, y) + 1}{bw} \right\rfloor \right\} \right|, 0 \leq n \leq L-1 \quad (3)$$



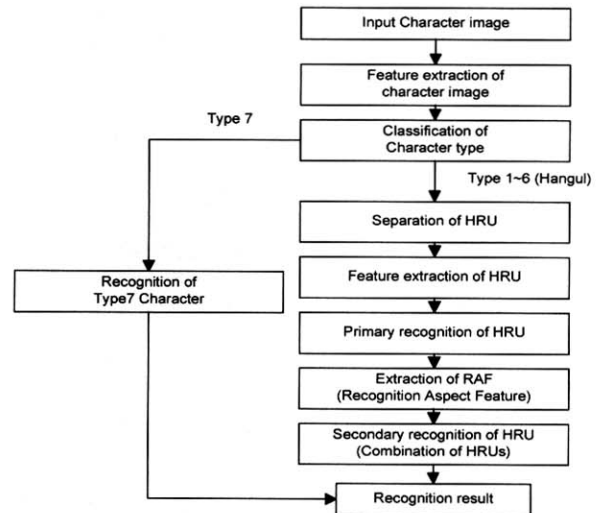
(그림 5) 방향각도 특징 벡터의 추출

3.4 문자 형식별 인식

3.4.1 문자인식 과정

본 논문에서의 문자영상 인식 과정은 (그림 6)과 같다. 우선 입력 문자영상 전체에 대해서 방향각도 특징 벡터를 추출하고 이를 이용하여 형식을 분류한다. 만약 7형식으로 분류된다면 7형식 문자인식기에 의한 결과를 최종 인식결과로 하며, 형식분류가 1~6형식의 한글이면 <표 1>에서 정의된 것과 같은 HRU들을 분리하여 인식한다. HRU들을 분리한 후, 인식하기 위해 각 HRU 영상에 대해서 특징이 추출되고 이 특징을 이용하여 각 HRU가 1차 인식된다. 인식

된 HRU들을 단순히 결합함으로써 인식결과가 종료될 수 있지만, 이들을 단순 결합하면 한 개의 HRU라도 오인식되었을 경우 문자인식 결과가 잘못된다. 본 논문에서는 HRU들을 인식한 후 인식결과를 단순히 조합하지 않고, 2차 인식하는 과정을 수행함으로써 HRU들에 대한 인식결과들을 단순 결합함에 따라 일어나는 오인식 가능성을 감소시키고자 한다. 각 HRU의 재인식을 위해서 1차 HRU 인식이 HRU를 인식하는 과정 중에 나타나는 인식영상들을 새로운 특징으로 추출하고, 이를 2차 HRU 인식기에 입력함으로써 HRU에 대한 인식이 완료된다. 인식영상 특징에 대해서는 뒤에서 설명하기로 한다. 마지막 단계로 인식된 HRU들을 결합하여 최종 인식된 문자를 출력하게 된다.



(그림 6) 문자인식 과정

3.4.2 문자 형식별 인식기 구성

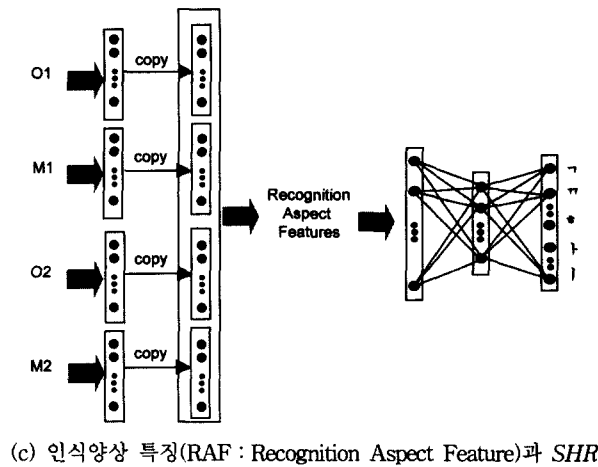
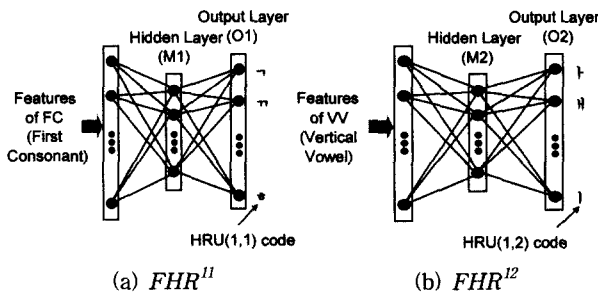
한글 형식별 문자인식기는 다음 식과 같이 표현된다.

$$CR^T = \{ FHR^{Tn}, SHR^T \}, 1 \leq T \leq 6, 1 \leq n \leq N_T \quad (4)$$

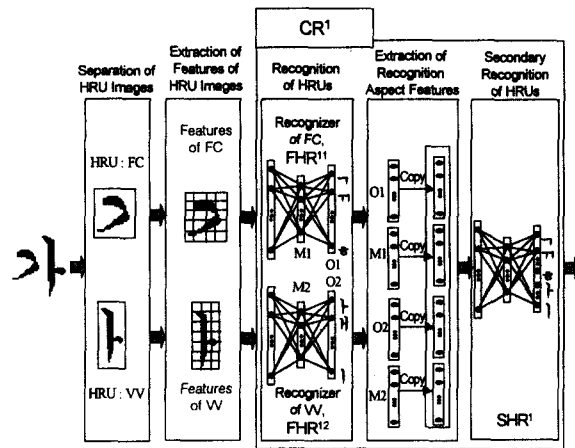
여기서 FHR^{Tn} 은 T 형식 문자의 n 번째 HRU인 $HRU(T, n)$ 를 인식하는 인식기이며 N_T 는 T 형식 문자의 HRU의 개수이다. SHR^T 는 N_T 개의 HRU에 대해서 재인식을 수행하고 결과들을 결합하여 최종인식 결과를 생성하는 인식기이다. 1형식 문자들에 대한 예를 들면, FHR^{11} 은 $HRU(1, 1)$ 즉, 1형식 문자의 첫번째 인식단위인 초성자음을 인식하는 인식기를 뜻하며, SHR^1 는 FHR^{11} 과 FHR^{12} 의 인식 후에 2차적으로 HRU를 인식하는 기능을 한다. CR^1 는 이들 3개의 인식기에 의해서 정의되는 1형식 문자인식기이며 인식 성능은 SHR^1 의 성능이 된다. 7형식 문자에 대한 인식기는 인식단위를 문자영상 전체로 하기 때문에 단일 인식기로 구성되며 CR^7 로 표현한다. 또한 문자 형식분류기는 CR^0 로 표현되며 CR^7 과 마찬가지로 단일 인식기로 구현된다. 이러한 각 인식기들을 <표 2>에 나타내었다. <표 2>의 모든 인식기는 MLP 신경망으로 구현이 된다. 즉, $CR^1, CR^2,$

<표 2> 문자형식별 인식기

Character type recognized	Recognizer
0	CR^0
1	$CR^1 = \{ FHR^{11}, FHR^{12}, SHR^1 \}$
2	$CR^2 = \{ FHR^{21}, FHR^{22}, SHR^2 \}$
3	$CR^3 = \{ FHR^{31}, FHR^{32}, SHR^3 \}$
4	$CR^4 = \{ FHR^{41}, FHR^{42}, FHR^{43}, SHR^4 \}$
5	$CR^5 = \{ FHR^{51}, FHR^{52}, FHR^{53}, SHR^5 \}$
6	$CR^6 = \{ FHR^{61}, FHR^{62}, FHR^{63}, SHR^6 \}$
7	CR^7



(c) 인식양상 특징(RAF : Recognition Aspect Feature)과 SHR^1



(d) 1형식 문자의 인식 흐름

(그림 7) 1형식 문자의 인식

CR^3 은 3개의 신경망으로 구성되어 있고 CR^4, CR^5, CR^6 은 4개의 신경망, CR^7 은 1개의 신경망으로 구성되어 있다.

한글 문자인식기인 $CR^1 \sim CR^6$ 에 대해서는 (그림 7)의 CR^1 을 예로 들어 설명한다. $CR^2 \sim CR^6$ 은 CR^1 과 동일한 방식으로 구현되고, CR^0 및 CR^7 은 전체 입력 문자영상에 대한 방향각도 특징 벡터를 입력으로 하는 단일 신경망으로 구현된다.

FHR^{11} 과 FHR^{12} 의 입력은 추출된 초성자음 영상 및 종모음 영상에 대한 방향각도 특징 벡터이다. 이들의 출력노드는 인식대상 자소의 전체 수와 같으며, 출력 값은 초성자음 및 종모음 각각에 대한 인식 신뢰도 값이다. 개별 HRU 추출은 매우 경험적인 방법을 이용하는 것으로 기존의 방법[1], 방법[3]과 같이 인식단위가 충분히 포함할 수 있도록 HRU 영역을 충분히 크게 두었다. FHR^{11} 과 FHR^{12} 의 인식 과정에서 생성되는 정보인 각 신경망의 중간층 값과 출력층 값은 인식양상 특징 값으로 정의되고 SHR^1 에 입력된다. 즉, 인식양상 특징 벡터의 차원은 각 FHR 인식기의 중간층 노드의 수와 출력층 노드의 수의 합이 된다. SHR^1 의 출력 노드의 개수는 FHR^{11} 과 FHR^{12} 의 각 출력 노드 개수들의 합과 같으며, 각 초성자음 및 종모음에 대한 인식 신뢰도를 출력한다.

4. 실험 및 결과

본 논문에서 구현되는 문자인식기에 대한 성능 실험은 200 dpi의 해상도로 입력된 실제 우편봉투 영상에서 추출한 인쇄체 문자영상에 대해서 수행되었다. 우편봉투 영상은 1999년부터 2001년간 서울 및 동서울 우편집중국에서 수집된 우편봉투 5만통에서 추출한 것으로 재질, 창의 유무, 글자체 등이 매우 다양한 형태로 되어 있다. 우편봉투 내에 존재하는 문자 영상은 기존의 인쇄체 한글 문자인식에 사용된 300 dpi 해상도의 문자에 비해 크기가 일반적으로 작은 것이 특징이다. 또한 이진화의 영향에 따라 문자의 획이 끊어지거나 잡영이 존재하는 등 매우 저품질의 문자영상이 다수 존재한다. 각 인식기의 학습을 위해서 301,070개의 문자 영상을 사용하였고, 테스트를 위해서 301,071개의 문자영상을 사용하였다.

본 연구에서는 <표 2>의 각 형식별 문자인식기를 신경망의 중간층 노드 개수를 변화 시키면서 다양한 크기로 구현하였다. 모든 신경망 인식기에 대해서 중간층의 노드 개수를 30, 40, 50, 60, 70개로 변화시키면서 구현하였는데, 신경망의 학습 파라미터는 학습율 0.1, 관성항 0.7, 학습 반복 횟수 100회로 동일하게 하여 오류역전파 방법 [16]을 이용하여 학습을 하였다. $CR^1 \sim CR^6$ 의 SHR 인식기는 FHR 인식기와 동시에 학습하지 않고, 각 FHR 인식기의 학습을 모두 완료한 후 학습하였다.

테스트 문자영상에 대한 모든 인식기의 인식 성능을 <표 3>에 나타내었는데, FHR 인식기의 각 HRU에 대한 인식률, FHR 인식기의 인식결과를 단순히 결합했을 때의 문자 인식률, SHR 인식기의 문자인식률 등을 신경망 중간층의 수에 따라 나타내었다. 각 HRU별 인식에 있어서는 5형식

의 수평모음 인식기인 FHR^{22} 와 6형식의 횡모음 인식기인 FHR^{62} 를 제외하고는 대부분의 HRU 인식기가 99%이상의 높은 인식 성능을 나타내었다. 주목할만한 것은 3형식을 제외하고는 모든 형식에 있어서 자음보다는 모음에 대한 인식률이 낮다는 것이다. 이는 모음이 작은 획 또는 점에 의해 구분되므로 인식단위 영상 내에 잡영이 포함되었을 경우 자음에 비해 잡영에 민감하여 인식이 어려워지기 때문이다. 모든 HRU 인식 성능 중 5형식 문자의 수평모음에 대한 인식률이 가장 낮는데, 이는 모음과 자음의 접촉이 다른 형식에 비해 심하기 때문이다.

각 형식별 문자인식률을 FHR 인식기의 인식결과를 단순히 결합한 경우와 SHR 인식기의 경우에 대해 살펴보면, SHR 인식기의 인식성능이 보다 우수함을 알 수 있다. 단순결합의 문자인식 성능과 SHR 문자인식 성능을 다음의 식 (5)와 같이 표현되는 오류감소율측면에서 분석해보면 SHR 인식기의 성능 향상을 명확히 알 수 있다. 식 (5)에서 ERR 은 오류감소율을 의미하며, SCE 는 단순결합에서의 오류율, $SHRE$ 는 SHR 인식기의 오류율을 의미한다.

$$ERR = \frac{SCE - SHRE}{SCE} \times 100(\%) \quad (5)$$

식 (5)를 각 형식별 문자인식기들에 적용한 결과를 (그림 8)에 나타내었다. 그림을 살펴보면, 모든 형식에 있어서 단순결합에 비해 성능의 향상이 있으며, 2형식의 오류감소가 상대적으로 뚜렷이 나타나고 1형식의 오류감소는 작게 나타남을 알 수 있다.

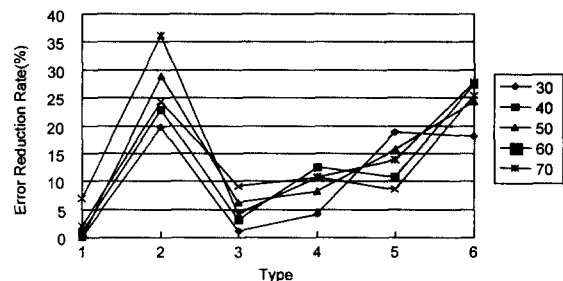
SHR 인식기의 성능은 5형식 문자를 제외하고는 모두 평균 98%이상의 높은 인식률을 나타내었다. 이러한 문자인식률은 기존의 문자인식에 관한 연구들의 결과와 비교하여도 매우 높은 인식 성능임을 알 수 있다. 그리고, 7형식 문자의 인식율은 99.22%로 단순한 단일 신경망 인식기로도 매우 우수한 인식 성능을 얻을 수 있었다.

<표 4>에는 <표 3>의 한글 형식별 인식기 중 인식 성능이 가장 우수한 인식기들의 인식 특성을 나타내었다. 표에서 '○'은 올바르게 인식되었음을 나타내고 '×'는 잘못 인식되었음을 나타낸다. 예를 들 <표 4>(a)에서 HRU(1,1)과 HRU(1,2)가 FHR^{11} 과 FHR^{12} 에 의해 모두 올바르게 인식되었을 경우, SHR^1 에 의해 올바르게 인식된 문자의 수는 52,633이고 틀리게 인식된 문자의 수는 47개임을 의미한다. 그리고, HRU(1,1)이 올바르게 인식되고 HRU(1,2)가 틀리게 인식되었지만, SHR^1 에 의해서는 올바르게 인식되는 문자의 수는 63개로, 인식단위의 인식 결과를 단순 조합하면 오인식되지만 SHR^1 에 의해 오류 교정된 문자가 63개임을 의미한다. 전체적으로 각 인식단위의 인식 결과를 단순 조합했을 경우 인식된 문자의 수는 전체 52,680이며 SHR^1 에 의해 인식된 문자의 수는 52,722(= 52,633 + 63 + 25 + 1)로 모두 42개의 문자가 올바르게 교정되어 인식되었다. <표 3>의 결과 분석에서와 마찬가지로 3형식의 경우를 제외하고는 두 번째 인식단위인 모음의 인식 성능이 자음의 인식 성

능에 비해 낮게 나타남을 볼 수 있다. 그리고, 2개 이상의 인식단위가 동시에 오인식된 경우는 매우 드물게 나타났다.

<표 3> 인식기별 인식 실험 결과

중간층 노드수 인식기	30	40	50	60	70
FHR^{11}	99.56	99.58	99.63	99.63	99.60
FHR^{12}	99.16	99.18	99.25	99.32	99.23
단순결합	98.75	98.79	98.89	98.97	98.85
SHR^1, CR^1	98.75	98.80	98.89	98.99	98.93
FHR^{21}	99.74	99.78	99.80	99.80	99.81
FHR^{22}	99.49	99.52	99.47	99.50	99.51
단순결합	99.24	99.30	99.27	99.30	99.31
SHR^2, CR^2	99.39	99.46	99.48	99.47	99.56
FHR^{31}	99.27	99.30	99.32	99.30	99.33
FHR^{32}	99.74	99.70	99.70	99.69	99.71
단순결합	99.02	99.01	99.04	99.01	99.05
SHR^3, CR^3	99.03	99.04	99.10	99.10	99.09
FHR^{41}	99.27	99.36	99.45	99.45	99.49
FHR^{42}	98.88	99.05	99.15	99.20	99.25
FHR^{43}	99.19	99.16	99.25	99.26	99.25
단순결합	97.41	97.63	97.91	97.96	98.04
SHR^4, CR^4	97.52	97.93	98.08	98.18	98.25
FHR^{51}	99.48	99.58	99.57	99.52	99.58
FHR^{52}	97.67	97.88	97.88	97.98	97.96
FHR^{53}	99.63	99.62	99.67	99.65	99.68
단순결합	96.85	97.14	97.19	97.23	97.26
SHR^5, CR^5	97.44	97.45	97.63	97.47	97.64
FHR^{61}	99.07	99.22	99.10	99.17	99.25
FHR^{62}	98.83	98.76	98.89	98.86	98.91
FHR^{63}	99.61	99.56	99.49	99.62	99.56
단순결합	97.57	97.62	97.57	97.72	97.76
SHR^6, CR^6	98.01	98.27	98.16	98.30	98.38
CR^7	98.96	99.08	99.13	99.12	99.22



(그림 8) HRU 인식결과 단순결합 방법과 비교한 SHR 인식기의 오류감소율

<표 4> 형식별 문자인식기의 인식 특성(○ : 인식 성공, × : 인식 실패)

(a) Type 1

HRU(1, 1)	HRU(1, 2)	SHR		단순결합	
		인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	52,633	47	52,680	0
○	×	63	338	0	401
×	○	25	174	0	199
×	×	1	11	0	12

(b) Type 2

HRU(2, 1)	HRU(2, 2)	SHR		단순결합	
		인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	36,845	10	36,855	0
○	×	87	96	0	183
×	○	13	59	0	72
×	×	0	0	0	0

(c) Type 3

HRU(3, 1)	HRU(3, 2)	SHR		단순결합	
		인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	12,218	10	12,228	0
○	×	4	30	0	34
×	○	12	69	0	81
×	×	1	2	0	3

(d) Type 4

HRU(4, 1)	HRU(4, 2)	HRU(4, 3)	SHR		단순결합	
			인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	○	74,982	101	75,083	0
○	○	×	154	393	0	547
○	×	○	56	493	0	549
○	×	×	0	13	0	13
×	○	○	53	317	0	370
×	○	×	1	9	0	10
×	×	○	0	10	0	10
×	×	×	0	1	0	1

(e) Type 5

HRU(5, 1)	HRU(5, 2)	HRU(5, 3)	SHR		단순결합	
			인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	○	47,021	122	47,143	0
○	○	×	39	108	0	147
○	×	○	217	757	0	974
○	×	×	1	3	0	4
×	○	○	51	141	0	192
×	○	×	0	3	0	3
×	×	○	0	8	0	8
×	×	×	0	1	0	1

(f) Type 6

HRU(6, 1)	HRU(6, 2)	HRU(6, 3)	SHR		단순결합	
			인식 문자 수	오인식 문자 수	인식 문자 수	오인식 문자 수
○	○	○	5,980	12	5,992	0
○	○	×	10	15	0	25
○	×	○	15	49	0	64
○	×	×	0	2	0	2
×	○	○	25	20	0	45
×	○	×	0	0	0	0
×	×	○	0	1	0	1
×	×	×	0	0	0	0

지금까지 논의된 형식별 문자인식률은 해당 문자 형식을 미리 알고 해당 인식기를 적용하여 얻은 문자인식률이다. 형식분류기 CR^0 를 포함한 전체적인 문자인식 성능의 실험 결과는 <표 5>와 같이 나타났다. 표를 살펴보면 1형식과 3형식, 4형식과 6형식 간의 형식분류 오류가 다른 형식간의 분류 오류보다 크다는 것을 알 수 있다. 이는 '과'와 '과' 등의 예에서처럼 단일 자음과 횡모음의 결합된 형태가 단일 자음과 유사하므로 발생하는 오류인 것으로 판단된다. 그러나 전체적으로 99.9%의 형식분류율과 98.62%의 문자인식률은 기존의 다른 연구에서 보고된 인식률에 비해 우수한 것으로 제안하는 방법이 인쇄체 문자의 인식에 효과적임을 증명하는 것이다.

<표 5> 형식분류율 및 형식별 문자인식기의 인식율

In\Out	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type Rec.(%)	Char. Rec
Type 1	53,254	0	20	6	0	0	12	99.93	98.87
Type 2	0	37,086	0	3	5	1	15	99.94	99.49
Type 3	24	1	12,310	10	0	1	0	99.71	98.84
Type 4	19	1	7	76,506	7	33	10	99.90	98.24
Type 5	1	1	0	3	48,463	0	4	99.98	97.72
Type 6	0	0	4	70	1	6,054	0	98.78	97.32
Type 7	8	15	1	9	7	1	67,098	99.94	99.13
계								99.90	98.62

5. 결 론

본 논문에서는 다중 신경망을 이용한 인식단위 결합 기반의 인쇄체 문자인식 방법을 제안하였다. 인식대상 입력 문자영상은 우편물 영상에서 추출된 200dpi의 저해상도 문자영상으로 한글 문자와 영어 알파벳, 숫자, 기호 등이며, 한글의 경우는 자소의 조합 형태에 따라 6가지의 형식으로 분류되어 인식된다. 문자영상은 한글 6형식 또는 한글 이외의 문자 형식을 구분하는 형식분류기에 입력되어, 형식분류기의 출력에 따라 각 형식별 문자인식기로 전송된다. 한글 문자는 2단계의 MLP 신경망 인식기에 의해 인식된다. 첫째 단계에서는 한글 문자를 자소의 조합 형태에 따라 2개

또는 3개의 인식단위로 나누고, 각 인식단위를 문자영상 화소의 방향각도 특징 벡터를 입력으로 하는 MLP 신경망으로 인식한다. 둘째 단계에서는 첫째 단계의 인식 결과를 단순 조합하지 않고, 첫째 단계의 MLP 신경망의 인식양상을 다른 MLP 신경망에 입력하여 최종 한글 문자인식을 한다. 한글 이외의 기타 문자의 인식을 위해서는 단일 MLP 신경망을 사용한다.

이러한 과정으로 구성된 제안된 방법은 실제 우편물 50,000통 영상으로부터 추출된 문자영상 데이터베이스 상에서 구현되었다. 실험 결과 문자의 형식분류율은 99.9%, 문자인식율은 98.62%로 제안한 방법이 인쇄체 문자인식에 유용함을 알 수 있었다.

참 고 문 헌

[1] 권재욱, 조성배, 김진형, "계층적 신경망을 이용한 다중 크기의 다중활자체 한글문서 인식", 한국정보과학회논문지, 제19권 제1호, pp.69-79, 1992.

[2] S. B. Cho and J. H. Kim, "Hierarchically structured neural networks for printed Hangeul character recognition," *International Joint Conference on Neural Networks*, Vol.1, pp.265-270, 1990.

[3] 이진수, 권오준, 방승양, "개선된 자소 인식 방법을 통한 고인식을 인쇄체 한글 인식", 한국정보과학회논문지, 제23권 제8호, pp.841-851, 1996.

[4] 이판호, 장희돈, 남궁재찬, "동적자소분할과 신경망을 이용한 인쇄체 한글 문자인식에 관한 연구", 한국통신학회논문지, 제19권 제1호, pp.2133-2145, 1994.

[5] 최동혁, 류성원, 강현철, 박규태, "계층구조 신경망을 이용한 한글 인식", 대한전자공학회논문지, 제28권 B편 제11호, pp.1-7, 1991.

[6] 김우태, 윤병식, 박인규, 진성일, "인쇄체 한글 문자인식을 위한 특징성능의 비교", 한국정보과학회논문지, 제20권 제8호, pp.1103-1110, 1993.

[7] S. I. Chien, "Hangeul(Korean) and English OCR system using multiple hypothesis driven neural nets," *Korean-French Character Recognition Workshop*, pp.37-52, 1994.

[8] 장명욱, 천대녕, 양현승, "연결화소를 이용한 문서 영상의 분할 및 인식", 한국정보과학회논문지, 제20권 제12호, pp.1741-1751, 1993.

[9] 이성환, "다양한 활자체 및 크기를 갖는 대용량 한글의 고속 인식을 위한 최적 트리 분류기", 한국정보과학회논문지, 제20권 제8호, pp.1083-1092, 1991.

[10] 김정우, 이행세, "인쇄체 한글 및 한자의 인식에 관한 연구", 한국통신학회논문지, Vol.17, No.11, pp.1175-1184, 1992.

[11] H. Kim and J. Kim, "Hierarchical random graph representation of handwritten characters and its application to Hangeul recognition," *Pattern Recognition*, Vol.34, pp.187-201, 2001.

[12] S. Jeong, K. Lim and Y. Nam, "A combination method of two classifiers based on the information of confusion matrix," *International Workshop on Frontiers in Hand-*

writing Recognition, pp.519-523, 2002.

[13] K. Kim, J. Kim, and C. Suen, "Segmentation-based recognition of handwritten touching pairs of digits using structural features," *Pattern Recognition Letters*, Vol. 23, pp.13-24, 2002

[14] G. Kim and V. Govindaraju, "A Lexicon driven approach to handwritten word recognition for real-time applications," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.19, No.4, pp.366-379, 1997.

[15] 김호연, 임길택, 김두식, 남윤석, "서장 우편물 자동처리를 위한 우편영상 인식 시스템", 정보처리학회논문지B, 제10-B권 제4호, pp.429-442, 2003.

[16] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning internal representations by error propagation," *Parallel Distributed Processing*, Vol.1, pp.319-362, 1986.



임길택

e-mail : ktlim@etri.re.kr

1993년 경북대학교 전자공학과(공학사)
1995년 경북대학교 전자공학과(공학석사)
1999년 경북대학교 전자공학과(공학박사)
1999년~현재 한국전자통신연구원
우정기술연구센터 선임연구원

관심분야 : 패턴인식, 문자인식, 영상처리, 컴퓨터비전, 신경망 등



김호연

e-mail : hoyon@etri.re.kr

1992년 연세대학교 전산학과(이학사)
1994년 한국과학기술원 전산학과(공학석사)
1997년 일본 NHK放送技術研究所
방문연구원
1999년 한국과학기술원 전산학과(공학박사)

1999년 독일 SIEMENS ElectroCom 방문연구원
1999년~현재 한국전자통신연구원 우정기술연구센터 선임연구원
관심분야 : 패턴인식, 문자인식, HMM, 영상처리, 신경망, 기계 학습 등



남윤석

e-mail : ysnam@etri.re.kr

1984년 아주대학교 산업공학과(학사)
1989년 Polytechnic Univ.(New York),
Dept. of the Industrial Engineering(공학석사)
1992년 Polytechnic Univ.(New York),
Dept. of the Industrial Engineering(공학박사)

1993년~현재 한국전자통신연구원 우정기술연구센터 자동구분 처리연구팀장
관심분야 : 소프트웨어 공학, 패턴인식 등