

잡음 성분을 포함한 한글 문자 인식

장 선영 조 동섭
이화여자대학교 전자계산학과

Recognition of Hangeul Characters with Input Noise

Sun-Young Chang Dong-Sub Cho

Department of Computer Science, Ewha Womans University

ABSTRACT

This thesis proposes a new scheme for the recognition of presegmented Hangeul characters. The proposed approach is rather insensitive to noise and variation by applying 2 dimensional convolution to learning patterns. In this thesis, the Hangeul recognition neural network is implemented in the basis of this scheme and recognition rate is analyzed in two cases of learning which are learning by binary patterns and learning by binary patterns and convoluted patterns together.

I. 서론

패턴 인식은 인공 지능의 중요한 연구분야이다. 특히 문자 인식은 실용성에 대한 기대감 때문에 오래전부터 연구되어 왔으며, 우리나라에서도 한글 문자 인식에 초점을 맞추어 많은 연구가 진행되어왔다. 하지만, 한글은 첫째로 문자수가 방대하며, 둘째로 자모가 극히 단조로우며 대부분의 문자가 유사성을 가지고 있다는 문제점때문에 훌륭한 조직체계를 가지고 있기는 하지만 이를 인식해 낼 적절한 알고리즘이 없어 기계적인 인식을 하는데 있어서 많은 난점을 가지고 있어왔다.

신경회로망이 등장하게됨에 따라, 패턴 인식에 필요한 많은 양의 데이터를 병렬처리 할 수 있고, 패턴인식과 같은 수학적 알고리즘의 적용이 곤란한 문제를 학습방법에 의하여 해결할 수 있으며, 패턴 인식에 있어서 자주 발생하는 잡음이 있거나 불완전한 데이터를 효과적으로 처리할 수 있다는 특징으로 기존의 Neumann형 컴퓨터에서는 비효율적이었던 한글 인식 문제에 대한 효율적인 해결안을 제시하였다. 한글을 인식하기 위해 신경망을 이용하는 경우, 기존의 인식 방법 즉 원형정합 방

법(template matching), 통계적 방법 및 구조적 분석 방법 등에서 제기된 문제점을 다소 해결해 주기는 하였지만 아직도 해결되어야 할 많은 문제점들이 남아 있다.

따라서 본 논문에서는 신경회로망을 이용한 한글 문자 인식의 문제점들 중 잡음 성분으로 인한 오인식이라는 문제를 해결하는 방안으로서 잡음 및 결손에 대한 민감도를 극소화 시킴으로써 인식도를 향상시킬 수 있는 방법을 제안하고 있다.

이 신경망은 전산 원고지를 이용한 한글 문자를 인식의 대상으로 삼으며, 초성, 중성(좌), 중성(우), 종성에 대한 4개의 자소별 인식 신경망으로 구성된다. 학습시 이진 패턴과 2차원 본범무선을 적용한 패턴을 함께 학습 패턴으로 사용함으로써 잡음에 대한 적응 능력을 향상시키고자 하였다. 또한 이진 패턴을 학습 패턴으로 사용한 경우와 본범무선된 패턴을 학습 패턴으로 사용한 경우를 인식률과 인식 속도면에서 비교 분석하여 보았다.

II. 잡음성분을 포함한 한글 문자 인식

1. 가우스 (Gaussian) 함수에 대한 본범무선(convolution)

하나의 화상은 화상 평면(image plane)상의 화소점(pixel:picture element)을 부여하는 하나의 함수로서 생각할 수 있다.

시각 처리 과정에서 화상에 대한 데이터 구조는 하나 또는 그 이상의 윤곽선 절편들로 구성된 특징을 묘사하며, 윤곽선 절편들과 이들로 구성된 특징들은 계조도가 급격히 변화하는 부분을 관찰함으로써 얻을 수 있다. 즉 2차 미분 변화율이 0점을 지나게 되는 점(영교차점:zero crossing)이 1차 미분 변화율의 최대값에 해당하는 점이며 그 위치가 윤곽선의 위치가 된다. 완만한 형태의 곡선에 대해서 영교차점을 찾아내기는 용이하지

만, 잡음이 섞여 있는 실제의 화상에 대해서는 잡음으로 인해 발생될 수 있는 정상값 때문에 여러 위치에서 영고차점이 발견될 수 있어서 확실한 윤곽선을 찾아내기가 어렵다. 따라서 이러한 경우 잡음으로 인한 영향을 최소화시키기 위해서 이웃점들과 평균을 취한 값으로 곡선을 평활(smoothing)하게 되는데 수식으로 표현하면 다음과 같다. $w(i)$ 는 가중값을 나타내며, $g(i)$ 는 계조도 함수를 나타낸다.

$$\text{평활된(smoothed) } g(i) = \sum_{j=i-1}^{j=i+1} g(i-j)w(j)$$

가우스함수는 평균되어질 이웃화소들의 수와 정점값의 정도에 따라 여러가지 형태가 있다. 고해상도의 가우스함수는 잡음 등에 의한 정점값의 영향을 완전히 해제하지 못하며 저해상도 가우스 함수는 원화상에 비해 너무 넓게 평활을 취하여 명암의 차이를 약화시킨다는 특징을 가지고 있는데, 평활을 행할 때 적합한 가우스 함수를 선택하는 것 또한 중요하다.

2. 잡음 처리 능력을 향상시킨 학습 방법

실제 스캐너를 통하여 입력된 문자는 같은 문서의 글자라도 스캐너의 작동 상태등 기계적 여건에 따라서 다르게 나타난다. 이러한 이유로 제한된 학습 패턴으로는 현실적인 복잡한 문제에 대하여 적절한 일반화를 이루기 어렵기 때문에 좋은 성능을 기대할 수 없는데, 스캐너를 통해 입력된 문자를 인식하기 위한 신경회로망의 일반화 능력 실험에 의하면 인식률을 저하시키는 주된 요인은 전체 문자가 상하좌우로 이동된 것임을 알 수 있다.

예상되는 입력의 변형이 비교적 크지 않은 문제에서는 신경망의 구조 자체를 복잡하게 하는 것 보다는 입력에 가능한 변형이 포함된 데이터를 학습하는 것이 효과적이다. 한글에도 위치나 크기의 변형이 심한 실제 문제를 신경망으로 해결하기 위해 변형을 첨가한 학습 방법을 이용한 예가 있는데, 가능한 입력의 변형으로 일반화 하기 어려운 상하좌우로 이동시킨 5쌍의 패턴을 학습에 참여시킴으로써 변형된 입력 문자를 인식할 수 있도록 하였다. 하지만 이러한 방법은 학습하여야 할 입력 패턴의 수가 너무 많아지게 되어 신경망을 안정화시키는데 오랜 시간이 지연된다는 단점을 가지고 있다. 본 논문에서는 이러한 잡음에 대한 적응도를 향상시키기 위한 방법으로 학습패턴에 본별부선을 적용하는 새로운 방법을 제안하고 있으며 그 이론은 다음과 같다.

표준 학습 패턴의 1에 해당하는 화소값은 본별부선한 학습 패턴에서 1과 같거나 작게, 그리고 0에 해당하는 화소값은 0과

같거나 크게 조정되어, 본별부선한 학습 패턴이 표준 패턴에서 보다 완만한 화소값을 갖게 됨을 알 수 있다. 따라서 인식의 단계에서 입력 화상에 잡음이 첨가될 경우 즉 흑화소(화소값=1)가 백화소(화소값=0)로 또는 백화소가 흑화소로 변하는 경우 잡음으로 인한 영향을 감소시킬 수 있다.

본별부선을 이용하여 재구성된 화상은 본별부선이 적용되기 이전의 화상에 비해 희미하게 퍼진 형태를 취하게 되는데, 이러한 특성으로 인해 스캐너를 통한 입력 문자 인식시의 문제점인 전체 문자의 이동에 대해서도 어느정도 융통성을 부여할 수 있게 되었다. 또한, 이러한 학습방법은 변형을 첨가한 학습 방법에 비해 적은 수의 패턴으로 학습효과를 얻을 수 있다는 장점을 가지고 있기도 하다.

III. 본별부선을 이용한 한글 문자 인식 시스템의 구현

1. 학습 패턴의 설계

본 논문에서는 전산 원고지를 사용한 한글 문자를 인식의 대상으로 삼는다. 전산 원고지는 이미 규격화되어 있는 영역을 분할하여 인식이 이루어지므로 사용시 주의를 요하는 단점을 가지고 있기는 하지만, 필기체 한글문자의 처리시 모아쓰기를 한 자료를 직접 입력으로 사용할 수 있고 입력된 문자가 신속하게 기본 자모로 분리(segmentation)될 수 있다는 장점을 가지고 있기도 하다.

한글은 한글 기본 자모를 모아쓰기하여 구성되며 모음의 형태와 중성의 존재 유무에 따라서 6가지 유형으로 분류될 수 있는데, 전산 원고지는 이러한 6가지의 구성 형식을 기본으로 설계되었다. 그림 1에 전산 원고지의 형태가 제시되어 있는데, 중자음과 복자음을 단자음으로 분리하여 인식 가능하게 하기 위하여 중자음과 복자음이 사용되는 중성 영역을 세분화하였다. 전산 원고지는 1개의 초성 영역, 4개의 중성 영역, 2개의 종성 영역으로 분할된다. 분할된 영역 중 1, 2, 4, 6, 7 영역만이 처리 대상이 된다.

1	3
	4
2	5
6	7

그림 1. 기본선

본 논문에서 사용하고 있는 학습 패턴은 이러한 전산원고지에서 제시된 영역에 근거하여 설계하였으며, 자형(font)으로는 고딕체를 선정하여 실험하였다. 초성으로는 20 x 16의 기본 자

모 19개의 패턴을, 중성(2영역)으로는 20 x 8의 5개의 자모와 복모음으로 사용될 때를 고려하여 σ , τ , ω 에 변형을 가한 3개의 패턴을 합하여 8개의 패턴을, 중성(4영역)으로는 16 x 10의 기본 자모에 해당하는 9개의 패턴을, 중성(6,7영역)은 동일하게 취급하여 20 x 16의 14개 자모를 학습 패턴으로 취하였다. 그리고, 전산 원고지를 사용하는 경우 임의의 영역에는 자모가 표현되지 않는 경우가 있으므로 자모의 유무를 좀 더 확실하게 구분하여 주기 위하여 각각의 영역에 공패턴(empty pattern)을 추가하여 학습패턴으로 설정하였다. 패턴의 크기는 전산 원고지에서 제시된 각 영역의 폭과 높이의 비율을 따르고 있다. 중성(6,7영역)의 경우 7영역(16x16)과 6영역(20x16)은 크기에 있어서 약간의 차이가 나는데 이 문제는 인식단계에서 크기를 확대 또는 축소할 해법으로써 해결하였다.

이와 같이 선정된 표준 학습 패턴에 2차원 컨벌루션을 적용함으로써 컨벌루션된 학습패턴을 만들어 낼 수 있는데, 컨벌루션 시 각점에서 주변의 점들에 부여될 가중치는 그림 2에 3 x 3의 가중치 윈도우(window)의 형태로 제시되어 있다.

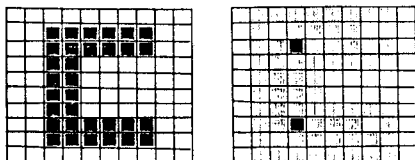
0.05	0.075	0.05
0.075	0.5	0.075
0.05	0.075	0.05

그림 2. 가중치 윈도우

2차원 화상의 컨벌루션은 원화상위에서 그림 2의 가중치 윈도우를 이동시키면서 원화상의 화소값에 가중치를 곱하여 새로운 화소값을 얻음으로써 이루어진다. 예를 들면, N x M 화상의 경우 다음의 식을 이용해 컨벌루션된 학습 패턴을 만들어 낼 수 있다. 가중치는 w라는 1차원 배열에 저장되어 있으며, 원화상은 p1이라는 2차원 배열에 그리고 컨벌루션된 화상은 p2라는 2차원 배열에 저장되어 있다고 하자. 이때 하나의 화소 p1(i, j)의 컨벌루션된(convoluted) 화소값은 다음과 같이 계산된다.

$$p2(i, j) = p1(i-1, j-1) * w(0) + p1(i-1, j) * w(1) + p1(i-1, j+1) * w(2) + p1(i, j-1) * w(3) + p1(i, j) * w(4) + p1(i, j+1) * w(5) + p1(i+1, j-1) * w(6) + p1(i+1, j) * w(7) + p1(i+1, j+1) * w(8)$$

단, p1(i, m) = 0.0 (1 < i <= m <= 0 또는 1 > i >= m >= N 인 경우)



a. 이진 화상. b. 컨벌루션된 화상.

그림 3. 컨벌루션된 화상의 예.

본 논문에서는 2가지 실험을 통해 그 결과를 비교 분석하였는데 실험 1에서는 표준 패턴만을 학습 패턴으로 취하고 있으며 실험 2에서는 표준 패턴과 콘벌루션된 패턴을 함께 내열하여 학습패턴으로 취하고 있다.

2. 신경회로망의 구성과 학습

본 논문에서 사용한 시스템은 4 개의 자소별 인식 신경회로망으로 구성되어 있다. 각각의 신경회로망은 2-layer perceptron의 신경회로망을 사용하고, 역전파 학습 알고리즘에 따라 학습이 진행된다. 시스템의 전체 구성은 그림 4에 나타나 있다. 각각의 자소별 인식 신경망은 다음과 같이 구성되어 있다. 입력층은 각 영역의 크기만큼의 노드로, 그리고 출력층은 분류되어야만 하는 패턴의 갯수만큼의 노드들로 구성하였으며, 은닉층의 노드의 갯수는 다음의 식을 이용하여 결정하였다.

$$2 * \sqrt{\text{입력 노드의 갯수} * \text{출력 노드의 갯수}}$$

그리고, 입력층과 은닉층/은닉층과 출력층간의 각 노드들은 완전 결합된(fully connected) 연결 구조를 가지도록 구성하였다.

신경회로망은 EBP 학습 알고리즘에 따라 학습이 진행되는데, 학습 단계에서는 어떤 특정 패턴에 대해서 단 하나의 출력층 노드만이 활성화되도록 지도 학습하였으며, 인식의 단계에서는 출력층의 노드 중 최대값을 가지고 활성화되는 노드를 선택함으로써 자소를 인식하였다.

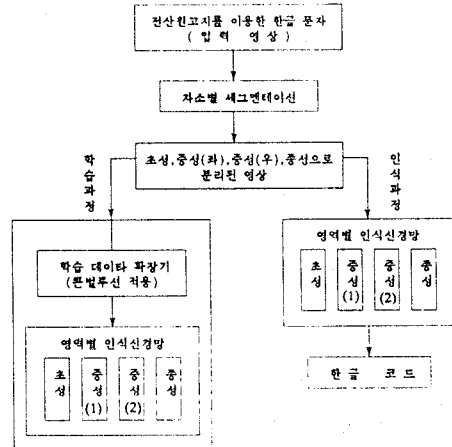


그림 4. 시스템의 전체 구성도

IV. 실험 및 결과 분석

1. 실험 환경

실험에 사용한 인식 시스템은 IBM PC 486(33 MHz) 상에서 C언어로 구현하였으며, 실험 1과 실험 2에서의 실험 환경은 표 1과 같다.

표1. 실험 환경

		실험1	실험2
입력노드의 갯수	초성	320	
	중성1	160	
	중성2	160	
	종성	320	
출력노드의 갯수	초성	21	
	중성1	9	
	중성2	10	
	종성	15	
은닉노드의 갯수	초성	36	
	중성1	26	
	중성2	26	
	종성	36	
학습패턴의 갯수	초성	21	42
	중성1	9	18
	중성2	10	20
	종성	15	30
학습률 (Learning Rate)		0.9	
모멘텀 (Momentum Rate)		0.7	
Total System Error		0.0002	
Pattern Error		0.001	

2. 실험 결과의 분석

본 논문에서는 잡음에 대한 적응도를 분석해보기 위하여 2 가지 실험을 행하였다. 실험 1은 정상적인 패턴으로 학습한 경우이고 실험 2는 콘벌루션된 패턴을 정상적인 패턴과 함께 학습 패턴으로 취하여 학습한 경우이다. 패턴에 대한 콘벌루션이 패턴 인식에 있어서 어떠한 영향을 주는지를 비교해 보기위해 실험 1과 실험 2에서 연상 패턴에 콘벌루션을 적용한 경우와 적용하지 않은 경우로 나누어 인식해 보았다.

임의로 적당한 정도의 잡음을 추가시킨 데이터로 인식했을 경우에는 실험 2에서 더 높은 반응을 보이며 인식하고 있었다.

표 2. 우측 상단으로 한 화소 이동된 초성 패턴에 대한 인식 결과 비교

		< 실험 1.>		< 실험 2.>	
		연상패턴에 대한 콘벌루션의 적용 여부			
연상패턴의 종류		적용하지 않은 경우	적용한 경우	적용하지 않은 경우	적용한 경우
오인식된 갯수	이동만시킨 경우	3	2	2	0
	이동 + 잡음 첨가	5	3	3	3
총 오인식된 패턴의 갯수		8	5	5	3

표3. 우측 상단으로 한 화소씩 이동된 초성 패턴에 대한 인식 결과 비교

		< 실험 1.>		< 실험 2.>	
		연상패턴에 대한 콘벌루션의 적용 여부			
연상패턴의 종류		적용하지 않은 경우	적용한 경우	적용하지 않은 경우	적용한 경우
오인식된 갯수	이동만시킨 경우	13	12	8	7
	이동 + 잡음 첨가	10	9	8	7
총 오인식된 패턴의 갯수		23	21	16	14

표 2와 표 3은 연상패턴이 이동된 경우와 이동된 패턴에 잡음이 첨가된 경우의 인식 결과인데 이것은 초성의 19개 패턴에 대해서 이동과 잡음을 첨가한 패턴으로 인식한 결과이다. 두 실험을 비교해 보았을 때, 이동된 패턴에 대해서는 콘벌루션을 이용한 실험 2에서 인식률이 향상되었음을 알 수 있다. 즉, 우측으로 이동된 패턴을 인식한 경우에 실험 2는 실험 1에서 오인식된 문자의 약 62.5%를 제대로 인식하고 있으며, 우측상단으로 각각 한 화소씩 이동된 패턴에 대해서는 39%를 인식하고 있음을 알 수 있다.

인식 속도에 있어서도 중성1(2명역)과 중성2(4명역)와 같이 분류 패턴이 적은 경우에는 거의 비슷한 속도로 수렴하지만, 초성과 같이 분류하여야 할 패턴이 많은 경우에는 표4에서 볼 수 있듯이 실험 2에서 더 빠르게 수렴하고 있음을 알 수 있다.

표4. 학습속도 비교

	학습갯수 (수렴속도)		
	초성	중성1	중성2
실험 1	5721	5021	5021
실험 2	3580	5082	5974

V. 결 론

본 논문에서는 문자 인식시 문제점으로 지적되는 잡음 및 변형으로 인한 오인식을 감소시킬 수 있는 학습 방법을 제안하였다. 그리고, 이를 적용한 한글 문자 인식 신경망을 구현하였으며 기존의 신경망으로부터 나온 학습 결과와 비교해 보았다.

신경망은 초성, 중성1, 중성2, 종성에 대한 각각의 자소별 신경망으로 구성되며, 잡음에 대한 학습 능력을 향상시키기 위하여 이진으로 표현되어 있는 표준 학습 패턴에 2차원 콘벌루션을 적용하여 각각의 자소별 인식 신경망을 학습시켰다. 그 결과, 미소한 신경망에 대해서는 기존의 학습방법과 거의 같은 인식률을 보이지만, 이동 및 변형등으로 왜곡된 패턴에 대해서는 콘벌루션을 이용한 학습 방법이 인식률과 인식 속도에 있어서 향상되었음을 알 수 있었다. 이것은 전산언공지를 이용하므로 내재 잡음을 제외한 변환 잡음과 매체 잡음에 대해서만 분석한 결과이다.

본 논문에서는 학습 패턴에 가우스 함수에 대한 2차원 콘벌루션을 행할 때, 3 x 3 가중치 윈도우를 이용하였는데, 좀 더 폭이 넓은 형태의 가우스 함수를 적용함으로써 잡음에 대한 적응력을 향상시킬 수 있을 것으로 예상된다. 따라서, 본 논문에서 제시한 학습 방법을 기존의 학습 방법과 병행하여 사용함으로써 효율적인 패턴 인식을 위한 연구가 계속적으로 이루어져야 할 것이다.

참 고 문 헌

1. 이주근, "인공지능", 청문각, 1990.
2. "89 신경의로망 응용 Workshop", 한국전자통신연구소 한국정보과학회, 1989.
3. Yoh-Han Pao, "Adaptive Pattern Recognition and Neural Networks", Addison-Wesley Publishing Company, Inc., 1989.
4. Philip D. Wasserman, "Neural Computing (theory and practice)", Van Nostrand Reinhold, 1989.
5. 홍희경, "한글인식 전처리 과정에 있어서의 잡음제거 연산에 관한 연구", 이화여자대학교 대학원 석사학위논문, 1988.
6. 조성배, 김진형, "인쇄체 한글문자의 인식을 위한 계층적 신경망", 한국정보과학회논문지, Vol.17, No.3., pp.306-316, May, 1990.
7. 하지용, "전산 원고지름 이용한 한글 문자 인식", 이화여자대학교 대학원 석사학위논문, 1989.
8. "Neuro Shell", Ward Systems Group, Inc., 1988.
9. Branko Soucek and Marina Soucek, "Neural and Massively Parallel Computers", John Wiley & Sons, Inc., 1988.
10. "Pattern 인식을 위한 효율적인 Neural Network Model에 관한 연구 : 기존 Model들의 특성 비교분석", 한국전자통신연구소, 1989.