

연속 필기 패턴 인식을 위한 세그먼트 재조합 기반 통합 신경망 모델

장경익^{*}, 류정우, 박성진, 김명원
충실대학교 컴퓨터학부

Integrated Neural Networks Model for Handwritten Pattern Recognition using Segment Recombination

Kyoung Ik Jang, Joung Woo Ryu, Sung Jin Park, Myung Won Kim
School of Computing, Soongsil Univ

요 약

단일 문자 인식과 달리 연속 필기 패턴의 인식은 근본적인 필기 패턴의 형태적 특성을 충분히 고려할 필요가 있으며 다양한 형태의 패턴에 대한 특징이나 정보를 사용하여 종합적으로 판단 할 수 있는 모델의 유연성이 요구된다. 신경망의 학습 기능은 패턴의 왜곡과 잡음 등에 크게 영향을 받지 않으면서 인식에 필요한 특징의 추출이나 패턴 부류에 해당하는 노드의 반응을 스스로 학습시킬 수 있고, 다양한 형태의 정보를 쉽게 통합할 수 있는 유연한 구조를 제공한다. 퍼지 이론(Fuzzy theory)은 일정한 규칙이나 수학적 모델로 표현하기 어려운 패턴의 애매한 특징을 모델링 할 수 있기 때문에 인식 대상의 총체적 특징을 추출해 신경망에 효과적으로 적용할 수 있다. 본 논문에서는 연속 필기 숫자 패턴을 인식을 위한 신경망과 퍼지 이론을 이용한 통합 신경망 모델을 제안한다.

1. 서론

문자인식은 꾸준한 기술 발달로 우편번호인식, 제품 검사나 분류를 위한 기호인식, 도면인식, 팩시밀리를 통한 서식인식, 전표나 수표인식 등 여러 분야에서 실용화되어 사용되고 있다. 그러나 현재까지 상용화된 대부분의 오프라인 패턴 인식 시스템은 다양한 종류의 필기 패턴 및 그 변형 때문에 제한적으로 사용되고 있다. 이의 효과적인 인식을 위해서는 기존의 연구 범위를 유연성있게 연속 필기 패턴의 오프라인 인식으로 확대할 필요가 있다. 신경망을 이용한 방법 중에서 다층 퍼셉트론은 단일 필기 숫자 패턴 인식에 높은 인식율을 보여주고 있으며[1] 여러 응용 분야에서 좋은 성과를 거두고 있다. 본 논문에서는 기존의 패턴 추출 과정과 인식기에 퍼지 이론[2]을 도입하고 신경망의 학습기능을 이용해 연속 필기 패턴을 효율적으로 인식 할 수 있는 통합 신경망 모델을 제안한다.

이를 위해 우선 입력패턴을 세그먼트(segment)라는 인식 단위로 분할[3]하고 이들의 구조적 특징을 추출해 이들의 조합으로 패턴을 인식하게 된다.

2. 특징 추출

일반적으로 단일 숫자에서 사용되는 특징 추출 방법은 각 화소에서 기울기를 구한 다음 12개의 방향 성분으로 정량화(Quantization)를 한 후 패턴을 4x4로 분할화하여 각 영역에서 정량화된 값을 병행에 따라 그 개수를 히스토그램으로 나타내어 192개의 특징 벡터를 생성하고 있다[4].

본 논문에서 사용하는 특징추출 방법은 기존의 정량화 부분과 분할화 부분에 퍼지이론을 접목시킨 방법으로서 인식을 함상을 도모하였다. 예를 들면 그림 1-a)의 가)에서 "2"번 영역은 16~45도 까지 할당되어 있다 만약 45도라면 "2"번 영역에 속하지만 46도라면 "3"번 영역에 속하게 되어 단지 1도 차이에 다른 방향을 나타내는 특징 값이 되어 버린다. 또한 패턴을 분할하여 히스토그램화하는 데에도 마찬가지로 분할된 각 영역의 경계 부분에서 1개의 최소 자이로 다른 영역에 속하게 되어 다른 특징을 갖게 된다. 이처럼 정량화와 히스토그램화는 영역 경계 부분에 있어서 민감한 반응을 보이는 문제점을 가지고 있다. 따라서 이러한 경계부분의 민감성을 완화시키기 위해 정량화에서는 퍼지 기울기(fuzzy

gradient)를, 분할화에서는 퍼지 블로킹(fuzzy blocking)[5]을 사용하여 보다 일반적인 특징 벡터를 추출할 수 있게 하였다.

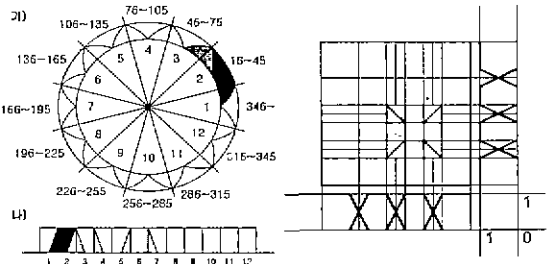


그림 1-a) 퍼지 기울기

그림 1-b) 퍼지 분할화

2) 퍼지 기울기와 퍼지 분할화에 대한 단일 숫자 인식 실험

실험1은 기존의 특징 추출방법을 그대로 사용하여 실험한 것이고, 실험 2는 퍼지 히스토그램을 <그림2>의 (가)와 같은 소속함수를 사용하여 실험한 것인데, 그림에 나타나 있는 것은 X축에 대해서만 나타난 것이다. 기울기 정보 이미지 상의 1~4는 분할화를 하는 구역을 나타낸 것이

은닉 노드	실험 1	실험 2	실험 3	실험 4	실험 5
70	93.7	94.85	94.15	93.75	94.8
75	93.95	94.95	94.25	94.1	95.1
80	94.1	95.0	94.85	94.35	95.25
85	94.1	95.25	94.85	94.45	95.3
90	94.2	95.3	95.05	94.75	95.35
95	94.3	95.5	95.05	94.8	95.35
100	94.45	95.65	95.5	94.8	95.8

표 1. 퍼지 분할화과 퍼지 기울기를 사용한 실험

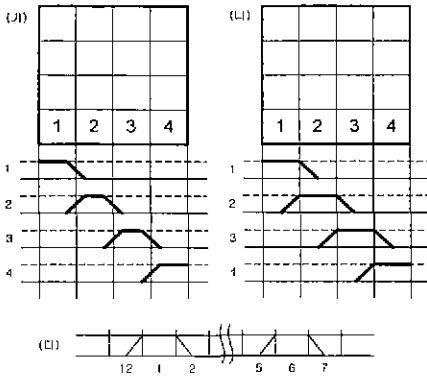


그림 2. 퍼지 히스토그램과 퍼지 기율기에서의 소속함수

고, 그 밑의 그래프는 그 구역의 소속함수를 나타내고 있다 Y축에 대한 소속함수도 X축과 같다 실험3은 퍼지 히스토그램을 <그림2>의 (나)와 같은 소속함수를 사용하여 실험2와 같이 실행한 것이고, 그리고 실험4는 퍼지 기율기를 <그림2>의 (다)와 같은 소속함수를 사용하여 실행한 것이다 마지막으로 실험5는 퍼지 히스토그램과 퍼지 기율기를 모두 사용하여 실행한 것인데, 실험 2에서 사용한 <그림2>의 (가)소속함수와 (다)의 소속함수를 사용하였다

결과를 보면 퍼지 개념을 사용한 모든 실험들이 아무것도 쓰지 않은 실험1보다 테스트 패턴들에 대해 결과가 좋게 나왔고, 특히 퍼지 히스토그램과 퍼지 기율기를 모두 사용한 실험5가 평균적으로 가장 좋은 결과를 나타내었다

3. 인식기 구조

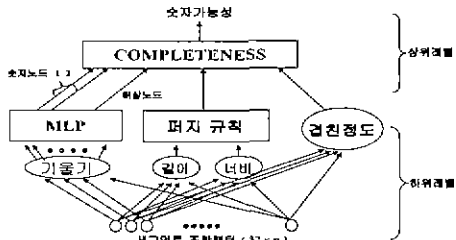


그림 3. 인식기의 구조

세양한 인식기의 하위레벨에서는 입력패턴으로부터 세그먼트를 추출하여 하나이상의 세그먼트들을 결합시켜서 세그먼트 조합패턴을 생성시키고 생성된 이러한 패턴이 숫자인지를 판별하기 위해 다음과 같이 세가지의 척도를 고려하였다. 첫 번째로는 허상을 학습한 신경망을 이용하여 세그먼트 조합패턴이 허상인지 숫자인지를 판별하였고, 두 번째 특징은 퍼지 규칙을 이용하여 신경망에서 고려되지 못하는 세그먼트 조합패턴의 크기로 숫자가가능성을 판별하였으며, 마지막으로 각 세그먼트 조합패턴간의 관계를 보여주는 특성으로서 겹치는 정도를 고려하였다 상위레벨에서는 현 개이상의 세그먼트 조합패턴으로 구성된 그룹에 하위레벨에서 나타낸 각 세그먼트 조합패턴의 숫자가가능성을 이용하여 의미 부여하고 있다

1) 신경망

세양한 인식기에서 사용하고 있는 신경망으로는 난립문자에 대해서 인식율이 높은 다층 퍼셉트론(Multi-Layer Perception)을 사용하였다 여기서 사용하고 있는 MLP의 구조는 입력노드 192개, 은닉노드 100개, 출력노드는 허상노드[6] 1개를 포함해 11개로 되어 있으며 학습은 CENPARMI 숫자레이더와 허상데이터로 학습을 하였고 '특징'으로 '학

습 패턴을 32*n으로 1차 정규화(Normalization)한 이미지를 32*32로 다시 2차 정규화로 부터 얻어진 이미지의 퍼지 기율기값을 사용하였다 출력층의 11개 노드들의 출력값 중에서 숫자노드 중 가장 높은 값 두 개와 허상노드의 출력값 인이 상위레벨에서 숫자가가능성을 부여하는데 적용하고 있다

2) 퍼지 규칙

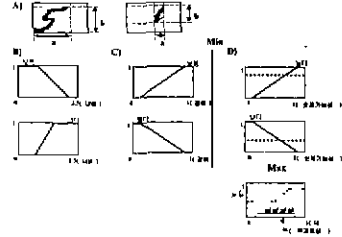


그림 4. 퍼지 소속함수

신경망은 패턴의 위상적 정보인 기율기를 학습 특징으로 하기 때문에 필기체 문자에 대한 다양한 크기를 인식할 수 없다. 이러한 결점을 보완하기 위해서 퍼지이론을 사용하여 이미지 크기의 가로 세로 비율을 고려함으로써 숫자의 가능성을 나타냈다. 사용된 커리의 변수의 언어학의 집합과 그에 따른 소속함수는 <그림4>과 같이 정의하였다.

3) 겹친정도(Overlap)

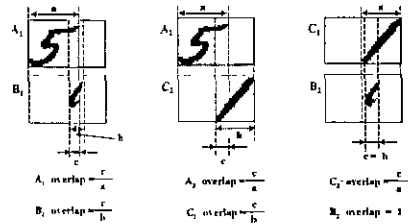


그림 5. 겹친정도

하위레벨에서 숫자가가능성을 표현하기 위해 사용하고 있는 마지막 척도로서 세그먼트 조합 패턴간의 관계를 고려한 겹친정도를 사용하였다 즉, 세그먼트 조합 패턴 내의 한 세그먼트와 나머지 세그먼트들 사이의 겹치는 정도 중에서 가장 큰 값을 그 세그먼트의 겹친정도로 정의하였다

예를 들면 아래의 <그림5>에서와 같이 그룹에 속해 있는 조합패턴 A, B, C 각각의 겹친정도 Ao, Bo, Co는 다음과 같다

$$A_o = \text{Max}(A_1, A_2), B_o = \text{Max}(B_1, B_2), C_o = \text{Max}(C_1, C_2)$$

이러한 정의는 <그림5>에서 B의 이미지는 앞에서 정의한 두 가지 척도로는 숫자 "1"로 오인식 하게된다 그러나 겹친정도를 고려하면 이미지 C에 의해서 이미지 B는 숫자가가능성이 낮아진다. 왜냐하면 겹친정도가 크면 숫자의 가능성이 낮아지고 반대로 작으면 숫자의 가능성이 높아지기 때문이다

4) Completeness

지금까지 하위레벨에서 세그먼트 조합패턴들의 숫자 가능성을 세 가지 척도-신경망, 퍼지 규칙, 겹친 정도-에 의해서 나타내었다. 상위레벨에서는 이러한 각각의 패턴의 가능성을 가지고 다음과 같은 가중치의 합(weighted summation)을 사용하여 가장 최대값을 가지는 그룹을 숫자 모델로서 최종 인식을 수행하도록 하였다

$$\text{Completeness} = \frac{w_1(\text{max}_{i=1}^n \text{seg}_i) + w_2(\text{max}_{i=1}^n G_i) + w_3 \text{EuzzeOutput} - w_4(1 - \text{Overlap})}{w_1 + w_2 + w_3 + w_4}$$

<그림6>은 세양한 인식기에서 세그먼트 조합 패턴이 하위레벨에서 고너되는 숫자 가능성의 척도에 대해서 어떠한 값을 가지며 상위레벨에

n	세그먼트 조합패턴
n	세그먼트 조합패턴들의 개수
max	1 번째 패턴에 대한 신경망(MLP)에서의 숫자 출력 노드 중 가장 높은 값
send	1 번째 패턴에 대한 신경망(MLP)에서의 숫자 출력 노드 중 두 번째로 높은 값
G	1 번째 패턴에 대한 신경망(MLP)에서의 허상 노드 값
FuzzyOutput	1 번째 패턴에 대한 비지 규칙에 의해서 생성된 숫자 가능성의 정도
Overlap	1 번째 패턴에 대한 겹침 정도

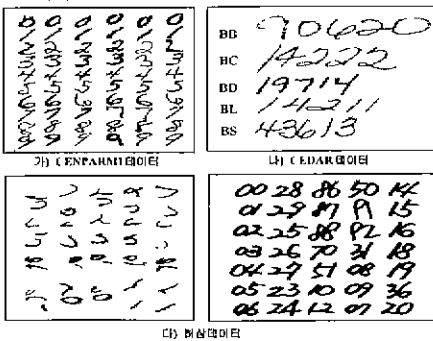
서 그룹들이 어떻게 값을 최종 인식을 수행하고 있는지를 보여주고 있다

	조합패턴	MLP		Garling	FuzzyOutput	Overlap	Completeness
		max	send				
그룹 1	5	1.00	0.01	0.00	0.9412	0.2632	0.920331
	4	0.93	0.02	0.19	1.0	0.3333	
그룹 2	5	1.00	0.01	0.00	0.9412	0.2632	0.863265
	1	0.97	0.06	0.00	1.0	0.3333	
	1	0.79	0.03	1.00	0.6011	1.0	

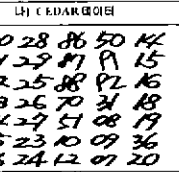
그림 6. completeness의 예

3. 실험 및 분석

1) 실험 데이터



(가) CENPARMI 데이터



(나) CEDAR 데이터

그림 7. 학습 데이터

05 03 12 03 04 19	44 28 200 880 40 38	
06 22 21 0002 22 00		
88 35 16 20 32 26 05		
89 20 59 23 23 30 05		
87 20 36 95 20 23 02		
00 20 40 55 54 46 19		
73 20 43 37 55 50 22		
60 20 44 52 55 40 23		
인식된 패턴들		오인식된 패턴들

그림 8. 테스트 데이터

본 논문에서 제안한 인식기의 학습데이터로는 CENPARMI 와 CEDAR 데이터[7]와 <그림7>의 다)의 같이 세그먼트 조형시 생긴 수 있는 비숫자(non-digit)를 허상데이터로 만들어 학습시키고 있다

인식율을 측정하기 위한 테스트 데이터로는 CEDAR 데이터베이스에 세그먼트가 6개 미만이고 두 개 이상의 숫자가 연결되어진 데이터를 사용하였다.

2) 인식 실험

	Garbage+ A+B Set	T Set 2000
75	100	93.05
80	100	93.35
85	100	93.45
90	100	93.65
95	100	93.75
100	100	93.8

표 2. 인식기에 사용된 MLP실험

<표2>는 인식기에서 사용된 MLP의 실험으로 은닉노드개수 100이고 학습 상수가 0.01 모멘텀함은 0.1일 때 가장 높은 값 인식율을 나타냈다. 최종적으로 completeness 척도를 사용한 상위레벨 인식에서 세그먼트 조합패턴들의 숫자가능성을 하위레벨에서 받아 가중치의 합으로 그 그룹에 의미를 부여하는데 사용되는 가중치는 실험을 통해 경험적 (heuristic)으로 주어진다.

본 실험은 앞에서 학습된 신경망을 이용하여 제안한 인식기의 인식율을 CEDAR 데이터 베이스 중 <그림8>과 같은 61개의 실험 데이터로 실험하였고 실험결과는 61개의 데이터 중 6개를 오인식 하여 인식율은 90.16%로 나타났다

패턴들을 잘못 인식되는 원인을 살펴보면 하위레벨에서 패턴에 대한 숫자 가능성이 잘 못 처리되는 경우와 상위레벨에서 패턴들로 구성된 그룹의 의미를 잘 못 부여하는 경우로 나눌 수 있다 따라서 본 실험에서 제안한 인식기가 잘 못 인식되는 패턴들을 위의 두 경우로 나누어 보면, 6개중 5개의 패턴이 하위레벨에서 잘 못 처리되어 인식을 하지 못한 경우이다 예를 들어 28의 패턴에서 2를 제안한 인식기에서 MLP가 2로 처리하지 않고 허상패턴으로 처리하였기 때문에 28은 허상패턴과 8로 잘 못 인식하고 있다 다른 1개의 패턴은 상위레벨에서 잘 못 의미를 부여하므로서 생기는 오인식 패턴이다. <그림8>에서 패턴 200은 하위레벨에서는 정확하게 숫자 가능성이 처리되었으나 상위레벨에서 의미를 잘 못 부여하므로서 허상패턴과 2와 0으로 잘 못 인식하고 있다

4. 결론 및 향후 연구방향

본 논문에서 제안한 모델은 특징 추출 및 인식 시스템에 퍼지이론을 도입한 통칭 신경망 모델로서 기존의 신경망을 이용한 단일 숫자 인식 시스템을 연속 필기 패턴 인식 시스템으로 확장하여 인식기를 구성하고 있다 특징 추출을 위해 퍼지 기율기와 퍼지 분할회는 새로운 특징 추출 방법으로 기존의 단일 숫자 인식보다 높은 인식율을 보여주고 있으며 앞으로 새로운 특징 추출방법으로 적용하여 사용되리라 기대된다

향후 연구 방향은 보다 실용적이고 많은 데이터에 대한 실험을 추가 하여 인식기의 일반성을 검증하고, 현재의 세그멘테이션 과정에 사용된 인위적 수행을 배제하여 인지적 처리 기반의 자동화된 세그멘테이션을 구현하는 것이다

5. 참고문헌

- [1] Ching Y Sunen, "Computer Recognition of Unconstrained Handwritten Numerals", Proceedings of the IEEE, VOL 80, NO 7, July 1992
- [2] 이상형, 오길목 "퍼지이론 및 응용 (I 권 및 II 권)", 홍공과학출판사 1992
- [3] 최병민, 김병원, 유권자 알고리즘을 이용한 모델기반 상위레벨 패턴인식, 한국정보과학회 학술발표논문집 (B), 24, 1, pp 333-336, 1997.
- [4] Geetha Srikanth, "Gradient representation for handwritten character recognition", the Third International Workshop on Frontiers in Handwriting Recognition, pp 153-162, Buffalo, New York, USA, 1993 pp 318-323
- [5] J.-S R Jang, C.-T.Sun, and E Mizutani, "Neuro-Fuzzy and soft computing", Preuce-Hall International, Inc, 1997 pp. 436-441
- [6] Paul D Gader, James M Keller, Raghu Krishnapuram, Jung-Hsien Chiang Magd A Mohamed, Neural and Fuzzy Methods in Handwriting Recognition, IEEE Computers, pp. 79-86, 1997
- [7] Jonathan J Hull, "A Dalabase for Handwritten Text Recognition Research", IEEE Transactions on Pattern Analysis Machine Intelligence, April, 1993