카오스 신경망을 이용한 서체 숫자 인식

조 재 흥, 성 정 원, 최 한 고, 김 상 희, 이 상 재
금오공과대학교 전자공학부
FAX : 0546-467-4270, E-Mail : hgchoi@knut.kumoh.ac.kr

Recognition of Unconstrained Handwritten Numerals using Chaotic Neural Network

Jae-Heung Cho, Jung-Won Sung, Han-Go Choi, Sang-Hee Kim, Sang-Jae Lee
School of Electronics, Kumoh National University of Technology

Abstract : Several neural networks have been successfully used to classify complex patterns such as handwritten numerals or words. This paper describes the discrimination of totally unconstrained handwritten numerals using the proposed chaotic neural network (CNN) to improve the recognition rate. The recognition system in the paper consists of the preprocessing stage to extract features using Kirsch mask and the classification stage to recognize numerals using the CNN. In order to evaluate the performance of the proposed network, we performed the recognition with unconstrained handwritten numeral database of Concordia university, Canada. Experimental results show that the CNN based recognizer performs higher recognition rate than other neural network-based methods reported using same database.

1. 서론

최근에 숫자인식과 같은 복잡한 패턴분류를 위해 비선형 처리능력과 학습능력을 갖고 있는 신경망이 성공적으로 사용되어 왔다[1]. 지금까지 패턴인식에 사용되어온 신경망들은 네트워크의 구조, 뉴런의 활성화 함수, 유닛간의 연결강도 갱신을 위한 학습규칙 등에 따라 여러 가지 형태의 모델이 연구되었다. 신경망은 입력과 출력간의 패턴 대응관계에 따라서 정적과 동적 신경망으로 구분할 수 있다. 순방향 다층 신경망은 정해진 입출력간의 정적 맵핑 구현에 적합하므로 실제적인 데이터의 인식에 있어서 부분적인 해답만을 제공할 뿐 유사한 패턴에 대해서는 좋은 학습 능력을 보여주지 못하므로 상대적으로 단순한 패턴분류에 사용되고 있다[2-4]. 따라서 복잡한 패턴인식 및 분류를 위해서는 시간지연 신경망이나 리커런트 신경망과 같은 동적 신경망이 사용되고 있다. 이러한 신경망은 데이터의 temporal 정보를 처리하기 위한 메모리로 궤환 루프나 지연소자가 사용되어 동특성 구현 능력을 향상시키기 때문에 복잡한 패턴인식을 효과적으로 수행할 수 있다.

본 논문에서는 신경망을 이용한 무제약 서체 숫자(unconstrained handwritten numerals)의 인식을 다루고 있는데, 일반적으로 서체 숫자는 사람에 따라서 글씨체가 서로 다르며, 각 숫자의 크기, 모양 등이 다양하다. 그리고 사용되는 필기도구 및 매체에 따라서 숫자의 형태 또한 매우 다양하므로 이러한 서체 숫자를 인식하기란 쉽지 않다. 따라서 이러한 숫자의 인식을 위하여 리커런트 신경망 등의 동특성이 우수한 신경망을 이용한 패턴 분류기가 사용되어 왔다. 그러나 숫자인식과 같이 여러 패턴을 분류하기 위해 기존의 시그모이드 활성함수을 갖는 한 개의 출력단을 사용할 경우 활성함수의 형태에 따라 출력단의 포화영역의 표현에 한계가 있으므로 효율적인 패턴인식이 이루어지지 않는다. 따라서 대부분의 신경망 출력층에 다수의 출력 유닛을 갖는 구조를 사용하거나 또 다른 패턴 분류기도 동시에 사용하여 인식율을 높히고 있으므로 신경망 및 분류기 구조가 복잡하게 된다.

본 논문에서는 지금까지 사용되어 왔던 신경망에 의한 서체 숫자의 인식율을 보다 향상시키기 위해 기존의 신경망에 비해 동특성이 더 우수한 카오스 신경망을 이용한 숫자패턴 인식을 다루고 있다. 제안된 카오스 신경망은 각 유닛의 자체 궤환 루프를 갖고 있으며 다른 뉴런들과도 상호 연결되어 있어서 고차의 비선형 동특성을 갖고 있으므로 복잡한 패턴인식에 효과적으로 사용할 수 있는 구조이다. 일반적으로 패턴인식 시스템은 입력신호의 유용한 특성추출을 위한 전처리과정과 정확한 패턴인식을 위한 분류과정으로 나누어 수행된다. 본 논문에서 사용된 전처리과정은 원 숫자에 Kirsch 마스크를 사용하여 특징을 추출하였으며, 이러한 특징정보를 입력으로

사용한 카오스 신경망에서 숫자를 인식한다. 실험 데이터는 Concordia 대학의 데이터베이스를 사용하였으며 실험결과에 의하면 제안된 신경망을 이용한 인식율은 98.0%로서 같은 데이터 베이스를 이용하여 수행된 기존의 방법보다 더 높은 인식율을 보여줌으로써 본 논문에서 사용된 카오스 신경망이 복잡한 패턴인식에 유용함을 확인할 수 있었다.

2. 패턴 인식

2.1 데이터 베이스

이 논문에서 사용된 데이터베이스는 미국 전역에 있는 우체국에서 발송되지 않은 우편물 가운데 6000여 개의 우편번호를 캐나다의 콘코디아 대학에서 데이터 베이스화 시킨 것이다. 데이터베이스 숫자들은 대략 166 PPI의 농담을 가지면서 0.153 mm의 화소크기를 갖고, 64×224의 영상크기로 이치화 되어 있다.

이들 데이터 중 4000 개의 데이터는 학습 데이터로 2000 개는 시험데이터로 사용하였다. 그림 1은 데이터베이스에서 발췌된 샘플 데이터를 보여주고 있다. 그림에서 확인할 수 있듯이 각 숫자들의 크기, 폭, 기울기, 그리고 형태 등이 글씨를 쓰는 스타일에 따라서 서로 다름을 알 수 있다.

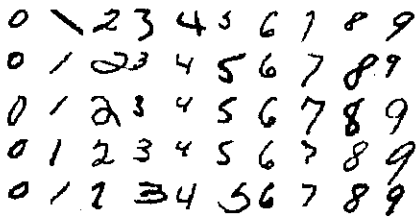

그림 1. 샘플 데이터

2.2 데이터 영상 추출

신경망의 입력데이터로서 숫자 영상을 그대로 사용할 경우보다 숫자 영상에 대한 사전 정보를 이용할 경우 즉, 입력 영상으로부터 특징을 추출하여 신경망의 입력으로 사용할 경우 신경망의 인식 성능이 향상된다. 따라서 인식에 필요한 중요한 정보의 손실을 최소화하는 특징 추출과정을 통해 특징 벡터를 추출할 필요가 있다. 본 연구에서는 영상 에지 추출방법으로 Kirsch 마스크를 사용하였다. 이 방법은 마스크 연산 시 주위 8 이웃 모두를 고려함으로써 다른 마스크에 비하여 4방향에 대한 보다 정확한 에지를 검출하므로 숫자영상의 에지 검출에 적합하다. 영상으로부터 에지를 검출하기 위한 Kirsch 방법은 다음과 같은 정의된다[5].

$$G(i,j) = \max_{k=0}^{7} [\,|\,5S_k - 3T_k\,|\,] \quad (1)$$

$$S_k = A_k + A_{k+1} + A_{k+2} \quad (2)$$

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7} \quad (3)$$

여기서 $G(i,j)$는 i번째 행과 j번째 열의 화소(Pixel)의 기울기 값이고, A는 8진수로 계산된 나머지 값이며, A_k (k=0,1, ... ,7)는 그림 2에서 보여준 것처럼 특정 화소(i,j)의 8개 근방(Neighbors) 화소들이다. 각 화소에 곱해지는 숫자의 의미는 방향성 기울기의 Scale factor를 나타낸다. 이는 원 영상에서 방향성으로 추출하고자 하는 부분의 화소와 그렇지 않은 부분의 화소에 5 와 -3이라는 값을 주어 각 숫자의 방향정보를 추출하게 된다.

본 연구에서는 크기가 모두 다른 Run Length Coding으로 된 원 데이터를 이치화 시켜 원래 크기의 영상으로 복원하여 압축된 영상을 시각적으로 확인한 후 다시 정규화된 크기로 재구성하여 각 숫자의 특징을 추출하였다[3]. 이 과정에서 원래의 영상 정보를 최대한 보존하기 위하여 각각의 영상을 전부 역치 처리(Thresholding)를 하였다. 입력 데이터의 영상은 16×16 화소의 크기로 일정하게 하였고, 각 영상의 4가지의 방향성 영상 벡터 즉, 수평, 수직, 좌, 우 방향에 대한 국부적인 특징 벡터 추출을 위해 식(4),(5),(6),(7)과 같이 방향정보를 계산하였다[3].

$$G(i,j)_H = \max\,(\,|\,5S_0 - 3T_0\,|,|\,5S_4 - 3T_4\,|\,) \quad (4)$$

$$G(i,j)_V = \max\,(\,|\,5S_2 - 3T_2\,|,|\,5S_6 - 3T_6\,|\,) \quad (5)$$

$$G(i,j)_R = \max\,(\,|\,5S_1 - 3T_1\,|,|\,5S_5 - 3T_5\,|\,) \quad (6)$$

$$G(i,j)_L = \max\,(\,|\,5S_3 - 3T_3\,|,|\,5S_7 - 3T_7\,|\,) \quad (7)$$

4쌍으로 이루어진 Kirsch 마스크를 통하여 16×16으로 정규화 된 입력 영상으로부터 추출된 방향 정보는 평균화 연산을 통하여 다시 4×4×4의 국부적인 특징 벡터로 압축된다. 여기서 평균화 연산은 2×2 화소를 한 개의 화소로 압축시키는데 4화소의 값을 모두 더해서 4로 나눈 값을 압축된 화소의 값으로 할당하는 연산이다. 무제약 필기체 숫자 인식에 있어서 이러한 국부적인 특징만을 이용할 경우, 입력된 문자의 전역적인 특징을 고려할 수 없으므로, 이를 위하여 정규화 된 16×16 영상을 직접 4×4 영상으로 압축하여 이를 전역 특징정보로 사용하였다. 따라서 본 논문에서는 모두 5×4×4 개의 특징 벡터(4×4×4 개의 국부적인 특징 벡터와 1×4×4 개의 전역적인 특징벡터)가 신경망의 입력으로 사용되어 숫자 인식을 수행하게 된다. 방향성으로 추출된 각 영상들의 화소는 원래 0에서 256까지의 값을 가지고 있는 Gray scale 영상인데, 신경망의 학습 데이터로 사용하기 위하여 각 영상(4×4 화소)이 갖고 있는 화소 중 가장 큰 값을 각각의 화소에 나눈 뒤 그 화소 값을 신경망의 입력 데이터로 사용하게 된다. 따라서 각 데이터는 0과 1사이의 값을 갖게 된다.

그림 3은 국부적인 특징 벡터 추출에 사용되는 Kirsch

마스크를 보여주며, 그림 4는 숫자 '0'의 국부적인 특징 벡터와 전역 적인 특징 벡터의 예를 보여주고 있다.

A_0	A_1	A_2
A_7	(i,j)	A_3
A_6	A_5	A_4

그림 2. 화소(i,j)의 8개 이웃하는 화소의 정의

5	5	5	-3	-3	-3	3	5	5	-3	-3	-3
-3	0	-3	-3	0	-3	-3	0	5	5	0	-3
-3	-3	-3	5	5	5	-3	-3	-3	5	5	-3
(a)						(c)					
-3	-3	5	5	-3	-3	-3	-3	-3	5	5	-3
-3	0	5	5	0	-3	-3	0	5	5	0	-3
-3	-3	5	5	-3	-3	-3	5	5	-3	-3	-3
(b)						(d)					

그림 3. 방향성 영상추출을 위한 Kirsh 마스크
(a)수평방향, (b)수직방향, (c)우대각방향, (d)좌대각방향

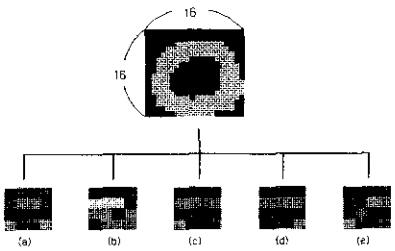

그림 4. 원 영상에서 추출된 압축 영상
(a)수평방향, (b)수직방향, (c)좌대각 방향, (d) 우대각 방향, (e)마스크를 사용하지 않은 압축 영상

3. 카오스 신경망의 구조

제안된 카오스 신경망(그림 5)은 기존의 카오스 신경망과는 달리 층내 뉴런들 사이에 단방향 연결을 갖는 결합계수를 2개씩 설정하였으며, 층간의 결합계수는 일반적인 다층 신경망의 경우와 같이 하위층에서 상위층으로의 단방향 결합계수를 갖도록 하였다. 카오스 신경망은 3층으로 구성되어 있으며 신경망의 결합계수들의 초기값은 임의로 선정되었다.

카오스 신경망 입력층의 i 번째 뉴런 입력

$$Z_i(n)=\left\{\sum_{k=0}^{N} I_i(n)\right\} \tag{8}$$

여기서 N은 외부입력 뉴런의 개수이다. 입력층의 뉴런은 비선형함수에 의해 처리되지 않고 그대로 히든층에 전달된다.

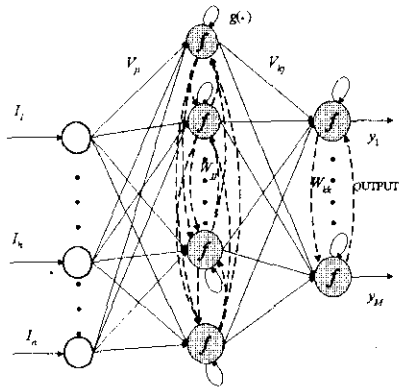

그림 5. 적용된 카오스 신경회로망의 구조

이러한 입력들에 대한 히든층의 j번째 뉴런의 내부 상태는 다음과 같다.

$$x_j(n+1)=k\cdot x_j(n)+\sum_{i=0}^{N} W_{ji}(n)Z_i(n+1) + \sum_{i=0}^{M} V_{ji}(n)f[x_i(n)]-f[x_j(n)] \tag{9}$$

여기서 $x_j(n)$는 출력층의 내부상태, N는 입력층의 뉴런개수, W_{ji}는 입력층과 히든층 사이의 연결계수, M은 히든층의 뉴런개수, V_{ji}는 히든층과 히든층사이의 연결계수, $f(\cdot)$는 시그모이드 활성함수를 나타낸다. 히든층의 j번째 뉴런 출력 $O_j(n)$는 다음과 같다.

$$O_j(n) = f[x_j(n)] = \frac{1}{1+e^{-s\cdot x_j(n)}} \tag{10}$$

히든층의 출력은 카오스 신경망 출력층의 입력으로 사용되는데, 출력층의 k번째 뉴런의 내부상태는 다음과 같이 나타낼 수 있다.

$$x_k(n+1)=k\cdot x_k(n)+\sum_{j=0}^{N} W_{kj}(n)O_j(n+1) + \sum_{r=0}^{P} V_{kr}(n)f[x_r(n)]-f[x_k(n)] \tag{11}$$

여기서 $x_k(n)$는 출력층의 내부상태, M는 히든층의 뉴런개수, W_{kj}는 히든층과 출력층 사이의 연결계수, P는 출력층의 뉴런개수, V_{kr}는 출력층의 뉴런 k와 출력층의 또 다른 뉴런 r사이의 연결계수를 나타낸다. 시그모이드 활성함수를 통과한 카오스 신경망의 최종 출력 $y_k(n)$은 다음과 같다.

$$y_k(n) = f[x_j(n)] = \frac{1}{1+e^{-s\cdot x_j(n)}} \tag{12}$$

제안된 구조의 카오스 신경망에 있어서 층간 결합계수와 층내 결합계수는 기존의 오차 역전파 알고리즘을 이용하여 학습하였다. 그리고 학습속도를 향상시키기 위해서 모멘텀 항이 추가되었다.

4. 실험 결과

카오스 신경망에 사용된 입력층과 히든층의 유닛수는 각각 80개이고, 출력층 유닛수는 1개이다. 상대적인 성능 비교를 위해 사용된 리커런트 신경망의 입력과 히든층의 유닛수는 160개(5x4x4=80개 특징벡터의 외부입력과 히든층으로부터의 80개 궤환입력)와 80개이며 출력층의 유닛수는 10개이다. 카오스 신경망의 학습율, 모멘텀율, 그리고 불응율은 각각 0.3, 0.6, 0.6을 사용하였으며, 리커런트 신경망의 학습율과 모멘텀율은 0.15와 0.6으로서, 이러한 변수 값은 최상의 인식율이 되도록 실험을 통해서 결정되었다. 학습 및 시험 데이터수는 두 신경망에 대해 각각 4000개와 2000개를 사용하였으며, 학습횟수는 더 이상 오차의 감소변화가 발생하지 않는 1500번을 학습하였고, 학습 후 시험 데이터에 대해 숫자인식을 수행하였다.

표 1은 학습과 시험 데이터에 대한 학습결과와 시험결과를 보여주고 있다. 표에서 RNN과 CNN은 각각 리커런트 신경망과 카오스 신경망을 나타낸다. 표에 의하면 카오스 신경망에 의한 인식율은 리커런트 신경망에 비해 학습 데이터에 대해서는 1.2%, 시험 데이터에 대해서는 2.2% 각각 향상되었다.

표 2와 3은 시험 데이터에 대한 confusion matrix로서 각 숫자에 대한 신경망의 인식율을 보여주고 있다. 두 신경망에 있어서 오판숫자의 분포가 매우 유사하므로 실험에 대한 신뢰성을 확인할 수 있으며 특히, CNN은 RNN에 비해 숫자 1과 2에서 인식율을 크게 향상시켰다. 실험결과를 요약하면 카오스 신경망은 리커런트 신경망에 비해 더적은 연결계수를 사용하면서도 인식율은 향상되었음을 확인할 수 있었다.

표 1. 학습 및 시험 데이터에 대한 인식율

	학습 data수	인식한 data수	인식율 (%)	시험 data수	인식한 data수	인식율 (%)
RNN	4000	3924	98.1	2000	1916	95.8
CNN	4000	3972	99.3	2000	1960	98.0

표 2. RNN의 각 숫자별 인식율

C	0	1	2	3	4	5	6	7	8	9	오판율 (%)	인식율 (%)
0	190			1			2	1	4	2	5.0	95.0
1	1	193	1		1			3		1	3.5	96.5
2		2	190	5		1			1	1	5.0	95.0
3	1		3	192			1		3		4.0	96.0
4	1	3		1	191	1		3			4.5	95.5
5	1		1	3		191	2		1	1	4.5	95.5
6	1	1				1	192		4	1	4.0	96.0
7		3	2		2			191		2	4.5	95.5
8	2	1		1			2		193	1	3.5	96.5
9	1	1		2				2	2	193	4.0	96.0
								평 균 치			4.2	95.8

표 3. CNN의 각 숫자별 인식율

C	0	1	2	3	4	5	6	7	8	9	오판율 (%)	인식율 (%)
0	195	1					2			2	2.5	97.5
1		198						1		1	1.0	99.0
2		1	196	2				1			2.0	98.0
3				197		1			2		1.5	98.5
4	1	2			195	1		1			2.5	97.5
5	1			1		195	2		1		2.5	97.5
6	1						196		2	1	2.0	98.0
7	1	2						195		2	2.5	97.5
8				2			1		197		1.5	98.5
9		1						2	1	196	2.0	98.0
								평 균 치			2.0	98.0

C : Class

5. 결 론

본 논문에서는 새로운 구조의 카오스 신경망을 이용하여 서체 숫자의 패턴인식을 다루고 있다. 제안된 카오스 신경망은 각 유닛에 자기 궤환과 인접 유닛과의 상호 연결된 구조로서 학습을 거듭할수록 뉴런자체의 과거값과 층내 결합에 의한 상호 커플링이 존재하므로 기존의 다른 신경망보다 동적 특성이 더 우수한 신경망이다. 무제약 서체 숫자 데이터베이스를 이용한 실험결과에 의하면 리커런트 신경망과 같은 기존의 다른 신경망에 의한 숫자 분류에 비해 제안된 신경망에 의한 숫자 인식률이 향상되었다[2][3]. 현재 사용된 신경망의 출력단은 단일 출력으로서 다출력을 사용할 경우 인식율을 더욱 더 향상시킬 수 있으며, 본 논문에서 사용된 카오스 신경망은 서체 글자 등과 같이 복잡한 다른 패턴인식에도 효과적으로 적용할 수 있으리라 판단된다.

참고문헌

[1] B. Bavarian, "Introduction to Neural Networks for Intelligent Control", *IEEE Control System Magazine*, pp. 3-7, Apr. 1988.

[2] Seong-whan Lee, Hee-heon Song. "A New Recurrent NeuralNetwork Architecture for Visual Pattern Recognition ", *IEEE Trans. on Neura Networks*, Vol. 8, No. 2. Mar. 1997.

[3] Sung-Bae Cho "Neural-Network Classifiers for Recognizing Totally Unconstrained Handwritten Numerals", *IEEE Trans. on Neural Networks*, Vol. 8, No. 1, Jan. 1997.

[4] R.J.Williams, D. Zipser "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks", *Neural Computation 1*, MIT. pp. 270-280, 1989.

[5] W. K. Pratt, *Digital Image Processing*, New york: Wiley, 1978.