

HOG 특징을 이용한 다항식 방사형 기저함수 신경회로망 기반 숫자 인식 방법의 설계

김봉연*, 오성권*
수원대학교 전기공학과*

Design of Digits Recognition Method Based on pRBFNNs Using HOG Features

Bong-Youn Kim*, Sung-Kwun Oh*
Department of Electrical Engineering, The University of Suwon*

Abstract - 본 논문에서는 HOG 특징을 이용한 다항식 방사형 기저함수 신경회로망 기반 숫자 인식 시스템의 설계를 제안한다. 제안한 숫자 인식 시스템은 HOG 특징을 이용하여 숫자를 입력 데이터로 사용하기 위해 특징을 계산한다. 다항식 방사형 기저 함수 신경회로망은 고차원 데이터의 임-출력 형태를 갖는 클래스를 분류하는데 용이하며, 활성함수의 중심점 및 분포상수는 Fuzzy C-Means(FCM) 알고리즘에 의해 초기 값을 설정한다. 또한 제안한 분류기의 최적화를 위해 Particle Swarm Optimization(PSO)를 사용하여 최적화된 분류기의 성능을 비교한다. 숫자 인식을 위하여 공인 데이터베이스인 MNIST handwritten digit database를 사용하여 분류기의 성능을 평가하고 분석한다.

1. 서 론

인간의 뇌신경 세포와 구조를 모방하는 신경회로망의 연구는 인공지능 분야에서 활발하게 진행되고 있다. 수기로 작성된 필기체 문자 및 숫자 인식은 수표, 서명, 우편번호 인식 등에서 다양하게 적용이 가능하여 패턴인식 분야에서 그 중요성이 더욱 부각되고 있다. 최근 Deep Learning 기법을 이용하여 인식을 개선할 위한 다양한 방법들이 연구되어 좋은 결과를 얻고 있지만, 구조가 복잡하고 고성능의 하드웨어를 필요로 한다는 단점이 있다. 또한 문자 및 숫자인식을 위한 데이터는 고차원으로서 많은 노이즈를 포함하고 있기 때문에 인식을 저하 시킬 수 있다. 본 연구에서는 이러한 단점을 보완하며 인식을 저하를 방지하기 위해 Edge성분의 벡터 기울기 및 크기를 HOG 특징으로 추출한다. 그리고 차원 축소 알고리즘인 주성분 분석법(Principal Component Analysis: PCA)을 이용하여 고차원의 숫자 이미지를 저차원으로 차원 축소하고, 이를 제안된 방사형 기저함수 신경회로망(Radial Basis Function Neural Networks: RBFNN) 패턴 분류기의 입력으로 이용한다. 또한 패턴 분류기의 은닉층은 가우시안 함수 대신에 Fuzzy C-Means(FCM) 알고리즘을 이용하여 구성한다.

2. 본 론

2.1 HOG를 이용한 특징벡터의 검출

본 논문에서는 제안된 숫자인식 시스템을 위한 이미지의 특징 추출을 위해 HOG(Histogram of Oriented Gradient)를 사용한다. HOG는 국소영역에 대한 밝기의 분포방향을 히스토그램화하여 이를 특징벡터로 나타내고자 하는 것에 사용된다. 이러한 HOG는 근접화소의 분포를 국소영역의 기하학적인 변화에 강건한 특징을 가지고 있다. 아래에 HOG를 이용한 특징 벡터를 계산하는 방법을 표현한다.



<그림 1> HOG 알고리즘의 흐름도

첫 번째 단계로 입력이미지에 대하여 감마 값 및 컬러 값을 Gray 변환한다. 이후 HOG 특징을 추출하기 위해 기울기 값을 계산하려면 영상의 각 픽셀로부터 y축과 x축 간의 변화의 크기와 y축과 x축 사이의 각도를 계산해야하는데 다음의 아래의 식을 이용하여 계산 할 수 있다. 식(1)의 $g_y(y, x)$ 는 y축에 대한 기울기에 관한 식이고, $g_x(y, x)$ 는 x축에 대한 기울기이며, 식(2)의 $M(y, x)$ 는 g_y 와 g_x 의 크기를 나타내고 식(3)의 θ 값은 x축을 기준 축으로 x축과 y축이 이루는 각도를 나타낸다.

$$g_y(y, x) = I(y+1, x) - I(y-1, x) \tag{1}$$

$$g_x(y, x) = I(y, x+1) - I(y, x-1)$$

$$M(y, x) = \sqrt{g_y(y, x)^2 + g_x(y, x)^2} \tag{2}$$

$$\theta(y, x) = \tan^{-1} \left(\frac{g_y(y, x)}{g_x(y, x)} \right) \tag{3}$$

2.2 PCA를 이용한 숫자 이미지의 차원 축소

주성분 분석법(Principal Component Analysis: PCA)은 본 논문에서 제안된 숫자 인식 패턴 분류기의 전처리 과정으로 이용한다. PCA는 대표적인 차원 축소 방법으로 숫자 이미지의 고차원 특징벡터를 저차원의 특징벡터로 정보 손실이 최소화 되도록 차원을 축소시킨다. 숫자 이미지의 크기가 $N \times N$ 이고, 이미지 개수가 M 개 일 때 PCA의 수행 단계는 다음과 같다.

[Step 1] 숫자 이미지의 벡터 집합 구성

$$X = \Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_N \tag{4}$$

[Step 2] 숫자 이미지의 평균 벡터를 계산

$$\Psi = \frac{1}{N} \sum_{i=1}^N X_i \tag{5}$$

[Step 3] 숫자 이미지의 벡터와 평균 숫자 이미지 벡터의 차를 계산

$$\Phi_i = \Gamma_i - \Psi \tag{6}$$

[Step 4] 공분산 행렬을 계산

$$C = \frac{1}{N} \sum_{i=1}^N \Phi_i \cdot \Phi_i^T = A \cdot A^T \tag{7}$$

[Step 5] 고유값 분석을 통해 공분산 C의 고유값 행렬 A과 고유벡터 행렬 U를 계산

$$C = UAU^T \tag{8}$$

[Step 6] N^2 개의 고유벡터에서 n 개의 고유값을 선택

[Step 7] 선택한 고유값에 대응되는 고유벡터 W를 생성

[Step 8] 차원 축소된 데이터 Y를 얻는다.

$$Y = W^T X \tag{9}$$

2.3 다항식 기반 RBFNN 패턴 분류기 설계

방사형 기저함수 신경회로망(Radial Basis Function Neural Networks: RBFNN)은 조건부, 결론부, 추론부로 구성된다. PCA 알고리즘을 통해 차원 축소된 숫자 데이터가 조건부의 각 노드에 연결된다. 일반적으로 활성함수로는 가우시안 함수를 사용하지만 제안된 패턴 분류기에서는 Fuzzy C-Means 클러스터링 방법을 사용하여 입력 데이터 특성을 반영

할 수 있도록 한다. 그리고 조건부의 각 노드에서 나온 출력과 결론부에서 구한 연결가중치를 이용하여 최종 출력을 구한다. 연결가중치는 기존의 상수항을 확장시킨 다항식 형태로 구성되고, 제안된 pRBFNNs의 구조는 그림 2와 같다.

Type 1 : 상수항(Constant)

$$f_j(x_1, \dots, x_k) = a_{j0} \quad (10)$$

Type 2 : 1차 선형식(Linear)

$$f_j(x_1, \dots, x_k) = a_{j0} + \sum_{i=1}^k a_{ji} x_i \quad (11)$$

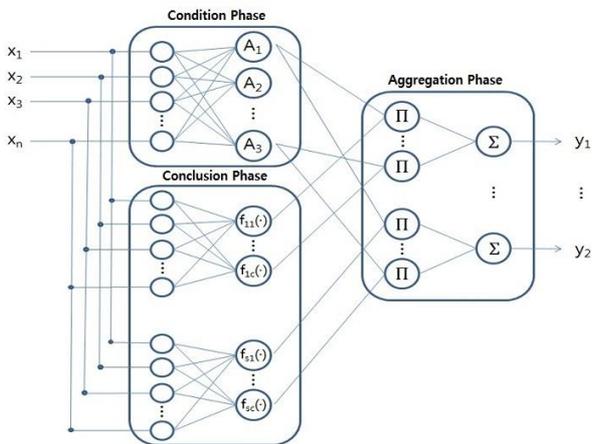
Type 3 : 2차 선형식(Quadratic)

$k = 2$:

$$f_j(x_1, \dots, x_k) = a_{j0} + \sum_{i=1}^k a_{ji} x_i + \sum_{i=1}^k a_{j(k+i)} x_i^2 + a_{(2k+1)} x_1 x_2 \quad (12)$$

$k \geq 3$:

$$f_j(x_1, \dots, x_k) = a_{j0} + \sum_{i=1}^k a_{ji} x_i + \sum_{i=1}^k a_{j(k+i)} x_i^2 + a_{(2k+1)} x_1 x_2 + \dots + a_{(k(k+3)/2)} x_{(k-1)} x_k \quad (13)$$



<그림 2> 제안된 패턴 분류기의 구조

4. 시뮬레이션 및 결과 고찰

본 논문에서는 MNIST 데이터를 아래의 그림 3과 같이 0부터 9까지 수기로 작성된 숫자들을 모아 놓은 데이터 집합으로써 하나의 이미지는 가로 28, 세로 28 픽셀로 구성하였다. MNIST 데이터의 수는 총 70000개이며, 각각 학습 및 테스트 데이터로 60000개, 10000개로 구성하였다. 그리고 입력 차원 수는 784개로 구성되어 있으며, 본 논문에서는 차원 축소 알고리즘인 PCA를 통해 고유 값이 0.1 이상인 값들만을 이용하여 784개의 입력 차원을 64개로 차원 축소하여 사용하였다. 제안된 분류기의 성능을 평가하기 위해 기존 논문에서 제안되었던 패턴 분류기의 성능과 비교한다.

표 2는 제안된 패턴 분류기와 기존 논문에서 제안된 패턴 분류기의 성능을 나타내고, 성능 평가 지수는 식 (14), (15)과 같이 오분류율을 이용하였다.



<그림 3> MNIST 데이터

$$\text{오분류율} = \frac{\sum_{i=1}^N f(y_i - \hat{y}_i)}{N} \times 100 \quad (14)$$

$$f(y_i - \hat{y}_i) = \begin{cases} 0, & y_i \neq \hat{y}_i \\ 1, & y_i = \hat{y}_i \end{cases} \quad (15)$$

표 2와 같이 제안된 분류기의 성능이 대부분의 다른 분류기 보다 우수함을 알 수 있다. Convolution net의 경우, 제안된 분류기의 성능보다 우수하지만 구조가 복잡하여 계산량이 많다는 단점이 있고, 제안된 분류기는 구조가 간단하여 계산량을 줄일 수 있다는 장점이 있다.

<표 2> 실험 결과

Classifier	Preprocessing	Test Error Rate (%)
Linear Classifier (1-layer NN)	None	12.0
Linear Classifier (1-layer NN)	Deskewing	8.4
1000 RBF + Linear Classifier	None	3.6
Convolution net LeNet-4	None	1.1
68PCA + 45 RBFNN	PCA	1.82
HOG + RBFNN (Proposed Classifier)	PCA	1.79

3. 결론

본 논문에서는 밝기에 대한 분포방향을 이용한 HOG알고리즘과 노이즈와 같이 고차원 숫자 데이터에서 발생하는 문제를 PCA 알고리즘을 이용하여 해결하고, 간단한 구조의 RBFNNs 분류기를 이용하여 숫자 인식 시스템을 설계하였다. 이와 같이 제안된 패턴 분류기는 기존에 제안된 다른 패턴 분류기에 비해 우수한 성능을 보였고, 데이터의 차원수도 축소하여 분류기의 계산량을 줄일 수 있었다.

감사의 글

본 연구는 한국 산업단지공단의 10차년도 산업집적지경쟁력강화산업계획의 생산기술사업화 지원사업으로 연구를 수행하였음 (NTIS-1415136442) 그리고 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음[GRRC 수원2015-B2, U-city 보안감시 기술협력센터]

[참고 문헌]

- [1] S. K. Oh, W. D. Kim, and W. Pedrycz, "Polynomial based radial basis function neural networks(P-RBF NNs) realized with the aid of particle swarm optimization," Fuzzy Sets and Systems, Vol. 163, No. 1, pp. 54-77, 2011.
- [2] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner "Gradient-Based Learning Applied to Document Recognition", IEEE, Vol. 86, pp. 2278-2324, 1998.