

방향각 특징 기반의 필기 숫자 인식

이상호, 김호연, 임길택, 남윤석

한국전자통신연구원 우정기술연구센터

e-mail : ileesh@etri.re.kr

Recognition of Handwritten Numerals Based on the Direction Angle Feature

Sang-Ho Lee, Ho-Yon Kim, Kil-Taek Lim, Yun-Seok Nam

Electronics and Telecommunication Research Institute

요 약

특징 추출은 입력 데이터를 인식이 더 잘 될 수 있도록 변환된 영역의 특징 벡터로 변환하는 과정으로 볼 수 있다. 특징벡터가 갖추어야 할 주요 특성은 손실되는 정보량이 가능한 적어야 된다는 것이다. 또한, 높은 인식률을 얻기 위해서, 동일 클래스에 포함된 특징 벡터의 편차는 적도록 만들어야 한다. 본 논문에서는, 방향각 누적 특징을 기반으로 개발된 몇 가지 새로운 특징을 필기 숫자 인식에 적용하였다. 특징을 추출하기 위하여, 입력된 이진 영상의 비선형 정규화, 영상의 크기에 의한 특징 정규화, 영상의 전경 영역에 의한 특징 정규화 등의 여러 가지 방법이 적용되었다. 실제 우편물에서 추출된 필기 숫자 데이터베이스를 실험에 사용하였으며, 제안된 방법이 필기 숫자 인식에 효과적으로 적용될 수 있다는 것을 결과에서 보여주고 있다.

1. 서론

지난 수십 년 동안 필기 숫자인식에 대해 많은 연구가 수행되어 왔다[1], [2]. 일반 자연어 인식에 비해 비교적 적은 수의 클래스가 인식 대상이 되기 때문에 필기 숫자 인식은 새로이 개발되는 인식 방법의 성능을 검증하는데 많이 채용되었다. 그러나, 필기자의 필기 습관에 따른 형태의 다양한 변화로 인해, 일반적으로 매우 높은 인식률을 얻기는 힘든 것으로 알려져 있다. 또한, 복잡한 문제를 해결할 수 있는 몇 가지 이론적인 모델도 다차원 공간에서 정의되는 인식 문제를 완벽하게 다루기는 힘들다고 할 수 있다.

기정의된 클래스를 표현하기 위한 랜덤 변수의 한 종류인 다차원 벡터는 불필요할 지도 모르는 많은 잉여 정보를 포함하고 있다. 그러한, 정보는 인식 문제를 더욱 어렵게 만든다. 또한, 다차원 데이터에서 중요한 특징에 집중할 수 있는 모델링 방법이 있지만, 입력 벡터의 차원이 점점 커진다고 해서, 입력 데이터를 잘 표현할 것이라고 단정할 수는 없다. 예를 들어, 작은 수의 입력 벡터로 학습된 인공 신경회로망이 큰

수의 입력 벡터로 학습된 것보다 입력 데이터의 모델링 및 인식률에서 앞서는 경우도 있다. 그러므로, 실용적인 목적으로 인식 시스템을 개발할 때는 원래의 입력 데이터를 잘 표현할 수 있는 중요 특징 벡터를 추출하는 것이 가장 중요하다. 본 논문에서는 실용적인 응용 분야와 중요 특징 벡터의 추출이라는 점에 주안점을 두고 필기 숫자 인식에 적합한 새로운 특징의 추출 방법을 제안하였다. 2장부터는 제안된 특징의 추출 방법을 설명하고, 3장은 실험 결과물, 그리고 4장에서 결론을 내리도록 할 것이다.

2. 특징 추출

2장에서는 필기 숫자 인식을 위한 특징 추출 방법에 대해 기술하고, 제안된 다차원 특징벡터의 차원을 줄일 만한 요소가 있는 지에 대해 제안된 방법을 조금씩 변형하여 보도록 하였다.

2.1 방향 각도 특징의 추출

먼저, 본 논문에서 제안되는 특징 추출 방법에서

주요한 요소로 사용되는 방향 각도 특징을 기술하기 위해, 임의의 한 점 $f(x, y)$ 에서 결정되는 방향벡터를 정의한다.

$$\nabla f = \left[\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right]^T = |\nabla f| \alpha(x, y) \quad (1)$$

$|\nabla f|$ 는 방향 벡터 ∇f 의 크기를 나타내며, $\alpha(x, y)$ 는 위치 (x, y) 에서 방향각도를 나타내는 것으로서 식(2)와 같이 주어진다.

$$\alpha(x, y) = \tan^{-1} \left(\frac{\sum_{a,b} b \cdot f(x+a, y+b)}{\sum_{a,b} a \cdot f(x+a, y+b)} \right) \quad (2)$$

식(2)에서, $-1 \leq a, b \leq 1$ 이다.

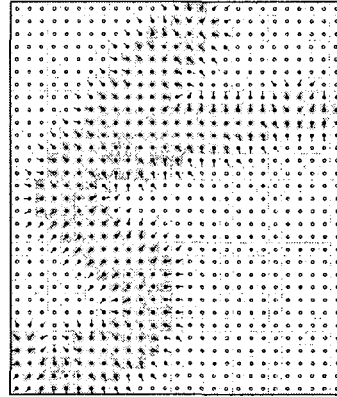
식(1)에서 정의된 방향 각도 $\alpha(x, y)$ 는 방향의 유무에 따라 8개 또는 4개의 방향을 가리키기 위한 값으로 코드화된다. 또한, 방향이 정의되지 않는 화소 자체도 인식에 도움을 줄 수 있는 정보를 제공할 가능성이 있으므로 특징 벡터에서 제외시킬 수는 없다. 그래서, <표 1>, <표 2>와 같이 방향이 존재하는 경우($|\nabla f| \neq 0$)와, 방향이 존재하지 않는 각 경우에 대해 코드화 방법을 정의하였다. 그리고, 방향이 없는 경우도 화소값에 따라서 두 개의 코드값으로 구분되도록 하였다.

<표 1> 방향 각도의 코드화 방법 I

방향각도	코드	화소
$\left[-\frac{\pi}{8}, \frac{\pi}{8} \right)$	0	
$\left[\frac{\pi}{8}, \frac{3\pi}{8} \right)$	1	
...	...	
$\left[\frac{13\pi}{8}, \frac{15\pi}{8} \right)$	7	
방향 없음	8	1
방향 없음	9	0

<표 2> 방향 각도의 코드화 방법 II

방향각도	코드	화소
$\left[-\frac{\pi}{8}, \frac{\pi}{8} \right)$ 또는 $\left[\frac{7\pi}{8}, \frac{9\pi}{8} \right)$	0	
$\left[\frac{\pi}{8}, \frac{3\pi}{8} \right)$ 또는 $\left[\frac{9\pi}{8}, \frac{11\pi}{8} \right)$	1	
$\left[\frac{3\pi}{8}, \frac{5\pi}{8} \right)$ 또는 $\left[\frac{11\pi}{8}, \frac{13\pi}{8} \right)$	2	
$\left[\frac{5\pi}{8}, \frac{7\pi}{8} \right)$ 또는 $\left[\frac{13\pi}{8}, \frac{15\pi}{8} \right)$	3	
방향 없음	4	1
방향 없음	5	0



(그림 1) 방향에 따른 특징 지도

- 방향 없는 흑화소(전경색)
- 방향 없는 백화소(배경색)

2.2 비선형 메쉬를 이용한 정규화

일반적으로 패턴 인식 분야에서 주로 사용되는 메쉬는 주어진 입력 영상을 임의의 크기로 재구성하여 균등한 면적을 가지도록 선형적으로 정의된다. 그러나, 본 논문에서는 주어진 이차원 영상 본래의 형태를 왜곡시키지 않도록 전경색의 밀도에 의해 비선형으로 메쉬를 구하는 정규화를 실시하였다. 비선형 메쉬는 다음과 같은 절차를 통하여 구한다.

1) 전경색인 흑화소 값에 기준하여 수평과 수직 각 방향의 히스토그램을 추출한다.

2) 식(3)과 같이 영상을 M개의 부분 영역으로 수평 분할한다. 이 때 각 부분 영역에는 동일한 개수의 흑화소가 존재하도록 하여야 한다.

$$\sum_{x=1}^{P_n} H_v(x) = \sum_{x=P_n+1}^{P_n} H_v(x) = \dots = \sum_{x=P_{n-1}+1}^{P_n} H_v(x) = \sum_{x=P_{n-1}+1}^W H_v(x) \quad (3)$$

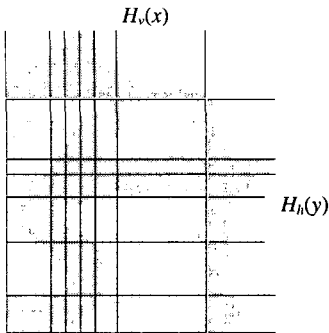
3) 식(4)와 같이 영상을 N개의 부분 영역으로 수직 분할한다. 이 때 각 부분 영역에는 동일한 개수의 흑화소가 존재하도록 하여야 한다.

$$\sum_{y=1}^{P_n} H_h(y) = \sum_{y=P_n+1}^{P_n} H_h(y) = \dots = \sum_{y=P_{n-1}+1}^{P_n} H_h(y) = \sum_{y=P_{n-1}+1}^H H_h(y) \quad (4)$$

식(3)과 식(4)에서 각 항목은 다음과 같다.

- $P_n = n$ 번째 부분 영역의 끝 위치
- $W =$ 영상의 가로 크기($W=P_M$)
- $H =$ 영상의 세로 크기($H=P_N$)
- $M =$ 수평 부분 영역의 개수
- $N =$ 수직 부분 영역의 개수
- $H_v(x) = x$ 번째 열의 흑화소 수
- $H_h(y) = y$ 번째 행의 흑화소 수

그림 2는 수평과 수직 히스토그램에 의해 계산되어진 비선형 메쉬 격자를 보여주고 있다.



(그림 2) 비선형 메쉬 격자

결과적으로 두 개의 히스토그램에 의해 $M \times N$ 개수의 비선형 메쉬 격자가 생성되고, 하나의 메쉬에는 방향 각도의 코드값을 누적시킬 수 있는 10개(코드화 방법II에서는 6개)의 버퍼를 마련한다. 그리고, 메쉬상의 각 점에 대해 방향 각도를 계산하고, 계산된 방향 각도를 코드화하여 그 값에 해당되는 버퍼에 누적시켜 각 메쉬마다 10개(코드화 방법 II에서는 6개)의 특징 벡터를 구하였다.

2.3 특징별 정규화

일반적으로 동일 클래스에 속하는 패턴들로부터 추출된 특징값의 분포를 일정하게 유지시키기 위해서 정규화를 수행한다. 전체 특징값들을 특징 벡터내의 최대 값으로 나누어서 0에서 1사이로 제한하는 방법이 많이 사용되는데, 여기서는 특징값이 추출된 화소의 종류에 따라 각기 다른 정규화 방법을 도입하였다.

1) 방향이 있는 특징

방향은 입력 패턴의 경계 지역에서 형성된다. 경계지역은 패턴의 크기에 비례하기 때문에 특징 벡터의 각 값을 $W+H$ 로 나누어서 정규화를 수행한다.

2) 방향이 없는 특징

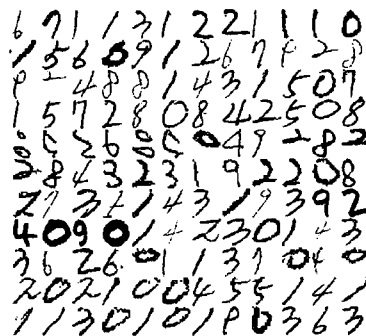
방향각도 값이 없는 특징은 패턴 경계의 내외부에서 정의된다. 패턴의 내부에서 정의된 특징값에 대해서는 특징 벡터의 각 값을 영상 내에 존재하는 흑화소(전경색)의 개수로 나누어 주고, 패턴의 외부에서 정의된 특징값에 대해서는 백화소(배경색)의 개수로 나누어 주었다. 즉, 각 경우에 대해 영상 내에서 차지하고 있는 해당 화소의 면적으로 정규화한다는 것을 의미한다.

3. 실험

3장에서는 실험 데이터, 인식기, 실험 결과치, 기존의 방법들과의 비교 등 본 연구 과정에서의 실험 과정에 대해 기술하였다.

3.1 실험 데이터

실험을 위하여 실제 필기 편지 봉투 15,000통에서 추출한 40,000개의 필기 숫자 데이터를 사용하였다. 여러 필기자에 의해 생성된 필기 편지 봉투의 각 필기 숫자들은 다양한 크기, 획두께, 형태 등을 가지고 있기 때문에, 소수의 필기자에 의해 수집된 한정적인 실험 데이터 보다는 실험 결과에 신뢰성을 더 부여할 수 있을 것이다. 특히, 실제 주소와 매칭해보지 않고서는 어떤 숫자인지 시각적으로는 구분할 수 없는 경우도 많다는 것이 편지 봉투에서 추출한 필기 숫자의 특징이라고 볼 수 있다. 그림 4에서 실제 편지 봉투로부터 추출된 필기 숫자 영상의 예를 나타내고 있는데, 매우 다양한 형태가 존재함을 볼 수 있을 것이다.



(그림 3) 필기 편지 봉투에서 추출된 숫자 영상의 예

3.2 실험 내용

본 논문에서 제안된 특징의 개발 과정 중에 비교 대상이 될 수 있는 몇 개의 유사 특징들이 있다. 실험에서는 제안된 특징에 의한 인식 성능의 비교 검증을 위하여 여섯 가지의 특징을 구현하여 인식 성능을 비교하였다. 편의상 구현된 특징을 표3에서와 같이 T1, T2, T3, T4, T5, T6로 구분하였다.

<표 3> 실험에서 사용된 여섯 가지 특징들

구분	종류	차원
T1	선형 메쉬 상의 흑화소 수 기반 특징	144
T2	선형 메쉬 상의 방향 각도 특징	288
T3	비선형 메쉬 상의 방향 각도 특징	288
T4	T3 + 방향이 없는 특징 추가	360
T5	T4 + 정규화 방법 개선	360
T6	코드화 방법II를 적용	216

<표 3>에서와 같이 구현된 여섯 가지의 특징을 사용해서 인식 성능을 측정하기 위해 하나의 은닉층만을 가진 단순 MLP (Multi Layer Perceptron) 신경회로망을 인식기로 채택하였다. MLP는 학습이라는 과정을 거쳐 입력 데이터를 구분해낼 수 있는 정보를 회로망 내의 가중치 벡터에 저장하는데, 그 구조적 단순함에 비해서 매우 신뢰성 있는 패턴 분별력을 가지고 있기 때문에 패턴인식 문제에서

가장 자주 사용되어 온 도구라고 볼 수 있다. MLP의 학습을 위해서 역전파 알고리즘을 사용하였으며, 은닉층의 노드 수는 60, 80, 100개씩 부여하여 각각 실험해보았다.

3.3 제안된 특징의 실험 결과

실험에 사용된 40,000개의 필기 숫자 중, 20,000개는 학습에, 나머지 20,000개는 인식 성능 검증에 사용하였다. T1부터 T6까지 은닉층의 노드 수를 변경해가면서 전체 18번의 학습과 인식 과정을 거친 결과를 <표 4>에 나타내었다. 그 결과, 여섯 가지 방식의 특징 중 제안된 특징에 의한 인식 성능이 가장 우수한 것으로 나타났다.

<표 4> 인식 성능 실험 결과

은닉층 노드수	특징 벡터 구분					
	인식률(%)					
	T1	T2	T3	T4	T5	T6
60	95.08	97.30	98.01	98.25	98.58	98.50
80	95.56	97.49	98.05	98.18	98.58	98.50
100	96.11	97.48	98.10	98.33	98.62	98.55

3.4 기존의 연구된 방법과 비교

제안된 특징 추출 방법에 의해 우수한 인식 성능을 구하였다고 하여도, 지금까지 발표된 다른 방법들과 비교해보지 않고서는 제안된 방법이 필기 숫자 인식에 있어서 효과적이라고 단언할 수는 없을 것이다. 그래서, 여러 논문에서 언급되고 실험 데이터로 사용된 캐나다 콩코디아 대학의 숫자 데이터 6,000자를 대상으로 동일 실험을 수행하여 보았으며, 그 결과를 <표 5>에서 나타내었다. 표5에서 볼 수 있는 바와 같이 제안된 방법이 가장 높은 인식률을 보임을 알 수 있다.

<표 5> 기존의 방법들과의 인식 성능 비교

방법	인식률 (%)
Krzyzak[1]	86.40
Krzyzak[1]	94.85
Lam[1]	93.10
Kim[3]	95.40
Kim[3]	95.85
Legault[1]	93.90
Mai[1]	92.95
Nadal[1]	86.05
Suen[1]	93.05
Cho[2]	97.35
Hwang[4]	97.90
제안된 방법(T5)	97.55
제안된 방법(T6)	98.05

Hwang[4]의 방법이 제안된 방법과 거의 비슷한 인식 결과를 가지고 있으나, 그 방법은 매우 복잡한

신경망 구성을 사용하여서 나온 결과이므로, 단순 MLP만을 사용한 본 논문의 인식기와 비교하기에는 어려운 측면이 있다.

4. 결론

본 논문에서 필기 숫자 인식을 위해 방향각 특징에 기반한 새로운 특징을 제안하였으며, 제안된 특징의 성능을 검증하기 위한 실험을 수행하였다. 첫 번째 실험에서는 제안된 특징을 개발하기 위한 중간 과정에 위치했던 특징들과 인식 성능을 비교해 보았는데 제안된 특징이 가장 우수한 성능을 보임을 결과로서 알 수 있었다. 그리고, 이전에 개발된 방법과 비교할 때, 매우 단순한 구조의 MLP만을 사용하였음에도 불구하고 높은 인식률을 획득 하였음을 두 번째 실험을 통하여 알 수 있었다. 그래서, 제안된 특징은 필기 숫자 인식에 효과적 이었다고 할 수 있을 것이다.

8방향만 사용한 T4와, 방향이 정의되지 않는 화소에서의 특징까지 포함한 T5를 비교해 볼 때, 방향이 없는 화소도 인식 성능을 향상시킬 수 있는 요소를 가지고 있는 것으로 분석할 수 있다. 또한, 8방향을 4방향으로 축소한 T6의 경우는 T5의 결과와 큰 차이가 없는 것으로 나타났는데, 이는 π 만큼의 위상차를 가지는 방향들을 별도의 것으로 구분하는 것은 큰 의미를 가지지 않는 것으로 분석되었다.

참고문헌

- [1] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam, "Computer recognition of unconstrained handwritten numerals"; *Proc of the IEEE*, vol. 80, no. 7, pp. 1162-1180, 1992.
- [2] S. B. Cho, "Neural-network classifiers for recognizing totally unconstrained handwritten numerals," *IEEE Trans. on Neural Networks*, vol. 8, no. 1, pp. 43-52, 1997.
- [3] Y. J. Kim and S. W. Lee, "Off-line recognition of unconstrained handwritten digits using multilayer backpropagation neural network combined with genetic algorithm," (in Korean), *Proc. of 6th Workshop on IPU*, pp. 186-193, 1994.
- [4] Y. Hwang and S. Bang, "Recognition of unconstrained handwritten numerals by a radial basis function neural network classifier," *Pattern Recognition Letters*, vol. 18, pp. 657-664, 1997.