
Zerinke 모멘트와 신경망을 이용한 온라인 필기체 숫자 인식

문원호* · 최연석* · 차의영*

*부산대학교

Recognition of Online Handwritten Digit using Zernike Moment and Neural Network

Won-ho Mun* · Yeon-suk Choi* · Eui-Young Cha*

*Dept. of Computer Engineering, Pusan National University

E-mail : aquamariner@nate.com

요 약

본 논문에서는 Zernike 모멘트와 backpropagation 신경망을 이용한 온라인 필기체 숫자 인식 방법을 소개한다. 마우스로 통해 입력된 숫자 정보는 전처리된 통해 시간에 순서적이고, 연속적인 좌표 정보로 변환된다. 전처리된 입력 좌표는 Zernike 모멘트(moment)와 각도 특징(angulation feature)을 이용하여 각 숫자가 가지는 고유의 특징을 만들어 낸다. 이러한 특징은 크기, 모양, 틀어진 정도에 상관없이 항상 일정한 성질을 가진다. 제안된 방법으로 추출된 특징은 패턴 구분을 위해 back propagation 신경망의 입력으로 사용된다. 본 논문은 200개의 필기체 숫자 데이터베이스를 이용하여 실험을 한 결과, 제시된 방법은 적은 학습데이터만으로 학습이 가능할 뿐만 아니라 좋은 인식률을 보여준다.

ABSTRACT

We introduce a novel feature extraction scheme for online handwritten digit based on utilizing Zernike moment and angulation feature. The time sequential signal from mouse movement on the writing pad is described as a sequence of consecutive points on the x-y plane. So, we can create data-set which are successive and time-sequential pixel position data by preprocessing. Data preprocessed is used for Zernike moment and angulation feature extraction. this feature is scale-, translation-, and rotation-invariant. The extracted specific feature is fed to a BP(backpropagation) neural network, which in turn classifies it as one of the nine digits. In this paper, proposed method not noly show high recognition rate but also need less learning data for 200 handwritten digit data.

키워드

Handwritten Digit, Zernike Moment, BackPropagation, Character Recognition

1. 서 론

문자 인식 기술은 입력 방법에 따라 온라인 인식(on-line recognition)과 오프라인 인식(off-line recognition)으로 크게 둘로 나뉜다. 온라인 인식은 사용자가 전자펜을 이용하여 필기하는 과정을 실시간으로 입력받아 인식하는 방법으로 입력패턴의 시간적 정보, 위치상의 공간적 정보, 압력 정보 등을 분석하여 인식한다. 반면에 오프라인

인식은 사용자가 노트위에 작성해 놓은 필기체 문서나 책 등의 인쇄된 문서를 스캐너 등의 영상 입력 장치로 복사하듯 입력 받아 인식하는 방법으로 입력 패턴의 위치상 공간적 정보를 분석하여 인식하는 기술이다.

온라인 인식에서 문자 패턴의 크기, 회전, 변형에 무관하게 인식하는 문제는 패턴인식에 있어서 매우 중요한 문제로써 최근 이러한 문제를 신경 회로망을 이용하여 해결하려는 많은 연구가 진행

되어 왔다 [1][2][3]. 이러한 연구로는 푸리에 변환(Fourier Transform)이나 모멘트(Moment) 등의 전처리 과정을 이용하여 불변 특징을 추출하여 신경회로망의 입력으로 사용함으로써 신경 회로망을 분류기로 사용하는 방법이 있고 [1][4], 네오코그니트론 모델처럼 신경회로망의 구조에 특징 추출 기능을 부가함으로써 불변 인식을 시도한 방법 등이 있다 [2]. 그리고 고차 신경 회로망은 입력 데이터의 활성화 관계를 이용하여 인식하려는 패턴의 이동, 크기, 회전에 불변인 특징을 추출할 수 있는 능력 때문에 패턴인식 분야에서 많은 각광을 받고 있다. 하지만 차수(order)가 증가함에 따라 연결선 수가 폭발적으로 증가한다는 문제점을 가지고 있으며, 패턴의 왜곡에 민감하기 때문에 필기자에 따라 변형이 심한 필기체 인식에서는 좋은 성과를 보이지 못하고 있는 실정이다.

본 논문에서는 비교적 많은 변형이나 왜곡을 흡수할 수 있는 필기체 인식에 적합한 Zernike 모멘트(Moment)와 기하학적 특징을 추출하여 특징 벡터를 생성한다. Zernike 모멘트는 회전 불변성에 대해서 이미 많은 연구가 진행되었으며, 회전된 문자나 숫자, 기호 등을 인식하는데 우수한 특성을 나타낸다 [5]. 위와 같이 추출된 특징 벡터는 backpropagation 신경회로망의 입력으로 이용하여 패턴을 분류하는 방법을 제안한다. 그리고 필기체 숫자 인식 실험을 통하여 제안한 시스템이 학습 패턴 수를 줄이고, 이동, 크기, 회전, 왜곡 등의 변형을 흡수할 수 있음을 보인다. 3장에서는 필기체 숫자의 전처리 과정을 언급하고, 4장에서는 Zernike 모멘트와 기하학적 특징을 추출하는 방법을 설명하고, 5장에서는 추출된 특징 벡터를 BP 신경 회로망을 통해 실험한 결과를 보이고, 6장에서는 결론을 맺는다.

II. 전처리 과정

온라인 문자인식의 경우 주로 마우스나 전자펜으로 패드에 입력된다. 입력되는 정보는 픽셀의 위치 정보뿐만 아니라 시간으로 정렬된 연속적인 dataset을 만들 수 있다. 그러므로 C_j 숫자의 μ 번째 픽셀 정보는 수식 (1)과 같이 표현된다.

$$C_i = [(x_{j0}^u, y_{j0}^u), (x_{j1}^u, y_{j1}^u), \dots, (x_{jn}^u, y_{jn}^u)] \quad (1)$$

그림 1은 숫자 '7'에 대한 시간에 따라 생성된 연속적인 좌표를 보여준다.

특징을 추출할 데이터로 사용할 입력 좌표의 개수는 시스템의 성능을 결정짓는 중요한 요소이다. 즉, 좌표의 개수가 너무 많으면 시스템의 속도가 느려지는 단점이 있는 반면 많은 정보를 활용할 수가 있다. 반대로 좌표의 개수가 너무 작으면 속도는 빨라질 수 있으나 활용할 수 있는 정보가 작아진다. 본 논문에서는 5-pixel을 기준으로

변화를 가진 정보를 입력 데이터를 시간으로 정렬한 좌표들을 입력데이터로 정의한다.

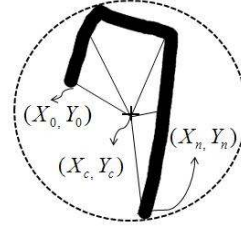


그림 1. 숫자 7의 전처리 예

III. 특징 추출

3.1. Zernike 모멘트

Zernike 모멘트는 모멘트에 근거한 방법으로서 물체가 표시되는 영역 전체를 이차원 함수로 간주하여 여러 차수 모멘트를 모델의 모멘트와 비교하여 물체만의 고유한 특징을 알아내는 방법이다. Zernike 모멘트는 직교 모멘트이므로 정보의 중복이 없고(information redundancy), 잡음 감도(noise sensitivity)가 다른 모멘트에 비해 가장 우수하다 [6].

Zernike가 소개한 크기가 1인 원의 내부에 완전한 직교함수의 형태를 가지는 복소 다항식의 형태는 수식 (2)와 같다.

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (2)$$

여기서 n 은 양의 정수 또는 0 , m 은 $n - |m| \geq n$, $n - |m|$ 이 짝수일 경우의 양 또는 음의 정수, ρ 는 원점에서 (x, y) 까지의 크기, θ 는 반시계 방향으로 x 축과 벡터 ρ 가 이루는 각도와 같다.

반지름 다항식(Radial polynomial)은 수식 (3)과 같이 정의 된다.

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s [(n-s)!] \rho^{n-2s}}{s! ((n+|m|)/2-s)! ((n-|m|)/2-s)!} \quad (3)$$

Zernike 모멘트는 직교 기본 함수들에게 영상함수로 투영한 것이다. 연속 영상 함수 $f(x, y)$ 에서 반복인자 m 을 가지는 차수 n 의 Zernike 모멘트를 디지털 영상에서 수식 (4)로 표현할 수 있다. 그림 2는 숫자 '2'에 대한 Zernike 모멘트를 3가지 샘플로 투영한 결과이다.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(x, y), \quad x^2 + y^2 = 1 \quad (4)$$

본 논문에서는 $f(x, y)$ 를 반지름이 r 인 원과 각 이웃하는 C_j 집합 내의 서로 이웃하는 좌표를 이은 선분이 교차하는 교차점의 집합을 말한다. 즉, 수식 (5)와 같이 표현 할 수 있다.

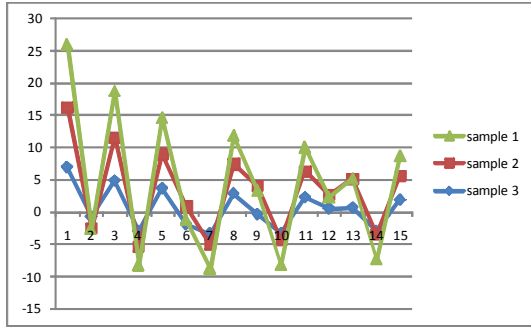


그림 2. 숫자 2에 대한 모멘트 크기

$$F(\phi) = [C_\phi \forall S(x, y)], \phi = 0, 1, 2, \dots, L \quad (5)$$

여기에서 ϕ 는 원의 숫자, $S(x, y)$ 는 C_j 의 이웃하는 좌표를 잇는 선분의 집합, C_ϕ 는 ϕ 번째 원, $F(\cdot)$ 는 각 원과 $S(x, y)$ 이 교차하는 교차점의 집합을 말한다.

3.2. Angulation feature

중심과 두 점을 연결한 두 선분 사이에 이루는 각도는 그림 3과 같다. 중심좌표를 (x_c^u, y_c^u) , 첫 좌표를 (x_0^u, y_0^u) 이고 두 번째 좌표는 (x_1^u, y_1^u) 이라 했을 때 세 좌표가 이루는 각도는 θ_0^u 로 표현한다. 이러한 각도 정보를 시간으로 정규화한 (x_k^u, y_k^u) 까지 데이터 집합은 수식 (6)과 같다.

$$\theta^u = [\theta_0^u, \theta_1^u, \theta_2^u, \dots, \theta_k^u] \quad (6)$$

위와 같이 방법은 크기와 회전에 상관없이 항상 같은 값을 가지기 때문에 각 숫자의 형태학적 특징을 결정짓는 요소가 된다.

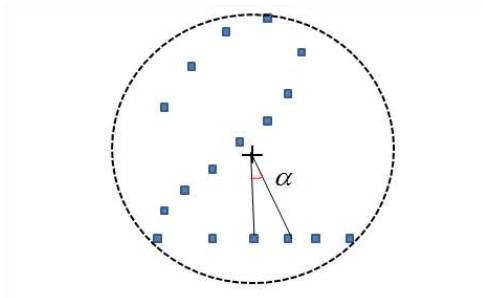


그림 3. 숫자 2의 k번째 각도 정보

IV. 패턴 분류

학습에 필요한 데이터는 숫자별 10개의 데이터를 학습하여 실험을 하였으며, 그림 4는 제안하는 시스템의 전체 구성도이다.

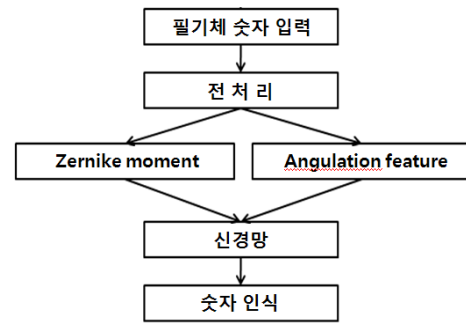


그림 4. 제안하는 시스템 전체 구성도

그림 5은 BP 신경망의 구조를 나타낸다. BP 신경망은 순방향 다층 신경망으로 패턴인식 분야에 널리 사용되어 왔다. 활성화 함수는 주로 시그모이드 함수를 사용한다. 본 논문에서는 Zernike 모멘트와 각도 특징으로 추출된 데이터를 BP 신경망의 입력으로 사용한다. 입력층(Input Layer)의 개수는 Zernike 모멘트 20개와 각도 특징 20개를 포함한 총 40개의 정보를 사용하고, 은닉층(Hidden Layer)의 개수는 10개, 출력층(Output Layer)의 개수는 10개(digit 0~9)를 사용한다.

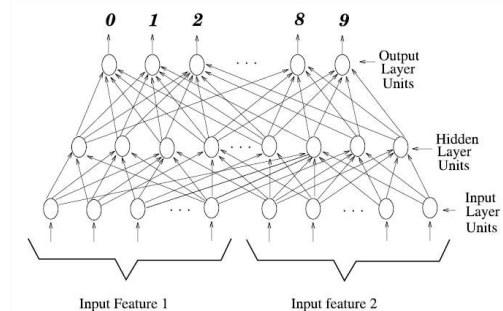


그림 5. BP 신경망의 구조

학습률(α)은 0.1, 오차 최대치(δ_{max})는 0.01로 사용했으며, 특히, 은닉층의 개수는 최대 100개까지 설정이 가능하며, 은닉층의 개수에 따라 다소 성능에 차이는 보인다. 본 논문에서는 제시하는 최적의 은닉층의 수는 10개이다.

표 1은 200개의 입력 패턴에 대한 인식률을 보여준다. 정인식과 오인식의 평균은 각각 86.5%와 13.5%이다. 숫자 '2'와 '3', 숫자 '6'과 '9'는 모멘트 투영 값이 유사하여 인식률이 상대적으로 낮게 나왔다. 반면에 뚜렷한 모멘트 특징을 갖는 숫자 '1'과 '8'과 같은 경우는 높은 인식률을 보였다.

표 1. 인식 결과(단위:%)

숫자	정인식	오인식
0	90	10
1	100	0
2	85	15
3	80	20
4	85	15
5	90	10
6	80	20
7	85	15
8	95	5
9	75	25
평균	86.5	13.5

V. 결 론

본 논문에서는 온라인 필기체 숫자를 인식하기 위해 크기, 회전, 변화에 강건한 Zernike 모멘트와 각도 특징 추출 기법을 사용하였다. n개의 원과 전처리를 통해 입력되는 좌표들을 잇는 선분이 교차하는 교차점의 집합을 Zernike 모멘트의 $f(x, y)$ 의 입력 값으로 사용하였다. Zernike 모멘트는 필기체 숫자 인식에 있어서도 우수한 회전 불변 성질을 갖는 다는 결론을 얻을 수 있었다. 실험 결과를 통해 86.5%의 인식률을 보였다. 다소 필기자에 따라 변형이 심한 숫자는 패턴 분류가 잘 되지 않았다. 제안된 방법은 오프라인에 방식과 비교해 볼 때 적은 신경망 입력층의 개수를 가지고 있기 때문에 실행속도가 빠르며, 적은 학습 데이터를 사용하여도 좋은 성능을 보여준다.

참고문헌

- [1] G.L.Cash, M.Hatamian, "Optical character recognition by the method of moments," Computer Vision, Graphics, and Image Processing, Vol.39, pp.291-310, 1987
- [2] K.Fukushima, "Neural Network Model for Selective attention in visual pattern recognition and associative recall," Applied Optics, Vol.26, No.23, pp.4985-4992, 1987
- [3] C.L.Giles and T.Maxwell, "Learning,

invariance, and generalization in high-order neural networks," Applied Optics, Vol.25, pp.4972-4978, 1987

[4] M.T.Y.Lai and C.Y.Suen, "Automatic recognition of characters by Fourier descriptors and boundary encoding," Pattern Recognition, Vol.14, pp.383-393, 1981

[5] P.M.Patil and T.R.Sontakke, "Rotation, scale and translation invariant handwritten Devanagari numeral character recognition using general fuzzy neural network," Pattern Recognition, Vol.40, pp.2111, 2006

[6] S.Belkasim, E.Hassan, T.Obeidi, "Explicit invariance of Cartesian Zernike moments," Pattern Recognition, Vol.28, pp.1969-1980, 2007