

고유벡터를 이용한 필기체 숫자인식

박중조* · 김경민** · 송명현***

Recognition of Handwritten Numerals using Eigenvectors

Joong-Jo Park* · Kyoung-Min Kim** · Myung-Hyun Song***

요약

본 논문에서는 고유벡터를 이용한 오프라인 필기체 숫자인식 기법을 제시한다. 본 기법에서는 KL 변환에 의한 고유벡터를 이용하여 통계적으로 숫자의 특징을 추출하며, 특징공간상에서 최소거리 기법으로 숫자를 인식한다. 본 기법에서 제안된 특징추출 방법에서는 많은 표본 숫자영상에서 각 숫자들의 특징을 가장 잘 표현하는 기저벡터를 찾아내고 이로부터 숫자의 특징을 구한다. 제시된 기법의 성능 평가를 위해 Concordia 대학의 무제한 필기체 숫자 데이터베이스를 사용하여 실험한 결과 96.2%의 인식률을 얻을 수 있었다.

ABSTRACT

This paper presents off-line handwritten numeral recognition method by using Eigen-Vectors. In this method, numeral features are extracted statistically by using Eigen-Vectors through KL transform and input numeral is recognized in the feature space by the nearest-neighbor classifier. In our feature extraction method, basis vectors which express best the property of each numeral type within the extensive database of sample numeral images are calculated, and the numeral features are obtained by using this basis vectors. Through the experiments with the unconstrained handwritten numeral database of Concordia University, we have achieved a recognition rate of 96.2%

키워드

Handwritten Numeral Recognition, KL Transform, Eigen-Vector, Nearest-Neighbor Classifier

1. 서론

최근 컴퓨터의 급속한 발전에 따라 정보 처리 시스템의 자동화 연구가 활발히 진행되고 있다. 특히 지난 40여년 동안 꾸준히 연구되어 온 문자인식은 사무 자동화 및 문자에 의한 분류 자동화 등에 근간이 되는 기술로서 그 응용범위가 매우 광범위하여 국내외적으로 많은 연구가 이루어지고 있다. 문자인식의 한 분야인 필기체 숫자인식은 온라인 및 오프라인 인식으로 나뉘어 연구되고 있으며, 이중 오프라인 숫자인식은 신용카드의 전표, 우편번호 자동 인식

및 생산공정에서의 분류 자동화 등 많은 응용분야에서 유용하게 사용될 수 있는데, 입력 숫자영상에 포함된 잡영과 왜곡, 필기자에 따라 다양하게 나타나는 필체의 변형, 필기 도구 및 필기 용지의 질감 특성 등 해결해야 할 많은 난제들이 있기 때문에 이러한 모든 문제를 극복할 수 있는 인식 시스템의 개발이 어려운 실정이다. 따라서 인식률의 향상을 위해 입력 데이터 처리 및 인식기 설계등 각 단계에서의 개선된 요소 기술의 개발에 대한 많은 연구가 이루어지고 있다[1-6].

일반적으로 인식 시스템의 성능은 인식기의 종류

* 경상대학교 전기전자공학부, 컴퓨터정보통신연구소 연구원
*** 순천대학교 정보통신공학부

** 여수대학교 전자통신 및 전기공학부
접수일자 : 2002. 10. 17

와 인식기가 사용하는 특징에 따라 좌우된다. 숫자의 특징이란 서로 이종인 숫자 형상의 차이점을 잘 표현하는 특성을 수치화한 것으로서, 국부적 특징과 전역적 특징이 있다. 이때 좋은 특징이란 높은 인식률을 성취하는 것으로서 연구방향에 따라 많은 종류가 보고되고 있다. 필기체 숫자 인식을 위한 대표적인 인식기의 종류로는 숫자를 구성하는 화소의 통계적 특징에 근거한 인식기(feature vector based classifier), 문자의 형태 특성에 근거한 구조적인 인식기(syntactic and structural based classifier) 및 신경회로망을 이용한 인식기(neural network based classifier) 등이 있다[5-10]. 기존의 필기체 숫자인식들은 대부분의 경우 효과적인 인식기 설계에 대한 연구가 활발한 반면 특징추출 기법의 연구는 미흡한 편이다. 그러나 이용할 수 있는 특징의 종류가 제한된 상태에서 인식기만을 개선하여 인식률을 높이는 데에는 한계가 있으며, 인식기의 성능이 입력 특징에 크게 의존함에 비추어 볼 때 필기체 숫자인식을 위한 새로운 특징추출기법의 개발이 필요한 실정이다.

본 연구에서는 고유벡터(eigenvector)를 이용한 필기체 숫자 인식기법을 제안한다. 본 기법은 숫자 형태 자체를 기반(holistic view approach)으로 하는 인식방법으로서, KL변환에 의한 고유벡터를 이용하여 통계적으로 숫자의 특징을 추출하며, 인식기로는 최소거리(Nearest Neighbor) 기법을 사용한다. 제안된 기법의 성능 테스트를 위해 캐나다 Concordia 대학의 필기체숫자 데이터베이스를 사용하여 실험하였으며, 실험 결과 96.2%의 인식률을 얻을 수 있었다.

II. 필기체 숫자영상의 전처리

필기체 숫자로부터 인식에 유용한 특징을 효과적으로 추출하기 위해서는 숫자 특징의 추출에 앞서 먼저 숫자영상의 전처리가 필요하다. 고유벡터를 이용하여 숫자의 특징을 추출하려면 숫자영상의 크기가 일정해야 한다. 따라서 전처리의 첫 단계에서는 보간법을 이용한 영상의 축소 및 확대 방법을 사용하여 숫자영상을 일정한 크기로 크기-정규화한다. 그 후 인계화를 통해 영상을 이치화한다. 보간법에 의해 크기-정규화된 숫자영상은 계조치영상이므로 다음

단계인 세선화를 위해서 영상의 이치화가 필요하다. 그후 영상을 세선화하고 팽창(dilation)연산을 수행하여 숫자를 이루는 선이 일정한 두께가 되도록 만든다. 이러한 방식으로 숫자선을 일정한 두께로 변환하고 이로부터 숫자의 특징을 추출하면 숫자 인식률의 향상을 도모할 수 있다. 그후 전처리의 마지막 단계로서 숫자영상에 평균화를 수행한다. 이 평균화 처리도 역시 숫자의 인식률 향상을 위해 필요하다.

III. 고유벡터를 이용한 숫자의 특징추출 및 인식

KL변환은 주로 영상압축 분야에서 사용되어 왔으나 1987년 Kirby와 Sirovich에 의한 얼굴인식 연구에 의해 패턴인식에서의 활용방법이 알려졌다. 이 방법은 많은 표본영상에서 각 영상들의 특징을 가장 잘 표현하는 벡터를 찾아내는 것으로서, 이 벡터는 주어진 많은 표본 영상들의 기저(basis)가 되는 벡터이다[11-13].

KL변환은 입력 데이터를 상관성이 없는 계수 집합으로 변환시키는 변환 방법으로서 입력 데이터의 2차 통계량(분산)에 의한 직교 기저함수로 데이터를 표현하기 때문에 가우시안 분포를 갖는 데이터에 대해서는 데이터에 함유된 고유의 특징을 잃지 않으면서 선형 변환하거나 낮은 차원으로 변환시키는데 있어서 최적의 변환방식을 제공한다[11].

N2 차원인 M개의 벡터 x_i 로 구성된 벡터의 모집단 x 에 대한 기대값(expected value) m_x 와 공분산 행렬(covariance matrix) C_x 는 아래와 같이 계산된다.

$$m_x = E\{x\} = \frac{1}{M} \sum_{i=1}^M x_i \quad (1)$$

$$C_x = E\{(x - m_x)(x - m_x)^T\} = \frac{1}{M} \sum_{i=1}^M (x_i - m_x)(x_i - m_x)^T \quad (2)$$

공분산행렬 C_x 는 실수값을 가지며 대칭이기 때문에, N2개의 정규직교인 고유벡터들(eigenvectors)을 구하는 것은 항상 가능하다. 이제 C_x 의 고유값들(eigenvalues)을 크기순으로 정렬($\lambda_1, \lambda_2, \dots, \lambda_{N^2}$)하고 그에 대응하는 각 고유벡터들 e_1, e_2, \dots, e_{N^2} 를 구한 후, 이를 이용하여 다음의 행렬 A를 만든다.

$$\mathbf{A} = [\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_{N^2}]^T \quad (3)$$

여기서, 행렬 A의 첫번째 행은 가장 큰 고유값에 대응되는 고유벡터이고, 마지막 행은 가장 작은 고유값에 대응되는 고유벡터이다. 그러면 이 행렬 A는 입력벡터 x를 계수벡터 y로 사상시키는 변환행렬이 되며, 이 행렬을 사용하여 다음과 같은 KL변환이 정의된다.

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad (4)$$

KL변환의 중요한 특성은 고유벡터와 계수벡터 y로부터 원래의 x를 완벽히 복원할 수 있다는 것이다. 즉, Cx는 실수값을 갖는 대칭행렬이므로 직교 정규인 고유벡터가 존재하게 되어 A⁻¹=A^T가 되고, 따라서 식 (4)의 역변환에 의해 다음과 같이 y로부터 x가 복원될 수 있다.

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} = \mathbf{A}^T\mathbf{y} \quad (5)$$

한편, Cx의 모든 고유벡터를 사용하는 대신에 K개의 가장 큰 고유치에 대응하는 K개의 고유벡터들만을 사용하여 변환행렬 AK를 구성한다면 식 (4)에서 벡터 y는 K차원이 되며 이 방식으로 구해지는 벡터 x는 더 이상 정확한 본래의 값이 되지 못하고 근사치가 된다. 이 근사치 x'는 다음 수식으로 구해진다.

$$\mathbf{x}' = \mathbf{A}_K^T \mathbf{y} \quad (6)$$

이때, x와 x'사이의 평균 제곱 오차는 다음 수식으로 구해진다.

$$e_{ms} = \sum_{j=1}^{N^2} \lambda_j - \sum_{j=1}^K \lambda_j \quad (7)$$

여기서, λ_j는 단조 감소로 정렬된 것이기 때문에 식 (7)은 가장 큰 고유값들과 관련 있는 K개의 고유벡터를 선택함으로써 오차를 최소화할 수 있다는 것을 보여준다. 즉, KL변환은 최소 제곱 오차의 의미에서 최적 변환이다. KL변환의 기본적인 개념은 기저 벡터를 선택하는데 있어서 벡터가 자료의 분산이 최대인 방향을 가지도록 하고, 또한 벡터들 간에 상호 직교성을 가지며 변환된 성분들이 비상관적이 되도록 하는데 있다.

본 연구에서 제시하는 숫자특징 추출 방법은 KL변환을 사용하는 것으로서, 샘플 숫자영상 집단을 기저 벡터의 집합으로 표현하여 이를 숫자인식에 사용하는 방법이다. 이 방법은 상관관계 정보를 영상에서

추출하고 KL변환 계수값의 유사도를 비교하여 숫자인식을 수행한다. 그런데 KL변환은 본질적으로 1차원 데이터에 사용할 수 있으므로 2차원 데이터인 숫자영상에 이를 적용하려면 이를 1차원 벡터형태로 변형시킨 후 사용해야 한다.

먼저, 2차원 데이터인 숫자영상에 KL변환을 적용하기 위해, 크기가 N×N인 M개의 숫자영상을 크기가 N²×1인 1차원 벡터 I_k, k=1,2,...,M로 변형시키고 그 후 다음 수식에 의해 이들의 평균영상 Ψ를 구한다.

$$\Psi = \frac{1}{M} \sum_{k=1}^M I_k \quad (8)$$

그리고, 각 숫자영상에서 평균영상을 뺀 차이영상 Φ_k를 구한다.

$$\Phi_k = I_k - \Psi, k=1, 2, \dots, M \quad (9)$$

이제, M개의 차이영상 Φ_k를 사용하여 주어진 숫자영상들의 분포를 가장 잘 나타내는 M개의 정규직교벡터 u_k를 구한다. 이 u_k는 식 (10)의 공분산행렬 C의 고유벡터를 계산함으로써 구할 수 있다. 이때 행렬 C의 가장 큰 고유값부터 시작하여 크기순으로 각 고유값에 해당되는 고유벡터를 구해야 한다.

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = \mathbf{B}\mathbf{B}^T \quad (10)$$

여기서, 행렬 B=[Φ₁ Φ₂... Φ_M] 이런 방식으로 얻어진 고유벡터를 원래의 영상 표현인 N×N으로 재배열하여 보면 숫자영상과 닮아있게 된다. 이 고유벡터는 샘플 숫자영상들의 주성분으로 해석될 수 있다.

일단, 고유벡터 u_k가 구해지면, 식 (11)에 의해 임의의 숫자영상 I에 대한 가중치 벡터 Ω=[w₁,w₂,...,w_M]가 구해지고, 이 가중치 벡터 Ω가 숫자영상 I에 대한 특징벡터로 사용된다.

$$\mathbf{w}_k = \mathbf{u}_k^T (\mathbf{I} - \Psi), k=1, \dots, M \quad (11)$$

한편 숫자영상의 개수 M이 1차원 벡터화된 숫자영상의 차수 N²보다 작은 경우에는 의미있는 고유벡터의 개수는 M-1개가 되고 나머지 고유벡터들은 0의 고유값에 해당된다. 이 경우에는 식 (10)의 행렬 C로부터 직접 고유벡터를 구하는 것보다 행렬 BTB의 고유벡터 v_k를 먼저 구하고 u_k=Bv_k를 사용함으로써 보다 적은 계산량으로 고유벡터 u_k를 구할 수 있다[12].

일단 고유벡터를 이용하여 숫자 특징이 추출되면 이 특징공간상에서 최소거리(Nearest Neighbor) 기법을 사용하여 숫자를 인식한다. 최소거리 기법을 사용하기 위해 먼저 필기체 숫자 데이터베이스에서 일정 개수의 표본데이터들을 학습용 숫자데이터로 선택하고 이로부터 각각의 숫자에 대한 특징값 집단을 구해 놓는다. 이렇게 준비된 각 숫자에 대한 특징값 집단은 미지의 숫자가 어떤 종류의 숫자인가를 판정하는데 사용된다. 즉 하나의 미지 숫자가 주어지면 그의 특징값을 추출하고 특징공간상에서 가장 가까운 학습용 숫자를 이 미지숫자의 실제 숫자 종류로 판정한다.

IV. 실험 및 결과

제안된 기법의 성능 테스트를 위해 PentiumII 400 PC에서 C언어를 사용하여 실험하였다. 본 실험에서 사용된 숫자 데이터는 캐나다 Concordia 대학의 필기체 숫자 데이터베이스로서 이는 총 6000개의 숫자들로 구성되는데, 본 실험에서는 이 중 4000개를 학습용 숫자 샘플로 사용하고 나머지 2000개를 테스트용 숫자로 사용하였다. 그림 1은 이 데이터베이스내의 숫자들 일부를 보인다.

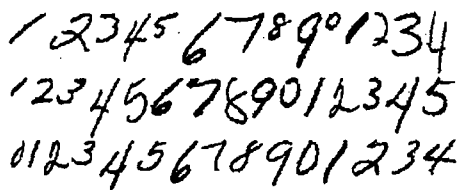


그림 1. 실험에 사용한 필기체 숫자들의 예
Fig. 1. Some examples of numeral data

필기체 숫자영상의 특징 추출에 앞서 먼저 다음과 같이 숫자영상에 전처리를 수행하였다. 즉, 먼저 입력된 숫자영상을 [25×25] 크기로 크기-정규화한 후 임계값 150을 사용하여 영상을 이치화하였다. 그리고 이 영상에 8회의 세선화를 수행하고 그후 1회의 확산연산을 수행하여 숫자선의 굵기를 일정하게 만들었다. 그후 평균화 연산을 2회 수행하여 숫자의 특징 추출을 위한 전처리를 완료하였다.

전처리된 숫자영상에 대해 고유벡터를 이용하여 숫자의 특징을 추출하였다. 전처리된 숫자영상의 크기는 [25×25]이므로 총 625개의 고유벡터가 얻어진다. 본 실험에서는 숫자특징의 추출에 있어서 총 625개의 고유벡터들 중 80개의 고유벡터를 사용함으로써 하나의 숫자영상에 대해 80개의 특징을 구하였다. 그리고 구해진 숫자 특징을 사용하여 특징공간상에서 최소거리(Nearest Neighbor) 기법을 사용하여 필기체 숫자인식을 수행하였으며, 제시된 기법에 의해 96.2%의 인식률을 얻을 수 있었다.

표 1은 제안된 특징추출 기법에 의한 숫자인식의 결과를 보인다. 이 표에서 보면 가장 많은 오인식을 나타낸 경우는 숫자 7과 1이 서로 혼동된 경우와 숫자 5와 3이 서로 혼동된 경우로서 각각 6회의 오인식을 나타내고 있다. 필기체 숫자를 심하게 휘갈겨 쓸 경우를 생각해 볼 때 다른 숫자에 비해 이들 숫자에서 보다 많은 오인식이 나타날 수 있을 것으로 생각된다.

표 1. 제안된 특징에 의한 인식결과
Table 1. Recognition result using proposed feature

class	0	1	2	3	4	5	6	7	8	9	error [%]	recog [%]
0	199	0	0	0	0	1	0	0	0	0	0.5	99.5
1	0	200	0	0	0	0	0	0	0	0	0.0	100.0
2	2	0	187	3	1	1	2	2	1	1	6.5	93.5
3	2	0	3	192	0	1	0	1	1	0	4.0	96.0
4	1	2	0	0	190	0	1	0	0	6	5.0	95.0
5	2	1	0	6	0	187	2	1	1	0	6.5	93.5
6	3	0	0	0	0	0	197	0	0	0	1.5	98.5
7	0	6	0	1	3	0	0	186	0	4	7.0	93.0
8	1	0	0	3	1	1	2	0	188	4	6.0	94.0
9	0	0	0	0	0	0	0	1	1	198	1.0	99.0
avg											3.80	96.20

표 2는 사용되는 고유벡터의 개수가 인식률에 미치는 영향을 보인다. 이 표를 고찰해 보면 좋은 인식률을 얻기 위해서는 일정 개수 이상의 고유벡터가 필요하며, 사용되는 고유벡터가 특정 개수를 넘어서면 더 많은 고유벡터를 사용하더라도 그 이상의 인식률 향상은 이루어지지 않음을 알 수 있다.

표 2. 사용되는 고유벡터의 개수와 인식률
Table 2. No of used eigenvectors vs recognition rate

No of engenvectors	Recognition rates
40	95.95
80	96.20
400	96.20
625	96.20

V. 결 론

본 연구에서는 고유벡터(eigenvector)를 이용한 필기체 숫자 인식기법을 제안하였다. 본 기법은 KL 변환에 의한 고유벡터를 이용하여 숫자의 형태 자체로부터 통계적으로 숫자의 특징을 추출하고 최소거리 기법을 사용하여 숫자 인식을 수행한다. 본 기법에서 사용된 특징추출 방법은 기존의 숫자 형태(shape)로부터 각 숫자의 차별적인 특성을 구하는 대신 많은 표본 숫자영상에서 각 숫자들의 특징을 가장 잘 표현하는 기저벡터를 찾아내고 이로부터 통계적으로 숫자의 특징을 구한다.

제안된 기법의 성능 테스트를 위해 캐나다 Concordia 대학의 필기체 숫자 데이터베이스를 사용하여 실험하였으며, 실험 결과 96.2%의 인식률을 얻을 수 있었다.

필기체 숫자인식의 성능은 여러 종류의 특징들을 복합적으로 사용하거나 여러 종류의 인식기를 결합하여 사용함으로써 더욱 개선될 수 있다. 향후 연구 과제로는 필기체 숫자의 인식률 제고를 위해, 본 연구에서 제시된 숫자특징 추출기법과 기존의 다양한 특징추출 기법으로부터 상호 독립적인 특징들을 추출하고 이를 복합인식기(hybrid recognizer)에 적용하는 방안을 연구하고자 한다.

참고문헌

[1] Øivind Due Trier, Anil K. Jain and Torfinn Taxt, "Feature Extraction Methods for Character Recognition-A Survey", Pattern Recognition, Vol. 29, No. 4, pp. 641-662, 1996.
[2] K.M. Mohiuddin and J.Mao, "A Comparative

Study of Different Classifiers for Handprinted Character Recognition", Pattern Recognition in practice IV, pp. 437-448, 1994.
[3] P.Gader, B. Forester et al, "Recognition of Hand-written Digits using Template and Model matching", Pattern Recognition, Vol. 24, No. 5, pp. 421-431, 1991.
[4] S.Knerr and L.Personnaz and G. Dreyfus, "Hand written Digit Recognition by Neural networks with Single-layer Training", IEEE Trans. Neural Networks, Vol. 3, No. 6, pp. 962-968, 1992.
[5] S.B.Cho and J.H.Kim, "Multiple Network Fusion using Fuzzy Logic", IEEE trans. Neural Networks, Vol. 6, No. 2, pp.487-501, March, 1995.
[6] Y.J.Kim, "A new type of recurrent neural network for handwritten character recognition", M.S. thesis of Chungbuk Univ., 1995.
[7] S.B.Cho, "Neural network Classifiers for Recognizing Totally Unconstrained Handwritten Numerals", IEEE Trans. on Neural Networks, Vol. 8, No. 1, pp. 43-53, 1997.
[8] L.Lam and C.Y.Suen, "Structural Classification and Relaxation matching of Totally Unconstrained Handwritten Zip code Numbers", Pattern Recognition, Vol. 21, No. 1, pp. 19-31, 1988.
[9] 임준호, 채수익, "단층 신경망과 이중 기각 방법을 이용한 문자인식", 전자공학회논문지(B) 제32권, 제3호, pp.119-129. 1995.
[10] 류한진, 이필규, "특징벡터를 이용한 무제약 필기체 숫자의 인식", 정보과학회학술발표논문집 Vol.22, No.1 pp.209-212. 1995.
[11] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces", IEEE Trans. on PAMI, Vol. 12, No. 1, pp.103-108, 1990.
[12] M. Turk and A. Pentland, "Eigenface for Recognition", J. Cognitive Neuroscience, vol. 3, pp. 71-86, 1991.
[13] A. Pentland, T. Starner, N. Etocoff, A. Masoiu, O. Oliyide, and M. Turk, "Experiments with Eigenfaces", IJCAI '93, Chamberry, France, 1993.

저자소개



박중조(Joong-Jo Park)
1981년 고려대 전기공학과 졸업,
동대학원 석사(1983), 박사(1995)
1996년~현재 경상대학교 전기전
자공학부 부교수
※관심분야: 컴퓨터비전, 생체인식.



김경민(Kyoung-Min Kim)
1988년 고려대 전기공학과 졸업,
동대학원 석사(1991), 박사(1996)
1997년~현재 여수대학교 전자통
신 및 전기공학부 조교수
※관심분야: 컴퓨터비전, 퍼지 및
신경회로망 응용.



송명현(Myung-Hyun Song)
고려대학교 전기공학과 졸업
학사, 석사, 박사.
1988년~현재 순천대학교 정보통
신공학부 교수
※관심분야: 제어시스템 설계, 퍼
지 및 신경회로망