

웨이블릿과 주성분분석을 이용한 홍채 특징 추출

김귀주⁰, 이일병⁺

⁰연세대학교 컴퓨터과학산업시스템공학과

⁺연세대학교 컴퓨터정보과학부 및 BERC

(kamang98, yblee)@csai.yonsei.ac.kr

Iris Feature Selection Using Genetic Algorithm

Kwi-Joo Kim⁰, Yillbyung Lee⁺

⁰Dept. of Computer Science and Industrial Systems Engineering, Yonsei University

⁺Division of Computer and Information Engineering, and of BERC, Yonsei University

요 약

패턴인식에서 사용되는 모든 영상을 해석할 수 있는 시스템은 현재 어려운 문제이므로 먼저 문제 영역을 해석하는 방법이 일반적이다. 이는 영상의 이해 및 특징 정보값을 사용하여 특징으로 사용하기 위하여 특징을 추출한다. 특징 추출시 정보의 손실 없이 데이터를 줄이는 작업이 연구되어오고 있지만 이는 사람의 주관적인 경험으로 알고리즘의 내부 파라미터를 결정하고 바꾸어야 하는 문제점을 가지고 있다.

본 논문에서는 웨이블릿 변환으로 얻어진 sub-band들을 특징으로 추출하고 결과를 통계적인 방법인 주성분분석을 이용하여 특징 차원을 감소시켜 인식의 정확도를 향상시키는 방법을 제안한다.

1. 서 론

현대 사회가 좀더 극명하고 명확한 보안을 요구함으로써 기존의 보안 방법으로는 현대 사회에서 보안 요구 사항을 충분히 만족시키지 못하고 있다. 이러한 요구로 생체 인식 방법이 많은 발전을 하고 있다.

생체인식은 인간의 생리적 또는 행동적인 특징을 기반으로 개인을 식별하는 방법으로[1] 최근 각광을 받고 있는 것으로 개인의 고유한 생물학적인 특성을 이용한 인증 방법인 얼굴, 지문, 망막, 홍채, 서명 등이 널리 활용되어 지고 있다. 이러한 생물학적인 특성을 이용한 개인 인증 방법 중에서 홍채는 오랜 시간이 지나도 개인이 가지는 특징의 변화가 없으며, 변별도가 매우 높은 것으로 알려져 있다.[2]

홍채 인식 과정에서 개개인을 인증 또는 인식하기 위해서는 홍채 영상이 가지는 유용한 정보를 손실 없이 영상을 대표하는 특징들을 추출하고 추출된 특징의 분석하여 최적의 데이터의 형태로 만드는 것이 필요하다.

본 논문에서는 홍채 인식을 위한 특징 추출 방법으로 웨이블릿과 주성분분석을 사용하였다. 이는 다량의 입력 데이터로부터 정보를 최대한 보존하면서 특징 벡터의 차원을 줄일 때 사람의 주관적인 판단 없이 원 영상이

=====

“본 연구는 KOSEF 생체인식연구센터(BERC)로부터 부분적인 지원을 받아 수행되었습니다.”

가지고 있는 데이터의 성분을 최대한 이용하여 특징을 선택할 수 있다. 이러한 이유로 홍채 영상에서 방대한(redundant) 정보와 잡영(noise)을 줄임(reduction)으로써 압축적이고(compact) 강인한(robust) 특징을 효율적으로 추출해낼 수 있다.

2. 특징추출

2.1 Wavelet을 이용한 영상 분석

본 논문에서 홍채 영상으로부터 특징을 추출한 방법으로 최근 패턴인식, 영상처리, 신호처리 등의 분야에서 널리 활용되고 있는 웨이블릿 변환을 이용하였다. 웨이블릿 변환을 이용하여 홍채 영상 분석시 다해상도(multiresolution) 분석 기법을 적용하였다.[3] 이는 공간적인 정보와 주파수적인 정보를 동시에 표현할 수 있다는 특징을 지니고 있기 때문에 영상처리 분야에 널리 활용되고 있다.

영상에 웨이블릿 변환을 적용하는 것은 2차원 신호에 웨이블릿 변환을 도입하는 것이라고 할 수 있으며, 1차원 웨이블릿 변환으로부터 2차원 웨이블릿 변환으로의 확장은 각 스케일링 함수와 웨이블릿의 곱으로 구해지는 1차원의 분해 가능한 웨이블릿을 이용하여 다차원 필터뱅크(multidimensional filter bank)를 구성하는 것으로 볼 수 있다.

본 논문에서는 카메라로부터 획득된 홍채 영상을 가장

효과적으로 표현할 수 있는 특징 벡터를 구성하기 위하여 다해상도 Daubechies 웨이블릿 기법을 사용하여, 원 영상을 3단계로 분할하였다.

웨이블릿으로 분할된 홍채 영상은 다음과 같이 표현될 수 있다.



그림 1. Daubechies wavelet을 이용한 홍채 영상 3단계 분할

2.2 PCA를 이용한 특징추출 알고리즘

PCA의 접근 방법은 대표적으로 두 가지가 있다. 첫째 분산최대를 이용한 PCA 방법, 둘째 최소 MSE를 이용한 PCA 방법이다. 각각의 방법은 서로 다른 접근 방법을 이용하나 같은 결과를 얻어낸다[4].

2.2.1 분산 최대를 이용한 PCA

주어진 데이터의 분포에서 임의의 축으로 투영시 가장 분산이 크게되는 축을 찾아가는 방법이다.

아래와 같은 선형 변환을 고려해보자.

$$y_1 = \sum_{k=1}^m w_{k1} x_k = w_1^T x \quad \text{식 (1)}$$

식(1)에서 x 는 입력 데이터이고 y 는 새로운 축 w 로 투영된 데이터이다.

새로운 데이터 y 를 이용하여 분산최대를 이용한 PCA의 측정함수는 아래와 같다.

$$\begin{aligned} J_1^{PCA}(w_1) &= E\{y_1^2\} \\ &= E\{(w_1^T x)^2\} \\ &= w_1^T E\{x x^T\} w_1 \\ &= w_1^T C_x w_1 \end{aligned} \quad \text{식 (2)}$$

기본 선형대수 이론에서 w_1 이 C_x 의 첫 번째 고유벡터 일 때 $w_1^T C_x w_1$ 가 최대가 된다. 즉 식(1)의 선형 조합은 다음 식(3)과 같이 입력 데이터를 입력 데이터의 covariance 행렬의 첫 번째 고유벡터로 투영시킨 결과가 된다. y_1 은 첫 번째 principle compoenet가 된다.

$$y_1 = e_1^T x \quad \text{식 (3)}$$

이를 일반화(Generalization)하면,

$$\begin{aligned} y_m &= e_m^T x \\ e_m: m^{th} \text{ 고유벡터} \end{aligned} \quad \text{식 (4)}$$

와 같이 새로운 축을 얻을 수 있다.

이렇게 얻은 데이터 y 의 covariance를 보면 아래 식과

같이 0이 됨을 알 수 있어 새로운 데이터는 비상관관계가 된다.

$$\begin{aligned} E\{y_m y_k^T\} &= E\{(w_m^T x)(w_k^T x)^T\} \\ &= w_m^T C_x w_k \\ &= 0 \quad (k \neq m) \end{aligned} \quad \text{식 (5)}$$

2.2.2 최소 MSE(Mean Square Error)를 이용한 PCA

주어진 데이터의 분포에서 임의의 축으로 투영시 투영된 데이터와 원래 데이터간의 MSE가 최소가 되게 새로운 축을 찾는 접근 방법이다.

이런 접근 방법은 식(6)으로 나타낼 수 있다.

$$\begin{aligned} J_{MSE}^{PCA} &= E\{\|x - \sum_{j=1}^m (w_j^T x) w_j\|^2\} \\ &= E\{\|x\|^2\} - E\{\|\sum_{j=1}^m (w_j^T x) w_j\|^2\} \\ &= \text{trace}(C_x) - \sum_{j=1}^m (w_j^T C_x w_j) \end{aligned} \quad \text{식 (6)}$$

식(6)은 $\sum_{j=1}^m w_j^T C_x w_j$ 가 최대가 될 때 J_{MSE}^{PCA} 는 최소가 되고, $w_1 = e_1, w_2 = e_2, \dots, w_m = e_m$ 일 때

$\sum_{j=1}^m w_j^T C_x w_j$ 가 최대가 된다. 이때 e_1, e_2, \dots, e_m 는 C_x 의 고유벡터이다.

결국 분산최대를 이용한 PCA 방법과 최소 MSE를 이용한 PCA방법을 통해 구한 기저는 데이터의 Covariance 행렬로 동일하다는 것을 알 수 있다.

4. 실험 결과 및 분석

본 방법은 PentiumIV, Windows 2000환경에서 Visual C++ 6.0과 MatLab 6.5로 구현하였다. 사용 데이터는 고정 초점 CCD 카메라를 통해 그레이 레벨의 320×240의 해상도로 획득된 26명의 320개의 데이터를 사용하였다. 임의의 130개의 학습데이터와 190개의 실험데이터를 사용하였고 [그림 2]는 학습에 사용된 몇 개의 대표 홍채 영역 이미지는 다음과 같다.



그림 2. 홍채 원영상

영상을 분석하는 방법으로 PCA만을 단독으로 사용하는 경우 어느 한 집단은 함축적으로 표현하는 경우에는 효과적인 방법이지만 집단간의 분리는 잘 하지 못 한다. 이러한 문제점으로 인하여 wavelet을 이용하여 눈 영상의 홍채 영역을 웨이블릿 분석을 이용하여 손실을 최대한으로 줄이고 가장 특징적인 것을 추출할 수 있다.

홍채의 특징영역은 웨이블릿 변환 결과의 Approximation영역과 Horizontal Detail영역의 값으로 특징벡터를 구성하였다. Approximation영역은 전체 홍채영상의 압축적인 결과를 나타내고 Horizontal Detail영역은 홍채 영상에서 가장 많은 것이 수직적인 성분을 가장 많이 가지고 있어 이 부분을 특징으로 사용한다. 이를 통해서 원 영상의 7200차원이 896차원으로 홍채의 특징을 가장 많이 포함하는 영역을 추출하였다. 이렇게 얻어진 결과를 PCA를 통하여 개개인을 표현할 수 있는 차원으로 특징의 차원을 129차원으로 축소시킬 수 있다.

인식율을 알아보기 위하여 SVM을 사용한 결과는 다음과 같았다.

| 사용커널 | 인식률 | 오인식률 |
|------------|-----------|----------|
| linear | 99.0625 % | 0.9375 % |
| gaussian | 99.375 % | 0.625 % |
| polynomial | 99.0625 % | 0.9375 % |

표 1. SVM을 이용한 인식률 결과

5. 결론

본 실험에서는 웨이블릿을 이용하여 홍채 영상이 가지는 주파수 성분을 분석하고 주성분분석(PCA)라는 통계적인 방법을 이용하여 홍채 특징을 추출하여 인식을 하는 방법을 제안하였다. 획득된 홍채영상에 대하여 wavelet을 적용했을 때, 획득된 영상에서 영상의 손실을 줄이면서 가장 특징으로써 사용될 수 있는 Approximation과 Horizontal부분을 사용함으로써 주된 특징으로 추출할 수 있다. 추출된 결과를 주성분분석을 적용하였을 때, 통계적으로 중요한 정보를 유지하고 잡영(noise) 정보를 효과적으로 제거하여 특징의 차원을 효과적으로 줄일 수 있었다.

이와 같은 방법으로 실험한 결과 원 영상에서 추출된 홍채 영역 7200차원을 웨이블릿 변환으로 896차원의 특징을 추출하였다. 이렇게 얻어진 특징차원을 주성분분석을 적용하여 129차원으로 줄일 수가 있었다. 이는 원 영상의 홍채영역이 가지는 값을 최대한 손실 없이 유지하면서 각 집단을 함축적으로 표현할 수 있다. 그 결과

Gaussian kernel을 이용한 SVM 분류기로 인식율 99.375%를 얻을 수 있었다.

6. 참고문헌

[1] R. Clarke, "Human Identification in Information Systems : Management challenges and public policy issues", *Information Technology & People*, 7(4), pp.6-37, 1994

[2] Daugman J. "High-Confidence Visual Recognition of Persons by a Test of Statistical Independence", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(11), pp.1148-1161, 1993

[3] Rafeai C. Gonzalez, Richard E. Woods, *Digital Image Processing*, 2nd Ed., Prentice Hall, pp349-404, 2001

[4] Richard O. Duda, Peter E. Hart, David G. Stork "Pattern Classification", Wiley Interscience, 2001

[5] Aha, D.W, "Machine Learning", A tutorial presented at the 5th International Workshop on Artificial intelligence & Statistics, pp 1~67, 1995

[6] Simon Haykin "Neural Networks a comprehensive foundation", Prentice Hall, 1999