

스케치 연산자를 이용한 얼굴 인식

최진***, 정윤수***, 유장희*

*한국전자통신연구원

**과학기술연합대학원대학교 정보보호공학과

Face Recognition Using Sketch Operator

Jean Choi***, Yunsu Chung***, Janghee Yoo*

*Electronics and Telecommunications Research Institute

**Information Security Engineering, University of Science & Technology.

E-mail : jchoi@etri.re.kr, yoonsu@etri.re.kr, jhy@etri.re.kr

Abstract

본 논문에서는 스케치 연산자를 적용하여 건설한 얼굴인식 방법을 제안한다. 제안된 방법은 인식 대상의 중요한 특성인 에지(edge), 벨리(valley) 및 질감(texture) 성분을 효과적으로 표현하기 위한 방법으로써, BDIP(block difference of inverse probabilities)를 사용하여 얼굴의 특징을 스케치 영상과 같이 나타내는 얼굴 영상을 획득한다. 그리고, BDIP 처리된 얼굴 영상은 입력 데이터의 차원 축소 및 얼굴 특징 벡터의 추출을 위해 PCA(Principal Component Analysis)를 수행한 후, Nearest Neighbor 분류기를 통해 인식을 수행한다. 제안된 방법의 성능을 평가하기 위하여, 일반적으로 많이 사용되는 HE(Histogram equalization)를 사용한 얼굴 인식 방법과의 비교를 수행한다. 실험결과, 본 논문에서 제안한 방법이 고유값이 적은 경우에 가장 높은 인식률을 나타내는 것을 알 수 있었다.

I. 서론

9/11 테러 이후 미국의 '미국 방문자 및 이민자 신분인식 기술'인 US-VISIT의 본격적인 도입이 이루어짐에 따라 생체인식 기술이 어느 때보다 크게 주목받으면서 지문, 음성, 정맥, 서명, 홍채, 얼굴 인식과 같은 생체 측정 방식을 이용한 연구가 국내외적으로 활발하

게 전개되고 있다. 이러한 생체인식 방법 중 얼굴 인식은 사용자가 가장 거부감의 적게 느끼며 활용도가 다양한 장점을 가지는 개인 인증 방식이다. 컴퓨터를 이용하여 얼굴 인식을 자동으로 처리하고자 하는 연구는 출입 통제 시스템이나, 기타 보안 시스템, 신용카드나 운전 면허증, 여권과 같은 개인 확인 시스템, 범죄자 관리, 그리고 미야 자동 검색 등 다양한 분야에 유용하게 활용될 수 있다. 얼굴 인식 방법에서 중요시 되어야 할 점은 표정, 연령, 시점, 조명의 변화, 잡음 등으로 인한 영상 변화에 영향을 받지 않고 안정적인 얼굴 인식 성능을 구현하는 것이다. 그러나 인식할 영상의 조명, 포즈, 표정 등의 변화에 민감하기 때문에 해결해야 할 문제로 남아있다.

본 논문에서는 입력된 얼굴 영상에 스케치 연산자를 적용하여 안정적인 얼굴인식 방법을 제안하였다. 제안된 방법은 입력 영상의 전처리, 얼굴 특징의 추출, 그리고 인식을 위한 분류 과정의 3 단계로 구성되어 있다. 첫 번째 단계에서는 에지 뿐만 아니라 물체의 인식에 있어서 중요한 특징인 벨리 및 질감을 효과적으로 추출하도록 하기 위해 스케치 연산자인 BDIP(block difference of inverse probabilities)를 사용하고 두 번째 단계에서는 입력 데이터의 차원을 줄이고 얼굴 특징 벡터를 구하기 위해 PCA(Principal Component Analysis)를 수행하였고, 마지막으로 Nearest Neighbor 분류기를 통해 인식률을 측정하였다.

본 논문의 구성은 다음과 같다. II장에서는 조명효과
의 제거 및 얼굴의 특징을 추출하기 위한 전처리 기
법을 소개하고, III장에서는 제안한 얼굴인식 방법을 설
명하고, 학습기법으로서 PCA 의 이론적인 원리를 소개
한다. IV장에서는 얼굴인식 실험을 수행하여 결과를 분
석하였다. V장에서는 결론을 맺고 추후 연구 방향을 제
시한다.

II. 스케치 연산자

기존의 얼굴 인식 방법에서 간단하고 효과적인 전
처리 방법으로써 보편적으로 사용하는 것이 Histogram
equalization(HE)이다. BDIP 를 사용했을 때 성능이 어느
정도 향상되는지에 대한 객관적인 지표를 얻기 위하여
HE 를 사용한 방법과 스케치 연산자를 적용한 방법을
비교하였다.

히스토그램은 이미지 명암값의 빈도수를 나타낸다.
HE 는 일정한 분포를 가진 히스토그램을 생성하는 것이
다. 즉, 명암값 분포를 재분배한다. HE 는 기존의 명암값
은 새로운 값으로 설정되지만 명암값의 실질적인 개수
는 입력영상의 명암값의 개수와 동일하게 된다. HE 의
처리 단계는 먼저 히스토그램을 생성하고 히스토그램의
정규화된 합을 계산한 후 마지막으로 입력 영상을 변형
하여 결과 영상을 생성한다. HE 의 효과는 영상이 어두
운 영역에 밀집되어 있거나 밝은 부분에 밀집되어 있는
경우에 효과적으로 영상을 향상한다.

$$HE(S_k) = \sum_{j=0}^k \frac{n_j}{n}, k = 0, 1, 2, \dots, L-1 \quad (1)$$

L 의 명암값을 갖는 이미지일 때 식 (1)와 같이 표현
된다. Sk 는 k 번째 명암값이다. 256 의 명암값을 갖는 이
미지일 경우 누적된 명암값을 전체 이미지의 화소수로
나눠서 새로운 값으로 대치시킨다.

스케치 연산자로 사용된 BDIP 는 각 마스크내의 평
균 밝기 변화량을 최대 휘도값으로 정규화한 형태를 가
지는 연산자로 입력 영상의 밝은 영역뿐만 아니라 어두
운 곳의 에지, 벨리, 질감을 강조하고 눈썹이나 입에서
의 질감 정보도 일부 표현하여 스케치 영상과 같은 결
과를 출력한다. 따라서 BDIP 를 적용한 얼굴 영상은 입
력된 얼굴 영상과 비교할 때 얼굴의 특징은 나타내 주

면서 조명 효과는 거의 제거한 상태가 된다.

BDIP 는 최근에 내용기반 영상검색에서 사용된 연
산자로 마스크내의 최대값으로 정규화 하는 과정을 통
해, Weber 의 법칙으로 알려진 바와 같이 밝은 영역에서
의 밝기 변화보다 어두운 영역에서 동일한 밝기 변화를
더 잘 인지하는 인간의 비선형적인 시각적 특성을 잘
반영한다. [1] BDIP 는 다음과 같이 각 마스크내의 평균
밝기 변화량을 최대 휘도값으로 정규화한 형태를 가진
다.[2]

$$BDIP(i, j) = \frac{\frac{1}{N_w^2} \sum_{(i-k, j-l) \in B} \left(\max_{(i-k, j-l) \in B} C(i-k, j-l) - C(i-k, j-l) \right)}{\max_{(i-k, j-l) \in B} C(i-k, j-l)} \quad (2)$$

(2)에서 B 는 $N_w \times N_w$ 크기, C(i-k,j-l)은 블록 픽셀의 명
도값을 나타낸다. DIP(Difference of Inverse Probability)는
영상의 에지와 벨리를 포함하는 스케치 특성을 추출하
는 연산자이다. DIP 는 윈도우 내의 전체 화소의 명도값
의 합에 대한 한 픽셀의 비율을 나타낸다. BDIP 는
DIP 의 블록 기반 버전이라고 할 수 있다.[2] 즉, $N_w \times N_w$
의 윈도우 중에 가장 큰 명암값을 가지는 화소값에서
윈도우 내의 모든 화소와의 차이값을 더한 후 윈도우의
크기로 나눠서 최대 명암값과 윈도우 내 전체 화소간의
차이의 평균을 구하고 다시 최대 명암값으로 나눠서 정
규화한 값이 BDIP 값이 된다



그림 1 입력 영상(왼쪽)과 DIP 적용 영상(중앙),
BDIP 적용 영상(오른쪽)의 예

<그림 1>에서 BDIP 적용 영상은 에지와 벨리를 모
두 추출하면서 어두운 부분의 에지와 벨리도 강조함을
알 수 있으며, 눈썹이나 입에서의 질감 정보도 일부 표
현함을 알 수 있다. 이와 같이 BDIP 는 에지만을 추출
하는 연산자에 비해, 에지만만 아니라 벨리도 잘 추출
하고 어두운 곳의 에지, 벨리, 질감을 강조하여 스케치
영상과 같은 결과를 출력한다. 이러한 BDIP 는 벨리를
중요시 하고 어두운 곳의 에지나 벨리에 민감한 인간
시각시스템의 인지 과정의 특성을 잘 반영함을 알 수
있다.

III.얼굴인식

PCA 는 평균과 분산까지의 통계적 성질을 이용한 2 차 통계적 기법이다.[3] PCA 는 입력 자료에 대하여 최대 공분산의 각 방향을 가리키는 직교 정규화된 일련의 축들의 집합을 찾는다. 이는 입력 자료의 가장 중요한 축들을 찾아 효율적으로 자료의 차원을 줄일 수 있는 장점을 갖게 됨을 의미한다. 그러나 PCA 는 2 차 통계 자료만을 사용하기 때문에 영상에서 가장 기본적인 특징이라 할 수 있는 에지를 나타내기가 어려운 단점이 있다.

주어진 자료를 X 라 놓고, n 개의 관찰된 표본이 있다면 X 를 $X=[x_1, x_2, \dots, x_n]$ 과 같이 정의 할 수 있다. 이때 X 의 각 표본 x_i 는 m 개의 데이터로 구성되어 있다. 만약 얼굴 영상의 경우라면 이때의 m 의 값은 해당 얼굴의 픽셀 수가 될 것이며 1 차원 벡터로 나타낼 수 있다. 계산의 편리를 위해 X 의 각 행의 값의 평균을 0 으로 만들어 주었다고 가정한다.

$$C_x = E\{(X_i - \mu_x)(X_i - \mu_x)^T\} \quad (3)$$

이때 주요 요인 축들은 (3)의 공분산 행렬의 고유값들이다. 이때 해당 고유값들은 그 축들의 분산값을 나타낸다. 첫째 주요 축은 최대의 분산값을 가지고, 둘째 주요 축은 최대의 분산 값을 가지고, 둘째 주요 축은 첫 번째 축에 수직하며 그 다음으로 가장 큰 값을 가진다. 이러한 방법으로 나머지 축들이 구성된다. 따라서 식 (3)을 이용하여 공분산 행렬의 고유벡터와 고유값을 구한 후에 고유값을 크기에 따라 정렬하고 이때 고유벡터도 해당 고유값의 위치대로 정렬한다. 일반적으로 뒤쪽에서 어느 정도의 고유값들은 0 또는 0 에 가까운 값을 가지게 되어 이를 버릴 수 있어 입력 데이터의 차원을 줄일 수 있다.



그림 2 훈련 얼굴 영상의 Eigenface

PCA 로 데이터를 표현하는 방법은 다음과 같다. 우선 PCA 로 데이터가 표현되는 것을 R 이라 놓고 이때 각 행이 원래 데이터의 표본에 매칭된다. 위에서 언급한 고유 벡터가 열로 들어있는 행렬을 V 라고 하였을

때 $R=XTV$ 와 같이 R 을 구할 수 있다. 고유벡터 V 가 대칭적이고, 직교 정규화되어 있으므로 $VVT=I$ 의 성질을 가지며, 역으로 데이터를 변환하는 것은 $XT=RVT$ 와 같이 구할 수 있다.

입력 얼굴 코드는 $R=[r_1, r_2, \dots, r_m]$ 이고, 임의의 학습 얼굴의 코드를 $T=[t_1, t_2, \dots, t_n]$ 라고 할 때, 유클리드 거리는 다음과 같다.

$$d = \sqrt{\sum_{i=1}^n (r_i - t_i)^2} \quad (4)$$

모든 학습 얼굴에 대하여 유클리드 거리를 구하여 가장 작은 d 값을 갖는 것을 출력으로 내는 방법이 Nearest Neighbor 방법이다.

IV. 실험

제안한 방법의 인식 성능 비교를 위해 자체적으로 제작한 얼굴 영상 데이터베이스를 이용하여 실험하였다. 이 데이터베이스는 64x64 크기의 55 명, 각 20 장의 총 1100 개의 256 그레이 레벨 이미지로 구성되어 있다. 각기 다른 표정과 조명의 방향에 따른 변화 그리고 안경 착용등과 같은 얼굴인식에 방해 요소로 작용하는 조건들을 고려한 영상들로 이루어져있다. 그림 3 은 데이터베이스의 영상 예이다. 대부분의 영상에서 얼굴은 영상의 중앙에 위치한다. 50 명의 영상 1000 개는 훈련 영상으로 나머지 100 개의 영상을 실험 영상으로 사용했다.



그림 3 실험에 사용된 데이터베이스의 예

우선 전처리를 하지 않은 1000 개의 얼굴 영상을 PCA 에 투영시켜 고유값을 찾는다. 훈련 영상에 적용할 PCA 고유값의 개수를 0 에서 300 까지 10 단위로 증가시키면서 각각의 Equal Error Rate(EER)과 이에 따른 인식률을 계산한다. 마찬가지로 방법으로 BDIP 를 적용한 실험 영상과 HE 를 적용한 실험 영상에 PCA 를 투영시켜서 각각의 고유값과 EER, 인식률을 계산한다.

에러율에는 False Acceptance Rate(FAR)과 False Reject Rate(FRR)이 있다. FAR 은 타인의 영상이 본인으로 인식되는 경우의 빈도를 의미하고 FRR 은 본인의 영상이 타인으로 인식되는 경우의 빈도를 말한다. FAR 과 FRR 의 정의는 다음과 같다.

$$FAR(\%) = \frac{I_A}{I_T} \times 100(\%)$$

$$FRR(\%) = \frac{C_R}{C_T} \times 100(\%) \quad (5)$$

여기서 I_A 는 타인이 본인으로 인식된 경우의 수이고 I_T 는 타인으로 인식실험을 한 전체 경우의 수이다. C_R 은 본인이 타인으로 인식된 경우의 수이고 C_T 은 본인으로 인식실험을 한 전체 경우의 수이다. 훈련 영상의 실험을 통해 PCA 의 고유값의 개수마다 본인과 타인을 나눌 수 있는 기준을 산출해야 한다. EER 은 FAR=FRR 이 되는 지점을 말하며 일반적으로 실험에서 본인과 타인을 구분하는 기준값으로 많이 쓰인다.

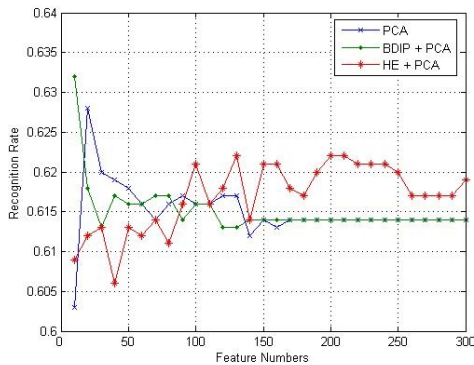


그림 4 인식률 비교 (고유값 개수 0~300)

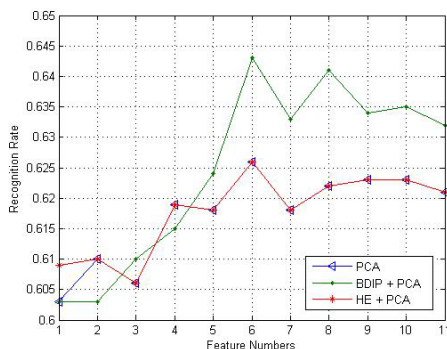


그림 5 인식률 비교 (고유값 개수 1~10)

고유값의 개수에 따른 인식률을 계산한 결과 개수가 150 이상인 경우 PCA 만을 적용한 방법과

BDIP+PCA 방법이 0.614 로 유사한 인식률을 나타내었고 HE+PCA 의 방법이 0.62 정도의 인식률을 나타내었다. 특이한 것은 고유값의 개수를 1 에서 10 까지 1 단위로 증가시키면서 인식률을 계산했을 경우 BDIP+PCA 의 경우가 고유값 개수가 6 일 때 0.645 의 인식률로 최고 인식률을 보인 것이다. 이것은 BDIP 가 원영상의 특징을 잘 표현했기 때문에 PCA 의 고유값 개수가 작더라도 다른 방법들에 비해 인식률이 증가된 것으로 볼 수 있다. 고유값의 개수가 적을수록 PCA 고유값의 차원이 작아지기 때문에 계산량이 감소되는 장점을 갖게 된다.

V. 결론

본 논문에서는 자체 제작한 데이터베이스를 이용하여 스케치 연산자를 사용한 얼굴 인식 방법을 제시하였다. 첫 번째 단계에서는 얼굴의 특징을 나타내기 위하여 스케치 연산자를 적용한다. 그 다음으로 전처리한 영상의 인식 대상의 얼굴 영상과 전체 훈련 영상의 평균 영상과의 차영상을 입력으로 PCA 를 수행하여 데이터의 차원을 줄이고 얼굴 특징 벡터를 구한다. 마지막으로 최소 유클리드 거리를 구하는 Nearest Neighbor 를 통하여 얼굴을 인식하였다. 실험 결과 작은 개수의 고유값을 사용하는 경우 BDIP 를 적용했을 때 가장 우수한 성능을 나타냄을 알 수 있었다. 향후 본 논문에서 제안한 얼굴 인식방법에 대해 좀 더 다양한 환경의 얼굴 데이터베이스에 대한 비교 연구가 필요하다.

참고문헌

- [1] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison Wesley, 1993.
- [2] Y. D. Chun, S. Y. Seo., and N. C. Kim, "Image retrieval using BDIP and BVLC moments", IEEE Trans. Circuits Syst. Video Techn., vol. 13, no.9, pp.951-957, Sep. 2003.
- [3] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
- [4] Martinez, A. M., Kak, A. C., "PCA versus LDA", IEEE Transactions, Pattern Analysis and Machine Intelligence, Volume 23, Issue 2, pp228-233, Feb. 2001.