

논문 2006-43SP-4-10

주파수 영역에서 에너지 확률을 이용한 얼굴 특징 추출

(Facial Feature Extraction Using Energy Probability in Frequency Domain)

최 진*, 정 윤 수**, 김 기 현**, 유 장 희**

(Jean Choi, Yns-Su Chung, Ki-Hyun Kim, and Jang-Hee Yoo)

요 약

본 논문에서는 얼굴 영상의 에너지 분포 특성을 이용한 새로운 특징추출 방법을 제안한다. 제안된 방법은 얼굴 영상의 에너지 확률과 에너지 맵을 이용해서 데이터 차원이 축소된 유효정보의 추출 및 유효정보의 LDA 해석에 기반을 둔다. 일반적으로, 얼굴 영상은 고유한 에너지 분포 특성을 가지고 있다. 그러나 기존의 많은 DCT 기반 방법들은 이러한 얼굴 영상의 특성을 효과적으로 이용하지 못하는 단점이 있다. 제안된 방법은 이러한 기존 방법의 단점을 개선하기 위해 다음의 3단계 방법을 사용한다. 먼저, DCT 도메인에서 얼굴의 에너지 확률 개념을 정의하고, 이러한 에너지 확률로부터 얼굴의 에너지 맵을 생성한다. 마지막으로, 에너지 확률 지도에 위치한 주파수 계수들에 대한 LDA 적용 및 해석을 통하여 특징 벡터 추출 및 인식을 수행한다. 제안된 방법은 ETRI 데이터베이스에서 96.8%, ORL 데이터베이스에서 100%의 인식률을 보인다. 실험을 통하여 인식 성능의 개선뿐만 아니라, 특징 벡터의 차원 축소에도 효과가 있음을 알 수 있다.

Abstract

In this paper, we propose a novel feature extraction method for face recognition, based on Discrete Cosine Transform (DCT), Energy Probability (EP), and Linear Discriminant Analysis (LDA). We define an energy probability as magnitude of effective information and it is used to create a frequency mask in DCT domain. The feature extraction method consists of three steps; i) the spatial domain of face images is transformed into the frequency domain called DCT domain; ii) energy property is applied on DCT domain that acquire from face image for the purpose of dimension reduction of data and optimization of valid information; iii) in order to obtain the most significant and invariant feature of face images, LDA is applied to the data extracted using frequency mask. In experiments, the recognition rate is 96.8% in ETRI database and 100% in ORL database. The proposed method has been shown improvements on the dimension reduction of feature space and the face recognition over the previously proposed methods.

Keywords : 얼굴 인식, DCT(Discrete Cosine Transform), LDA(Linear Discriminant Analysis), 특징추출

I. 서 론

얼굴 인식에 있어서 특징 추출은 데이터의 분류 성능을 높이면서 입력 데이터의 차원을 축소하고, 데이터의 차원 축소를 통해 분류 단계에서의 계산 복잡도를 감소시키는 역할을 수행한다. 일반적으로, 이러한 특징 추출

방법에는 공간 영역에서의 특징추출 방법과 주파수 영역에서의 특징추출 방법이 있다. PCA와 LDA는 이러한 공간 영역에서의 특징 추출을 위한 대표적인 방법이다. 그러나 이러한 방법들은 공간 영역에서의 데이터의 통계적 해석에 기반한 방법으로써 얼굴 영상의 주파수 특성을 활용하지 못하는 약점을 가지고 있다. 이러한 약점을 극복하기 위한 노력으로 DCT(Discrete Cosine Transform)를 이용한 주파수 영역에서의 특징 추출 관련 연구가 최근 증가하고 있다.

최근 주파수 영역에서의 특징추출에 관련된 연구 중 하나로 Hafed와 Levine에 의해 DCT가 사용되었다^[1]. DCT계수 중 저주파 7x7영역을 사용하는 것은 얼굴 특

* 학생회원, 과학기술연합대학원대학교 정보보호공학 (University of Science and Technology)

** 정회원, 한국전자통신연구원 바이오인식집중연구팀 (Electronics and Telecommunocations Research Institute)

접수일자: 2006년2월24일, 수정완료일: 2006년6월16일

징 추출에서 가장 대표적으로 이용되는 PCA와 흡사한 수준의 인식률을 보인다. 이것은 DCT가 얼굴 인식에 이용될 수 있다는 가능성을 보였다. 이는 의미 있으나 인식률 면에서 LDA를 비롯한 그 변형들의 성능과 비교할 때 실제 얼굴 인식에 사용될 수 있을 정도의 인식률을 보이지 못한다. 이를 시작으로 DCT 영역을 여러개의 밴드로 분할하여 특징을 추출하는 Jing과 Zhang의 방법^[2], DCT 영역을 지그재그 스캔 한 후 상위 55개의 DCT 계수를 추출하여 특징 벡터로 사용하는 방법이 소개되고 있다^{[3], [4]}. 이전 방법에 비해 많은 성능 향상을 보이긴 했으나 얼굴 영상에 최적화되었다고 보기는 어렵다. 이와 같이 이전에 연구된 방법들은 DCT 계수를 단순히 블록 혹은 밴드로 분할하여 사용하였기 때문에 얼굴 고유의 특성을 반영하는 유효 정보만을 포함했다고 할 수 없다. 이러한 특징추출은 적합한 기준이나 척도에 의한 유효 정보의 선택이 아니므로 얼굴인식에 필요한 정보와 불필요한 정보가 공존할 수밖에 없다. 즉, 불필요한 정보의 비율이 낮을수록 좋은 인식 성능을 보이게 되며 필요 정보만을 취사선택하여 특징 추출에 이용한다면 이전의 방법들에 비해 좋은 성능을 기대할 수 있다.

이와 같은 방법들의 변환 계수 선택 방법은 저주파수 대역에 에너지가 몰려 있다는 일반적인 영상의 특성을 전제하고 있다. 따라서 이러한 방법들은 수직, 대각 주파수 대역 보다는 수평 주파수 대역에 에너지가 더 밀집하는 얼굴 영상의 특성을 극대화해서 활용하지 못하는 단점을 가지고 있다. 본 논문에서는 이러한 이전 연구의 단점을 극복하기 위하여, 특정 주파수 계수를 중심으로 에너지가 밀집하는 얼굴 영상의 특성을 이용한다. 또한 유효정보의 기준으로 주파수 영역의 특성 중 하나인 에너지를 이용한 에너지 확률을 정의하고 에너지 확률에 의해 이전 연구 방법들의 블록 혹은 밴드에 상응하는 에너지 맵을 생성하여 필요 정보만을 선택하는 특징추출 방법을 제안한다.

제안된 방법은 다음의 3단계로 구성된다. 먼저, DCT를 이용하여 얼굴 영상을 주파수 변환한다. 다음으로 주파수 변환된 얼굴 영상으로부터 에너지 확률을 획득한 후 유효정보만을 포함하는 에너지 맵을 생성하고 이를 통해 축소된 데이터를 추출한다. 마지막으로 LDA를 통하여 최종 특징 벡터를 추출한다. 추출된 특징 벡터의 성능을 평가하기 위해 Nearest Neighbor Classifier를 사용하여 본인 여부를 판단하고 인식률을 측정한다. 이와 같이 제안된 방법은 얼굴 영상의 고유

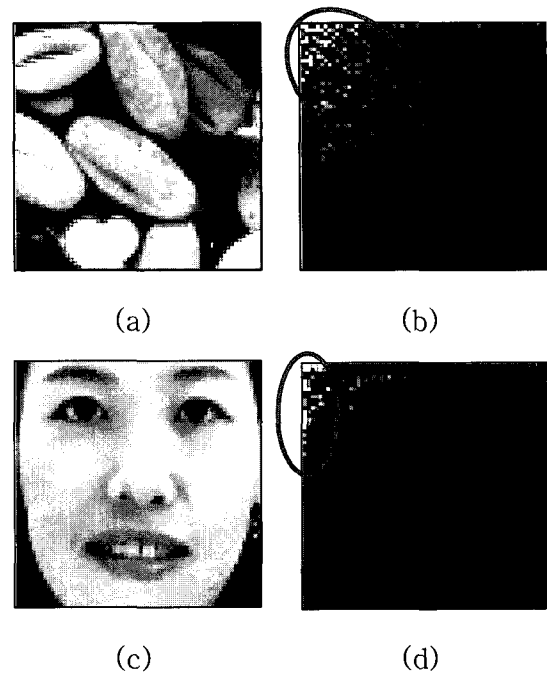


그림 1. 일반 영상(a)과 DCT 변환 (b) 예와 얼굴 영상(c)과 DCT 변환 (d) 예

Fig. 1. Example of a general image (a) and its DCT transformed image (b), and example of face image (c) and its DCT transformed image (d).

한 특성을 최대한 반영할 수 있도록 에너지 확률을 이용하여 특징벡터를 추출하기 때문에 얼굴 인식의 성능을 높일 수 있다는 장점이 있다.

본 논문의 구성은 다음과 같다. II장에서는 본 논문에서 정의한 에너지 확률과 에너지 맵에 대한 설명을 하고, III장에서는 제안된 에너지 확률과 에너지 맵을 이용한 특징추출 및 이를 이용한 얼굴인식 방법을 설명한다. IV장에서는 제안된 방법과 기존 방법들 간의 비교 실험을 통해 결과를 분석한다. 마지막으로, 결론 및 추후 연구 방향을 V장에서 제시한다.

II. 에너지 확률과 에너지 맵

본 장에서는 얼굴의 유효정보를 측정하는 기준으로 정의한 에너지 확률과 이를 이용한 에너지 맵의 정의 및 생성 방법에 대해 상세히 설명한다.

1. 에너지 확률 (Energy Probability, EP)

얼굴 영상과 일반적인 비얼굴 영상을 공간 영역에서 주파수 영역으로 변환하면 얼굴 영상의 주파수 변환된 영역은 비얼굴 영상과 비교할 때 일정한 패턴에 따라

변환 계수가 분포하는 것을 관찰 할 수 있다. 예를 들어 얼굴 영상과 비얼굴 영상을 주파수 영역으로 변환한 종류 중 하나인 DCT 계수 분포를 비교 해보면 그림 1과 같이 일반 영상의 DCT 계수인 (b)의 경우 DCT 계수가 큰 부분이 대각선 상단의 좌측 부분에 고르게 분포되는 것을 볼 수 있다. 반면에 얼굴 영상의 DCT 계수인 (d)의 경우 DCT 계수의 큰 부분이 상단의 가장 좌측으로 몰려 있는 것을 확인 할 수 있다. 일반적으로 DCT계수의 좌측 열 부분은 영상의 수평 변화에 관련된 특성을, 상단의 행 부분은 영상의 수직 변화에 관련된 특성을 나타낸다^[5]. 이와 같이 얼굴 영상은 주파수 영역에서 얼굴 영상만의 고유한 주파수 분포 특징을 갖는다. 얼굴의 이러한 특징을 최적화 하는 영역을 선택하기 위해서는 얼굴 인식에 유효한 정보인지를 가늠할 수 있는 기준이 필요하다. 본 논문에서는 얼굴 인식에 대한 유효정보의 판단 기준으로 에너지 확률을 정의한다. 에너지는 신호처리 분야에서 신호가 가지는 특징을 나타내기 위해 사용하는 신호 수량이며 신호를 비교하는데 유용하게 쓰인다. 영상처리 분야에서 주파수 크기의 측정 기준으로 사용되는 에너지의 정의는 다음과 같다^{[5],[6]}.

$$Energy_F = \sum_{u=1}^N \sum_{v=1}^N |F(u,v)|^2 \quad (1)$$

$F(u,v)$ 는 영상의 주파수 변환 결과 이며 N 은 영상의 가로, 세로 크기이다^[7]. $Energy_F$ 는 하나의 영상에 대해서 하나의 값을 갖는다. 즉, 각각의 영상 자체에 대한 특징은 될 수 있으나 얼굴의 고유한 주파수 분포 패턴을 최적화 하는 영역을 추출하는데 사용하기에는 부적합하다. 따라서 에너지를 이용한 얼굴의 특징을 추출하기 위해 다음과 같이 $EP(u,v)$ 을 정의하며, 이를 에너지 확률 이라 한다.

$$EP(u,v) = \frac{|F(u,v)|^2}{Energy_F} \quad (2)$$

$Energy_F$ 는 (1)에서 정의한 주파수 변환된 결과에 대한 하나의 에너지 값이며 계산된 $EP(u,v)$ 크기는 유효한 정보에 대한 측정 기준으로 사용된다. $N \times N$ 의 영상변환 결과인 N^2 개의 $EP(u,v)$ 값이 의미 하는 것은 전체 이미지가 갖는 에너지 값 중 u,v 의 위치에서 갖

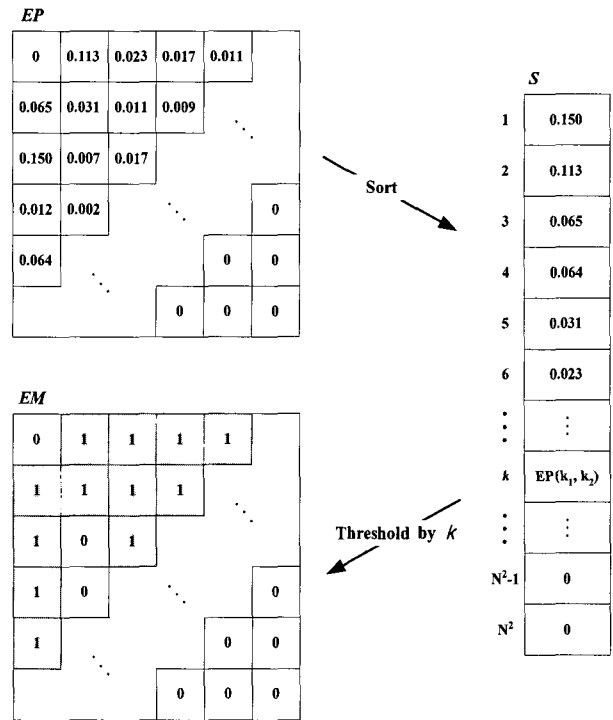


그림 2. 에너지 맵의 생성 과정
Fig. 2. The creation procedure of energy map.

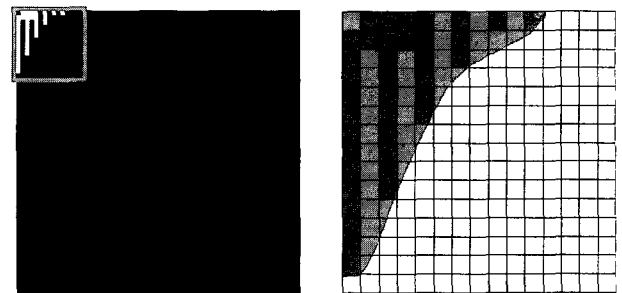


그림 3. 실제 생성된 주파수 마스크 (a) 와 15x15 영역의 확대(b)
Fig. 3. (a) is the generated the Energy Map and (b) is its magnification of 15x15 block.

는 에너지 값의 비율이다. 즉, $EP(u,v)$ 의 값이 크다는 것은 u,v 의 위치에서 에너지가 차지하는 비율이 높다는 것을 의미하며 $EP(u,v)$ 의 값이 커질수록 얼굴 특징으로서의 정보를 많이 포함하는 것으로 생각할 수 있다. 따라서 일정한 값 이상의 $EP(u,v)$ 의 위치만을 선택하면 얼굴의 고유한 주파수 분포 패턴을 최적화하는 영역을 추출 할 수 있게 된다.

2. 에너지 맵 (Energy Map, EM)

에너지 맵은 에너지 확률을 이용해서 추출한 유효한 정보의 패턴 위치를 기억한 정보이다. 에너지 확률이

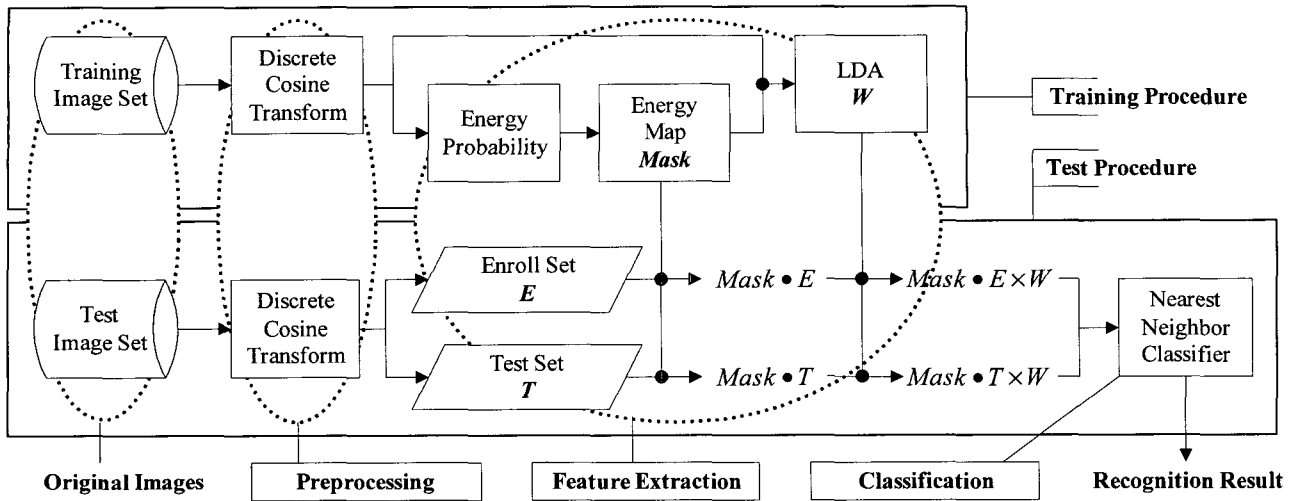


그림 4. 제안된 방법을 이용한 얼굴 인식의 흐름도
 Fig. 4. Face Recognition Procedure of the Proposed Method.

주파수 영역에서 $N \times N$ 의 각 화소가 얼굴인식에 얼마나 유용한지를 나타내는 기준이라면 에너지 맵은 그 기준에 의해 유효하다고 판정된 화소의 위치를 모아놓은 정보이다. 얼굴 영상 하나에 대해 에너지 맵을 생성한다면 해당 얼굴 영상의 주파수 분포 패턴만을 최적화하게 된다. 그러나 앞서도 언급한 것처럼 얼굴 영상과 비얼굴 영상의 주파수 분포 패턴을 비교해 보면 얼굴 영상의 고유한 특징이 존재한다. 따라서 에너지 맵을 한 장의 얼굴 영상이 아니라 여러 장에 대해 생성한다면 일반적인 얼굴의 에너지 맵을 만들 수 있다. 각각의 얼굴 영상에 대해 각각의 에너지 맵을 생성할 것인가 또는 여러 장의 얼굴 영상을 이용해서 하나의 에너지 맵을 생성할 것인가에 대한 문제는 시스템의 설계 방법에 따라 달라질 것이다. 그것은 인식 혹은 인증, 검색 등의 목적에 맞게 선택하게 되는 것이므로 본 논문에서는 에너지 맵 생성의 기본이 되는 하나의 영상에 대한 하나의 맵 생성에 대해 설명한다.

에너지 맵 생성의 구체적인 방법은 다음과 같다. $N \times N$ 크기의 얼굴 영상을 주파수 변환 한 후 계산된 에너지 확률 EP의 2차원 정보는 그 크기에 의해 내림차순 정렬된 1차원 정보인 S로 변환된다. 에너지 확률이 정렬된 열 벡터 S에서 에너지 확률이 큰 순서부터 k 번째까지를 의미 있는 유효정보로 선택하며 k는 다음의 조건을 만족한다. 여기서 α 는 얼굴 인식 성능을 최대로 하는 임계값이며 일반적으로 0.6에서 0.75 사이의 값으로 실험에 의해 결정한다.

$$\sum_{i=1}^k S(i) \leq \alpha \quad (3)$$

$$EM(u, v) = \begin{cases} 1, & \text{if } EP(u, v) \geq S(k) \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

(3)의 조건을 만족하는 상위 k개가 선택되면 선택된 k개의 값은 1로 나머지는 0으로 하는 EM을 생성한다.

그림 2는 에너지 맵의 생성과정을 도식화 한 것이며 그림 3은 실제 생성된 에너지 맵(a)과 유효정보로 선택된 영역이 분포하는 15×15 영역을 확대 (b)한 그림이다. 생성된 에너지 맵에서 확인 할 수 있는 것처럼 유효정보의 주파수 분포 패턴은 전체적으로 좌측 상단에 삼각 형태로 분포하며 세로로 긴 막대 모양의 조합으로 생성된다. 대부분의 얼굴 영상은 이러한 형태를 기본으로 하는 특징을 갖는다.

III. 제안한 방법을 이용한 얼굴 인식

본 논문에서 제안한 얼굴 인식 방법은 입력 영상에 대한 전처리, 특징추출, 분류의 3단계로 나누어진다. 그림 4는 제안된 방법의 얼굴 인식 전체 과정을 나타낸 것이다. 첫 번째로 얼굴 영상의 전처리 부분에서는 공간 영역의 얼굴 영상을 주파수 영역으로 변환하기 위해 DCT를 적용하였다. 두 번째로 얼굴 특징추출 부분에서는 에너지 확률 기반의 에너지 맵을 통해 입력 데이터의 차원을 축소하며 LDA를 통하여 최종 특징 벡터를 추출 한다. 마지막으로 분류 과정에서는 유클리디언 거리 (Euclidean Distance)를 이용한 Nearest Neighbor Classifier을 사용하여 본인 여부를 판단 한다. 본 장에서는 제안된 방법에서 주파수 영역으로 변환하기 위해

사용한 DCT와 선형변환을 위해 적용한 LDA에 대해 자세히 알아보고 얼굴 인식 전체 과정에 대한 설명을 한다.

1. 전처리

입력된 얼굴 영상은 눈썹 윗부분부터 입과 턱의 중간 부분까지 얼굴의 윤곽선이 포함되도록 검출된 것을 사용한다. 실험에 사용되는 얼굴 영상은 훈련용과 실험용으로 나누어진다. 다음으로 용도별로 분리된 얼굴 영상의 공간 영역을 주파수 영역으로 변환해야 한다. 본 논문에서는 DCT를 이용하였다.

DCT는 디지털 신호처리 분야에서 다양한 문제들의 해법으로 폭넓게 이용되며 N개의 데이터를 N개의 Cosine 함수의 합으로 표현하는 방식이다. $f(x,y)$ 로 표현되는 공간 영역에서의 영상의 크기는 $N \times N$ 이다. DCT를 적용한 후 $f(x,y)$ 는 $F(u,v)$ 로 변환되며 $1 \leq u,v,x,y \leq N$ 이다. $u,v,x,y=1,2,3 \dots N$ 이고 x,y 는 공간 영역에서의 좌표쌍이며 u,v 는 DCT변환된 주파수 영역에서의 좌표쌍이다^[7]. 2차원 DCT의 정의는 다음과 같다.

$$F(u,v) = \alpha(u)\alpha(v) \sum_{x=1}^N \sum_{y=1}^N f(x,y) \times \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \tag{5}$$

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}}, & u,v=1 \\ \sqrt{\frac{2}{N}}, & otherwise. \end{cases} \tag{6}$$

DCT 계수 중 $F(1,1)$ 의 계수는 DC(Direct Current) 성분으로 영상의 평균값을 의미하며 DC를 제외한 나머지는 AC(Alternate Current) 계수이다. 영상의 DC 계수는 보통 영상의 밝기와 관련이 있다고 알려져 있다. 따라서 얼굴 인식의 방해요소로 작용하는 조명의 영향은 DC 계수를 제거함으로써 감소시킬 수 있다^[2]. 제안된 방법은 영상의 조명 특징을 표현하는 DC값을 제외하기 위해 0으로 하고 나머지 AC 값만을 이용한다.

2. 특징추출

두 번째 단계는 용도에 따라 분리된 영상에서 얼굴 인식에 사용할 최종 특징 벡터를 추출하는 것이다. 특징추출 단계는 우선 훈련 과정을 통해 에너지 맵과

LDA 변환 행렬을 생성한 후, 실험 과정에 이를 적용하는 두 부분으로 이루어진다.

LDA는 PCA와 더불어 대표적인 얼굴 인식 방법 중 하나이다. 선형 변환의 일종으로 클래스 분류를 목적으로 설계되기 때문에 높은 인식 성능을 보이는 장점을 가진다. LDA의 가장 기본적인 개념은 각 클래스 내부의 분산은 최소로 하면서 클래스 간의 분산은 최대로 하는 $y = w^T x$ 의 w 을 찾는 것이다. 클래스 내부의 분산을 나타내는 행렬(Within Class Scatter Matrix)을 S_w 라 하고 클래스 간의 분산을 나타내는 행렬(Between Scatter Matrix) S_b 라 하며, S_w 와 S_b 를 이용하여 w 를 찾는 함수 $f(w)$ 의 정의는 다음과 같다^{[8],[9],[10]}.

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T \tag{7}$$

$$S_b = \sum_{j=1}^c (\mu_j - \mu)(\mu_j - \mu)^T \tag{8}$$

$$f(w) = \frac{w^T S_b w}{w^T S_w w} f(w) = \frac{W^T}{W} \tag{9}$$

x_i^j 는 j 번째 클래스의 i 번째 샘플이며 μ_j 는 j 번째 클래스의 평균이다. N_j 는 j 번째 클래스의 데이터 수이며 c 는 클래스의 수, μ 는 전체 클래스의 평균이다. 각 클래스 내부의 분산은 최소로 하면서 클래스 간의 분산은 최대로 하는 것을 목표로 하는 w 를 계산하기 위해서 $\max f(w)$ 를 이용한다.

DCT가 적용된 학습용 얼굴 영상에 II장에서 정의한 에너지 확률을 계산 한 후, 이를 이용하여 에너지 맵인 Mask를 생성한다. 생성된 에너지 맵을 DCT 변환된 학습 영상에 적용한다. 에너지 맵을 통해 선택된 유효정보의 수가 k 이고 학습 영상의 수가 M 개라 할 때 $k \times M$ 크기의 Mask 벡터가 생성된다. 에너지 맵을 이용해서 얻은 Mask 벡터에 대해 LDA를 적용하면 앞에서 정의한 $\max f(w)$ 를 이용해서 행렬 W 를 계산할 수 있다. 학습을 통해 Mask 벡터와 W 행렬을 구하였으므로 인식률을 얻기 위한 분류 과정을 수행하도록 한다. 실험용으로 분류된 영상에 학습과정과 마찬가지로 DCT를 적용하고 이것은 다시 등록 셋 E와 테스트

셋 T 로 분할한다. Mask 벡터는 유효정보의 위치를 열 벡터로 저장하고 있으므로 데이터에 대한 적용을 $Mask \cdot E$ 와 같이 표현한다. 학습과정을 통해 계산된 Mask 벡터와 W 벡터를 E 와 T 에 각각 적용하면 $Mask \cdot E \times W$ 와 $Mask \cdot T \times W$ 로 계산되는 최종 특징 벡터를 얻을 수 있다.

3. 분류

마지막으로 등록과 테스트 각각의 과정을 통해 추출된 특징 벡터를 유클리디언 거리를 이용해 유사도를 측정한다. 유클리디언 거리 D 에 대한 정의는 다음과 같다

$$D = \sqrt{\sum_{i=1}^l (Mask \cdot E \times W_i - Mask \cdot T \times W_i)^2} \quad (10)$$

실험하고자 하는 영상에 대해 등록된 모든 영상의 유사도를 측정하고 최소 이웃 분류법을 사용하여 유사도가 가장 높은 영상을 통해 본인 여부를 판단한다.

IV. 실험 및 결과

본 장에서는 제안한 방법의 객관적인 성능 평가를 위하여 두 가지의 얼굴 영상 데이터베이스를 사용하였다. ETRI 데이터베이스와 얼굴 인식 시스템의 성능 평가를 위해 대표적으로 사용되는 Olivetti Research Laboratory (ORL) 데이터베이스이다. 제안된 방법은 PCA를 이용해 데이터의 차원을 축소하고 와 LDA를 이용해서 특징을 추출하는 방법과 DCT의 좌측 상단을 블록으로 설정하여 데이터의 차원을 축소하고 LDA를 이용해서 최종 특징을 추출하는 방법을 통해 비교 실험한다. 이러한 방법을 통해 제안한 방법의 객관적인 성능 평가를 할 수 있다.

1. ETRI Database

ETRI 데이터베이스는 64×64 크기의 55명, 각 20장씩 총 1100장으로 구성되어 있으며 256 명암도 영상이다. 각기 다른 표정과 조명의 방향에 따른 변화 그리고 안경 착용 등과 같은 얼굴인식에 방해 요소로 작용하는 조건들을 고려한 영상이 포함되어 있다. 그림 5는 실험에 사용된 데이터베이스의 얼굴 영상 예이다. 대부분의 영상에서 얼굴은 영상의 중앙에 위치한다. 전체 영상 중 30명의 영상 600장은 훈련 영상으로 나머지 500장의 영상은 실험 영상으로 사용한다. 얼굴 인식을 위한 특징



그림 5. 실험에 사용된 ETRI 얼굴 영상의 예
Fig. 5. Example faces in the ETRI database.

추출은 크게 두 단계로 나누어진다. 먼저 얼굴 영상으로부터 얻은 높은 차원의 얼굴 데이터를 효율적으로 가공하기 위한 데이터 축소 과정과 클래스간의 변별력을 높이기 위해 가장 효과적인 방법으로 알려진 LDA를 적용해서 얼굴의 최종 특징 벡터를 추출하는 과정이다. 본 논문에서 제안한 방법은 데이터의 차원은 최대한 줄이면서 유효한 정보의 최적화를 통해 원래의 데이터가 가지고 있던 정보의 손실을 최소화하는데 그 목적이 있다. 따라서 객관적인 성능 평가를 위해 데이터의 차원 감소에 얼굴 인식에서 가장 폭넓게 이용되는 PCA와 기존의 DCT 방법의 성능을 측정하여 제안한 방법과 비교하였다.

그림 6의 (a)는 축소된 데이터의 수를 고정하고 LDA를 통해 추출한 최종 특징 벡터의 수에 대한 인식률을 나타낸 것이다. PCA의 Principal Component 수와 DCT의 계수의 수는 64로 고정했으며 제안된 방법의 축소된 데이터의 수는 46개이다. 그래프에서 확인 할 수 있는 것처럼 최종 특징 벡터의 수가 많을수록 인식률은 높아진다. 본 실험에서는 LDA를 통해 추출된 특징 벡터의 수가 30일 때 공통적으로 가장 높은 성능을 보이는 것으로 나타났다. 다음으로 그림 6의 (b)는 LDA를 통해 최종 추출된 특징 벡터의 수가 30일 때 축소된 데이터의 개수에 대한 인식률을 그래프로 나타낸 것이다. 축소된 데이터의 범위는 0에서 100개이며 제안된 방법의 성능이 가장 좋은 것을 확인 할 수 있다. 그래프에서 PCA와 DCT 계수를 이용한 방법은 40을 기점으로 약간의 인식률의 향상을 보인다. 반면에 제안된 방법은 데이터의 수가 30일 때를 기점으로 인식률의 감소가 뚜렷하다. 실험 결과에 의하면 제안된 방법이 최적화된 유효한 정보만을 사용할 때는 향상을 보이지만 불필요한 정보, 즉 유효하지 않은 정보를 포함할 때는 성능의 저하를 유발하는 것임을 알 수 있다.

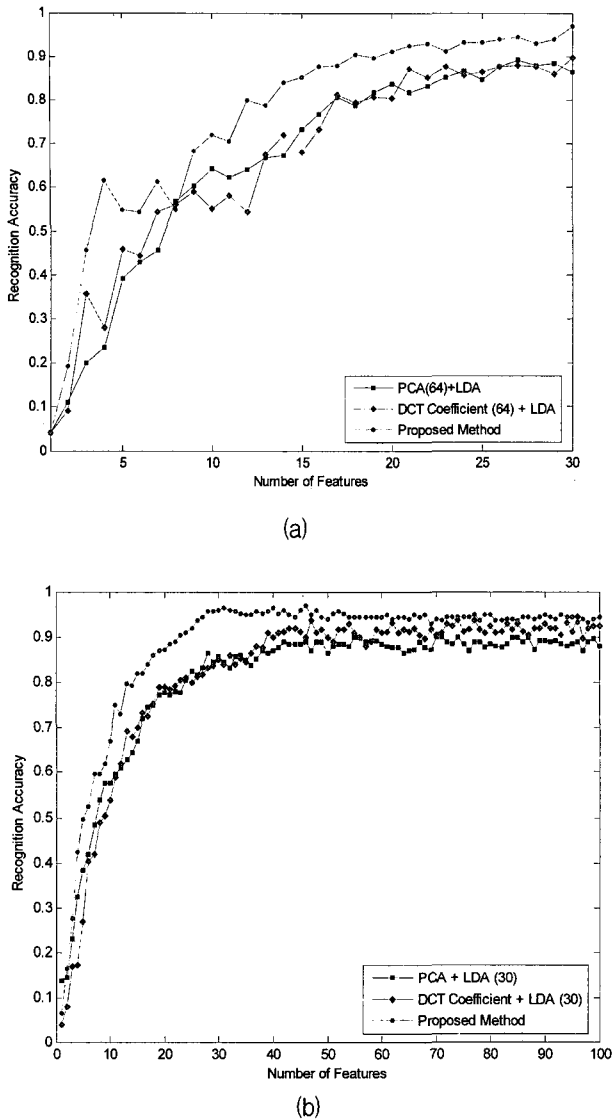


그림 6. (a)는 1) PCA, 2) DCT의 좌측 상단 8×8 영역 3) 에너지 맵을 이용하여 입력된 데이터의 차원을 64로 고정했을 때, LDA를 통해 최종 추출된 특징 벡터의 개수에 대한 인식률. (b)는 (a)와는 반대로 LDA를 통해 최종 추출되는 특징 벡터의 개수를 30으로 고정하고 각각의 데이터 축소 방법을 이용해서 축소된 데이터의 개수에 대한 인식률

Fig. 6. (a) is recognition accuracy with the number of reduced data, 64 and the increments of the number of extracted features (b) is recognition accuracy with the increments of the number of reduce data and number of extracted features, 30.

2. ORL Database

ORL 데이터베이스는 92×112 크기의 40명, 각 10장씩 총 400장으로 구성되어 있으며 256명압도 영상이다. 각 영상은 실험에 사용하기 위하여 64×64 의 이미지로 크

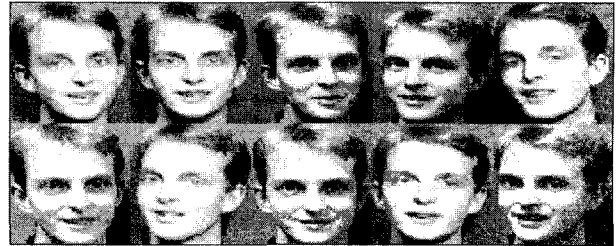


그림 7. 실험에 사용된 ORL얼굴 영상의 예
Fig. 7. Example faces in the ORL database.

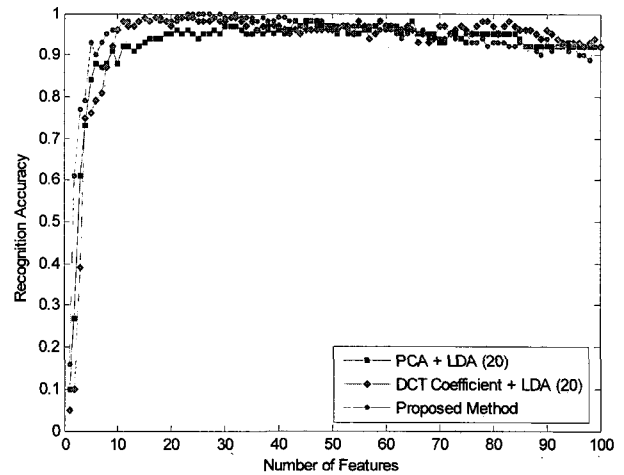


그림 8. LDA를 통해 최종 추출된 특징 벡터의 개수를 20으로 고정 했을 때, 각각의 데이터 축소 방법을 이용해서 축소된 데이터의 개수에 대한 인식률
Fig. 8. Recognition accuracy with the increments of the number of reduce data and number of extracted features, 20.

기를 조정하였다. 각각의 이미지는 조명과 표정, 포즈, 안경 착용 등에 변화가 있다. 모든 이미지는 동일한 어두운 배경에 정면 조명을 이용해서 촬영 되었다. 얼굴의 경사도와 회전각은 20% 이내이며 크기의 변화는 10% 이내로 조정된 영상이다. 그림 7은 실제 실험에 사용된 얼굴 영상 중 한 사람의 이미지 열 장을 나타낸 것이다. 전체 영상 중 20명의 영상 200장은 훈련 영상으로 나머지 200장의 영상은 실험 영상으로 사용한다.

그림 7은 LDA를 통해 최종 추출 된 특징 벡터의 수가 20일 때 축소된 데이터의 개수에 대한 인식률을 그래프로 나타낸 것이다. 축소된 데이터의 범위는 0에서 100개이다. 그래프에서 제안된 방법과 DCT 계수를 이용한 방법은 비슷한 인식률을 보인다. 실험결과 ORL 데이터베이스는 전반적으로 고르게 높은 인식률을 보인다.

3. 결과 분석

ETRI 데이터베이스의 경우, PCA, DCT 계수, 제안

표 1. ETRI 데이터베이스에서 성능분석 및 비교
Table 1. Comparison of classification performance using the ETRI database.

Method	PCA+LDA	DCT+LDA	Proposed Method
Reduced Data Dimension	74	96	46
Extracted Feature Dimension	29	30	30
Recognition Accuracy	0.900	0.944	0.968

표 2. ORL 데이터베이스에서 성능분석 및 비교
Table 2. Comparison of classification performance using the ORL database.

Method	PCA+LDA	DCT+LDA	Proposed Method
Reduced Data Dimension	45	23	23
Extracted Feature Dimension	20	20	19
Recognition Accuracy	0.980	0.990	1.000

된 방법을 데이터의 차원 축소 방법으로 사용하고 특징 벡터의 추출 방법으로 LDA를 사용했을 때 각각의 방법의 최고 인식률은 표 1에 나타난 것과 같이 90.0%, 94.4%, 96.8%이다. 또한 LDA에 적용된 축소된 데이터 데이터의 수는 각각 74, 96, 46개 이다. 제안된 방법이 가장 적은 차원이 데이터로 가장 높은 인식률을 나타낸 것을 알 수 있다.

ORL 데이터베이스의 경우, PCA, DCT 계수, 제안된 방법 모두 고르게 높은 인식률을 나타내었다. 수치상으로는 제안된 방법이 100%, PCA 방법이 98%, DCT 계수 방법이 99%를 나타내었으나 그 차이가 크지 않고 그래프 상으로도 유사한 경향을 보인다. 이것은 'state-of-the-art'로 불리는 LDA에 의한 결과이다. 두 데이터베이스에서 공통적으로 관찰 할 수 있는 것은 LDA에 입력되는 데이터의 수가 30 정도까지는 인식률의 급격한 증가를 보이지만 그 이후로는 완만하게 감소한다는 것이다. 이것은 유효한 정보 이외의 정보가 첨가되면 오히려 얼굴 인식에 방해요소로 작용해서 성능이 저하되는 결과를 도출하는 것을 의미한다. 따라서 본 논문에서 제안한 얼굴의 고유한 에너지 분포 특성을

이용한 특징추출이 얼굴 인식에 더 적합한 것을 알 수 있다.

V. 결론

본 논문에서는 얼굴 영상의 에너지 분포 특성을 이용한 새로운 특징추출 방법을 제안하였다, 얼굴의 에너지 분포 특성을 이용하지 못하는 기존 방법들의 단점을 극복하기 위하여, 제안된 방법은 얼굴 영상의 에너지 확률과 에너지 맵을 이용한 유효 벡터 추출 및 이러한 유효 벡터의 LDA 해석에 기반을 둔 특징추출 및 인식을 수행하였다. ETRI 및 ORL 얼굴 데이터베이스에 대한 실험 분석을 통하여 제안된 방법이 객관적인 인식 성능에서 뚜렷한 개선뿐만 아니라, 특징 벡터의 차원 축소에도 효과적임을 알 수 있었다.

참고 문헌

- [1] Z.M. Hafeed and M.D.Levine, "Face Recognition Using the Discrete Cosine Transform", *International Journal of Computer Vision* 43(3), 167 - 188, 2001.
- [2] Xiao-Yuan Jing and David Zhang, "A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction", *IEEE Transactions on Systems, Man, and Cybernetics-part b: Cybernetics*, vol. 34, no. 6, December 2004.
- [3] Meng Joo Er, Weilong Chen, and Shiqian Wu, "High-Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks", *IEEE Transactions on Neural Networks*, vol. 16, no. 3, May 2005.
- [4] Meng Joo Er, and Shiqian Wu, "PCA and LDA in DCT domain", *Pattern Recognition Letters*, vol. 26, 2474-2482, July 2005.
- [5] W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc., New York, 2001.
- [6] Gordon E. Carlson, *Signal and Linear System Analysis*, John Wiley & Sons, Inc., New York, 1998.
- [7] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison Wesley, Massachusetts, 1992.
- [8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces versus fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach.*

Intell., vol. 19, no. 7, pp. 711 - 720, Jul. 1997.
 [9] A. Martinez and A. Kak, "PCA versus LDA," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228 - 233, 2001.
 [10] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognit.*, vol. 34, no. 12, pp. 2067 - 2070, 2001.

저 자 소 개



최진(학생회원)
 2004년 숙명여자대학교 컴퓨터과학과 학사
 2004년 8월~현재 과학기술연합 대학원대학교 정보보호 공학전공 석사 과정
 <주관심분야 : 바이오인식, 패턴 인식, 영상처리>



정윤수(정회원)
 1993년 경북대학교 전자공학과 학사
 1995년 경북대학교 전자공학과 석사
 1998년 경북대학교 전자공학과 박사
 1999년~현재 한국전자통신연구원
 2005년~현재 SC37 Korea 생체인식전문위 전문위원,
 TTA PG103 바이오인식 프로젝트그룹 부의장
 <주관심분야 : 영상처리, 바이오인식, HCI, HRI>



김기현(정회원)
 1989년 경북대학교 전자공학과 학사
 1991년 경북대학교 컴퓨터공학과 석사
 1991년 2월~현재 한국전자통신연구원 근무
 <주관심분야 : 영상정보처리, 생체인식, 정보보호>



유장희(정회원)
 1989년 한국외국어대학교 전자물리학과 학사
 1990년 한국외국어대학교 경영정보대학원 전산학과 석사
 2004년 영국 University of Southampton 전자 및 컴퓨터 박사
 1989년 11월~현재 한국전자통신연구원 근무
 <주관심분야 : 실시간 영상처리, 패턴인식, 생체인식, Robotics 및 HCI>