논문 2010-47SP-1-6

극좌표계 변환에 기반한 얼굴 인식 방법

(Face Recognition Based on Polar Coordinate Transform)

오 재 현*, 곽 노 준**

(Jae-Hyun Oh and Nojun Kwak)

요 약

본 논문에서는 기존 평행좌표를 이용하는 얼굴영상 대신 극 좌표계 변환을 이용한 얼굴 인식 방법을 제안한다. 이 방법은 얼굴의 중심부분의 한 점을 극으로 삼아 이 점을 기준으로 360도의 각 방향으로 일정 길이만큼 얼굴 영상을 샘플링하여 새로운 얼굴 영상을 제작하고 이를 바탕으로 기존의 특징 추출 방법들을 이용하여 얼굴 인식의 성능을 높인다. 극 좌표계의 특성상 극에 가까운 부분은 세밀하게 묘사되고 극에서 멀리 떨어질수록 영상의 정확도가 떨어진다. 일반적으로 얼굴 영상은 얼굴의 중심부에 가까운 영역에 눈, 코, 입 등의 주요 부위가 밀집되어 있다. 따라서 이러한 극 좌표계를 얼굴영상에 적용한다면 같은 화소를 이용하는 기존 평행좌표를 사용할 때보다 눈, 코, 입 등 주요 부위를 보다 세밀하게 표현할 수 있다는 장점을 갖는다. 제안된 방법을 Yale데이터와 FRGC데이터에 적용한 후 기존의 특징 추출 방법인 LDA와 NLDA를 이용하여 얼굴인식을 수행한 결과 평행좌표에 기반한 원 영상을 그대로 사용했을 때 보다 인식률이 향상됨을 확인할 수 있었다.

Abstract

In this paper, we propose a novel method for face recognition which uses polar coordinate instead of conventional cartesian coordinate. Among the central area of a face, we select a point as a pole and make a polar image of a face by evenly sampling pixels in each direction of 360 degrees around the pole. By applying conventional feature extraction methods to the polar image, the recognition rates are improved. The polar coordinate delineates near-pole area more vividly than the area far from the pole. In a face, important regions such as eyes, nose and mouth are concentrated on the central part of a face. Therefore, the polar coordinate of a face image can achieve more vivid representation of important facial regions compared to the conventional cartesian coordinate. The proposed polar coordinate transform was applied to Yale and FRGC databases and LDA and NLDA were used to extract features afterwards. The experimental results show that the proposed method performs better than the conventional cartesian images.

Keywords : Polar coordinate, transfrom, important feature, LDA, NLDA

I. 서 론

얼굴인식 분야에서 패턴은 입력 변수의 집합으로 표현될 수 있으며, 얼굴인식을 잘하기 위해 입력 변수에서 새로운 특징을 만들어 내는 특징 추출이 중요하다. 이 때, 서로 다른 각각의 얼굴을 분류해 낼 수 있는 정보를 유지한 채 입력 변수의 차원을 줄이는 것이 필요하다. 입력 차원을 줄이면, 특징 벡터를 간결하게 다루는 효과를 얻으며, 입력 차원이 커질 때 발생하는 차원의 저주(the curse of dimensionality) 같은 문제를 해결하는 효과를 얻을 수 있다. 얼굴 인식 분야에서 많이 사용되는 차원 감소(dimensionality reduction) 방법에는 아공간(subspace) 방법들이 많이 사용되며 이들 중 대표적인 것으로 LDA (Linear Discriminant Analysis)[1], PCA (Principal Component Analysis)[2], 그리고 ICA (Independent Component Analysis)[3] 등의 방법이 있다. 이 중 LDA는 각 클래스 간 떨어진 정도(class separability)를 최대화하는 것이 목적이다. LDA

* 학생회원, ** 정회원, 아주대학교 전자공학과
(Division of Electrical and Computer Engineering, Ajou University)

※ 이 연구는 2008년도 한국학술진흥재단 지원(KRF-2008-313-D00942)과 2009학년도 아주대학교 교내 연구비 지원(20094720)에 의하여 이루어졌음.
접수일자: 2009년7월27일, 수정완료일: 2009년12월28일

는 특징 벡터의 차원을 감소시키는 기법 중의 하나로서, 클래스간 분산과 클래스내 분산의 비율을 최대화하는 방법을 이용하여 데이터에 대한 특징 벡터의 차원을 축소한다. LDA의 변형인 NLDA (null space based LDA)는 클래스내 분산의 영공간(null space)에서 클래스간 분산을 최대화 하는 것이 목적이다. 이는 클래스간 분산의 사영된 결과가 0이 아닌 한, 클래스내 분산의 영공간에 서로 다른 클래스들을 분류 할 수 있는 중요한 정보가 많이 들어있다는 점에서 나온 개념이다. PCA는 재구성 오차(reconstruction error)를 최소화 하는 방법을 사용한다.

대부분의 영상 인식 방법들은 원 영상에 위에서 설명한 특징 추출 방법을 적용하여 특징을 추출한 후 분류기(classifier)를 통해 최종 분류 작업을 시행하는 것이 일반적이며, 특히 얼굴 인식에서는 대부분 평행좌표에 기반한 원 영상을 이용한다. 이에 반해 극 좌표계 변환 방법은 홍채 영상에서의 특징 검출[4]이나 물체 인식[5]을 위해 사용되었다. 본 논문에서는 극좌표계 변환 방법을 이용하여 얼굴 인식률을 향상 시키는 방법을 제시하고자 한다. 극좌표계 변환 방법은 얼굴을 분류하기 위한 중요한 특징들의 대부분은 얼굴의 중심부 근처에 위치한다는 사실에 기반한다. 이러한 사실은 원 얼굴 영상에서 머리카락 같은 얼굴의 외곽부분은 제외하고, 눈, 코, 입과 같은 얼굴 부분만을 포함하게 얼굴 영상을 잘라내어 이렇게 잘려진 얼굴 영상의 인식률과 원 얼굴 영상의 인식률을 비교해 보면 확인할 수 있다. 즉 원영상에 대해 LDA, NLDA 등의 특징 추출 방법을 적용해 얻은 얼굴 인식률과 머리카락 같은 얼굴의 외곽부분을 제거하고 눈, 코, 입 주변 부위만 잘라낸 영상에 동일한 특징 추출 방법을 적용해 얻은 인식률을 비교해 볼 때 일반적으로 후자의 경우가 인식률이 더 높은 것을 확인할 수 있다.

본 논문에서 사용하는 극 좌표계 변환은 얼굴의 중심부를 기준점으로 일정한 반지름을 갖는 원을 그리면서, 각 360도 방향으로 각각 원의 중심에서부터 반지름의 끝 부분까지 샘플링을 하여 얼굴 데이터를 재구성한다. 본 논문의 실험에서는 360도의 방향 중 같은 간격으로 총 90개의 방향을 사용하여 실험을 수행하였다. 만약 극좌표계와 평행좌표계(원영상)에서 사용하는 화소수가 같다면, 극좌표계를 이용할 때 얼굴을 분류하는데 중요한 역할을 하는 눈, 코, 입 등 얼굴의 중앙 부분이 원 영상을 그대로 사용할 때 보다 더 자세히 표현되므로 얼굴인식의 인식률의 향상을 도모할 수 있다.

이 논문의 각 장의 내용은 다음과 같다. 먼저 Ⅱ장에서는 얼굴 인식에서 차원 감소 방법 중에서 본 논문의 실험에서 사용하는 LDA와 NLDA 방법을 간단히 살펴보고, Ⅲ장에서는 극좌표계 변환 방법과 인식률 향상의 이론적 검토에 대해 살펴본다. 제Ⅳ장에서는 실험결과를 분석하며, 제Ⅴ장에서는 결론을 제시한다.

Ⅱ. LDA와 NLDA를 이용한 얼굴 인식

얼굴인식 분야에서 패턴은 입력 변수의 집합으로 표현될 수 있으며, 일반적으로 분류에 사용되는 입력 변수의 수가 너무 많으므로 클래스를 분류해 낼 수 있는 정보를 유지한 채 입력 변수의 차원을 줄이는 특징 추출이 중요하다. 이 장에서는 본 논문에서 사용한 특징 추출 방법으로 얼굴 인식에서 차원을 감소시키기 위해 일반적으로 사용되는 방법인 LDA와 NLDA에 대해 설명한다.

1. LDA를 이용한 얼굴 인식

LDA는 클래스간 분산 행렬(between-class scatter matrix)과 클래스내 분산 행렬(within-class scatter matrix) 간의 비율을 최대화하는 방법을 통해 데이터에 대한 특징벡터의 차원을 줄이는 방법이다. 즉 클래스 분리를 최대화 하는 벡터의 방향을 찾아 그 방향으로 사영을 시켜 차원을 감소시키게 된다. 같은 클래스내의 데이터들은 가능한 한 서로 가깝게 모여 있고, 동시에 다른 클래스의 평균 사이의 거리는 최대화하는 방향을 찾게 된다[6].

LDA에서는 최적화된 사영을 찾기 위해 클래스내 분산 행렬인 S_w와 클래스간 분산 행렬 S_B의 개념을 도입한다. 열벡터 $x_k(k=1,\cdots,N)$들로 이루어진 크기가 N인 표본 집합에 대하여 S_w와 S_B는 아래와 같이 정의된다[7].

$$S_i = \sum_{x_k \in w_i} (x_k - \mu_i)(x_k - \mu_i)^T, \mu_i = \frac{1}{N_i} \sum_{x_k \in w_i} x_k \quad (1)$$

$$S_w = \sum_{i=1}^{C} S_i \quad (2)$$

$$S_B = \sum_{i=1}^{C} N_i (\mu_i - \mu)(\mu_i - \mu)^T$$

$$\mu = \frac{1}{N}\sum_{k=1}^{N} x_k = \frac{1}{N}\sum_{x_k \in w_i} N_i \mu_i \quad (3)$$

여기에서 w_i는 i번째 클래스를 나타낸다. 또한 C 는 클래스의 수, N_i은 각 클래스에 속하는 샘플의 수, N 은 모든 표본 집합의 개수를 나타낸다.

LDA의 목적을 달성하기 위해 목적 함수 $J(w)$을 클래스간 중심이 멀리 떨어져 있게 하기 위해 데이터의 중심간 거리, 즉 클래스의 평균의 차이를 최대화하고, 동일 클래스의 표본들은 인접하게 분포할 수 있도록 LDA의 개념을 고려하여 (4)와 같이 정의한다.

$$J(W) = \frac{W^T S_B W}{W^T S_w W} \quad (4)$$

위의 $J(W)$ 값을 최대로 하는 W값은 $S_w^{-1}S_B$의 고유값 문제(eigenvalue problem)를 풀어서 얻을 수 있으며, 구체적으로 가장 큰 고유값에 해당하는 고유벡터를 구함으로써 해를 얻을 수 있다.

2. NLDA를 이용한 얼굴 인식

NLDA는 클래스내 분산행렬의 영공간에서 클래스간 분산행렬을 최대화 하는 방법을 통해 특징벡터의 차원을 줄이는 방법[8]이다. 이 방법은 클래스간 분산행렬에 의한 사영이 0이 되지 않는 경우에, 클래스내 분산행렬의 영공간에 클래스들을 분류할 수 있는 가장 중요한 정보가 포함되어 있다는 개념에 기반한다. LDA의 경우 클래스의 수를 C라고 했을 때, C-1개의 특징을 뽑아낼 수 있지만, NLDA의 경우 이러한 제한 없이 특징을 뽑아낼 수 있다는 장점이 있다. NLDA의 목적함수는 다음과 같이 정의된다.

$$\begin{aligned} &W = \arg max_W trace(W^T S_b W) \\ &s.t.\ W^T S_W W = 0,\ W^T W = I \end{aligned} \quad (5)$$

먼저 S_W의 영공간은 $V^T S_w V = 0$ $(V^T V = I)$을 통해 구할 수 있다. 그리고 S_b를 S_w의 영공간으로 사영을 시키면 $\widetilde{S}_b = V^T S_b V$ 결과를 얻을 수 있다. $\widetilde{S}_b$의 고유값 Λ중 가장 큰 값에 대해 고유벡터 U를 구한다. 최종 목적함수 W는 V와 U의 곱을 통해 구할 수 있다[9].

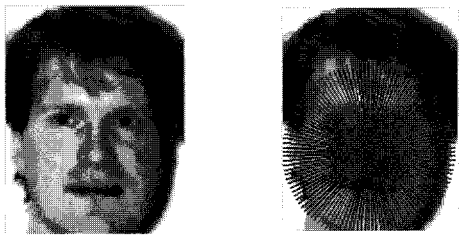
Ⅲ. 극좌표계 변환 방법과 인식률 향상의 이론적 검토

1. 극좌표계 변환 방법

얼굴 영상을 분류하기 위해서는 각 얼굴의 중요한 정보를 담고 있는 특징들을 뽑아내는 작업이 중요하다. 얼굴 인식에서 일반적인 얼굴 영상을 그대로 사용하게 되는 경우 영상의 모든 픽셀을 동일한 중요도를 갖고 차원벡터를 줄이게 되므로 머리 스타일의 변화가 있거나 모자를 쓴 경우에는 인식률이 저하될 수 있다.

극 좌표계 변환은 얼굴의 중심부의 한 점을 극으로 삼아 이 점을 기준으로 360도의 각 방향으로 일정 길이만큼 얼굴 영상을 각각 원의 중심에서부터 반지름의 끝부분까지 샘플링을 하여 얼굴 데이터를 재구성한다. 일반적으로 얼굴 영상은 얼굴의 중심부에 가까운 영역에 눈, 코, 입 등의 주요 부위가 밀접해 있게 된다. 이러한 극좌표계 변환 방법을 통해 얼굴을 분류하는데 중요한 특징들을 원 영상을 그대로 사용할 때 보다 더 많이 고려하게 되어 인식률의 향상을 도모할 수 있다. 이 방법은 특징 추출 전 입력 영상을 변환하는 방법으로서 모든 특징 추출 방법에 적용이 가능하며, 본 논문에서는 LDA와 NLDA를 대상으로 실험을 수행한다.

그림 1은 Yale데이터의 일부로서 원 영상과 극좌표계 변환방법을 사용하기 위한 개념을 나타낸 그림이다. 그림 1(b)는 얼굴의 중심부분의 고정된 한 점인 극을 중심으로 반지름을 40픽셀로 하여 각 방향으로 픽셀 값을 읽는 것을 나타낸 것이다. 이 때 원의 중심점인 극의 위치를 정하는 것이 중요한데 본 논문에서는 그림 1(a)와 같은 동일한 크기의 원 영상에서 코끝의 위치를 지정한 후 이 점을 극으로 삼았으며, 코끝의 위치를 직접



(a) (b)

그림 1. 얼굴 영상의 원 영상(a)과 반지름을 40픽셀로 하여 극좌표계 변환을 하는 방법(b)

Fig. 1. The original face image(a) and the method of polar coordinate transform with radius of 40 pixels(b).

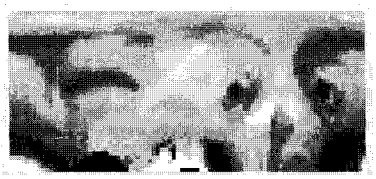


그림 2. 원 영상에 대해 반지름을 40픽셀로 극 좌표계 변환을 통해 얻은 영상

Fig. 2. The result of polar coordinate transform with radius of 40 pixels.

지정해 준 후, 각 영상에 대해 일률적으로 이 점을 중심으로 극좌표 변환을 수행하였다.

실험을 수행한 Yale데이터의 경우 극의 좌표는 (x, y) = (40, 65)로 정했으며, FRGC데이터의 경우 (x, y) = (30, 40)을 극의 위치로 정했다. 그림 1(b)는 360도의 방향 중 90개의 방향만을 동일한 간격으로 나타난 그림이다. 얼굴 중심의 극으로부터 외각 부분으로 픽셀 값을 샘플링 할 경우 얼굴의 중심 근처에서는 두 점이 겹쳐 나타나는 현상이 나타나는데, 이 경우에는 보간법(interpolation) 방법 중 쌍이차(bicubic) 보간법을 사용하여 픽셀 값을 구하였다.

그림 2는 원 영상에 대해 반지름을 40픽셀로 극 좌표계 변환을 통해 얻은 영상이다. 그림 2의 극좌표계 변환된 영상은 40x90 픽셀의 크기로 세로는 극좌표계 변환 시 반지름의 크기이고, 가로는 360도의 방향 중 90개의 방향에 해당하는 값들이 저장된다. 그림2의 아래쪽에 검은색으로 나타난 부분은 픽셀 값을 읽는 과정에서 원영상의 외부로 나갔기 때문에 해당 픽셀의 값을 읽을 수 없어, 0값으로 채워놓은 것을 나타낸 것이다.

2. 인식률 향상의 이론적 검토

극좌표계 변환 방법은 얼굴을 분류하기 위한 중요한 특징들의 대부분은 얼굴의 중심부의 주변에 위치한다는 사실에 기반 한다. 이는 그림 3(a)와 같은 원 얼굴 영상과 그림 3(b)와 같이 머리카락 같은 얼굴의 외곽부분은 제외하고, 눈, 코, 입과 같은 얼굴 부분만을 포함하게 잘라낸 영상의 인식률을 비교해 보면 알 수 있다. 그림 3에 나타낸 Yale 데이터에 대해 LDA, NLDA를 사용하여 인식률을 나타내었으며, 각각 14개의 특징을 사용하였다. 특징을 추출한 후 1-최근접 이웃(1-nearest neighbor) 알고리즘으로 분류를 수행하였으며, 11겹 교차 검증(11-fold cross validation)에 대한 인식률을 비교해 보면 표 1에서 볼 수 있듯이 후자의 경우가 인

(a) (b)

그림 3. 원 영상 100x80 픽셀(a)과 원 영상을 얼굴 중심 주변을 55x50 픽셀로 잘라낸 영상(b)

Fig. 3. The original face image of 100x80 pixels (a) and the result of cropping near at the center of face with 55x50 pixels(b).

표 1. Yale 데이터에 대해 원 영상과 얼굴의 중심부를 제외한 머리카락 부분을 제거한 영상에 대해 LDA와 NLDA를 수행한 인식률

Table 1. The recognition rate for original image and cropped image for Yale database using LDA and NLDA.

	LDA	NLDA
original	78.78	87.88
crop image	86.05	89.07

식률이 더 높은 것을 알 수 있다. 즉 얼굴의 코 주변의 중심부분에 위치한 특징들이 얼굴 외곽에 위치한 머리카락 등의 정보보다 더 중요한 특징이라는 것을 확인할 수 있다. 즉 극좌표 변환 방법은 얼굴 중심 부분의 사람을 분류하는데 중요한 특징에 대해 가중치를 더 줌으로써, 얼굴 분류에 오히려 방해요소로 작용할 수 있는 모자를 쓰거나, 머리 스타일의 변화 등의 인식률 저하의 요인을 줄일 수 있게 된다. 극좌표 변환 방법이 얼굴의 중심 부위에 더 가중치를 두었다는 사실은 그림 1을 통해서도 확인할 수 있는데, 얼굴의 중심부에 위치한 극 주변에서는 샘플링을 한 선의 밀도가 외각 부분보다 더 높은 것을 볼 수 있다.

그림 4는 FRGC 데이터의 일부이며, 상위 10개의 사진이 학습영상으로 사용된 영상이고, 하위 10개의 사진이 테스트 영상으로 사용된 영상이다. 하위 10개의 사진 중 왼쪽에서 3번째와 4번째 사진의 경우 원영상에 대해 그대로 LDA를 적용한 경우 학습 영상에 모자를 착용한 사진을 정확히 인식을 하지 못하는 것을 볼 수 있다. 반면 극좌표 변환 방법을 이용한 경우 정확하게 인식을 하는 것을 볼 수 있다.

또한 그림 5에서 볼 수 있듯이 Yale 데이터의 경우 사진의 배경에 그림자가 진 영상이 포함되어 있기 때문

그림 4. FRGC 데이터의 학습 영상, 테스트 영상과 LDA를 수행한 결과를 원 영상과 극좌표 변환을 통한 영상에 대해 나타낸 그림

Fig. 4. A Training image, a testing image and the recognition rate for original face image and image using polar coordinate transform for FRGC data.

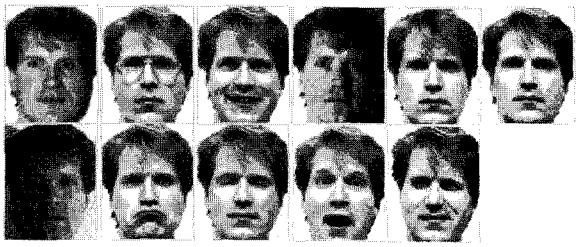


그림 5. Yale 데이터 중 일부

Fig. 5. A part of Yale database.

에 극좌표 변환 방법을 이용하면 이러한 그림자가 진 영상의 인식률 향상에도 도움이 된다. 즉 얼굴 외곽 주변에 그림자가 진 부분에 대한 가중치를 적게 두고, 얼굴 중심 주변의 눈, 코, 입 같은 특징 값들의 가중치를 크게 줌으로써 인식률 향상을 도모할 수 있다.

Ⅳ. 실 험

가. 학습 데이터의 입력

얼굴 인식에서 사용되는 Yale데이터와 FRGC데이터는 하나의 행이 한 사람의 얼굴 영상의 정보를 포함하고 있다. Yale데이터는 총 165장의 얼굴 영상으로 구성되어 있으며, 15명의 사람에 대해 각각 11장씩 구성되어 있다. 각각의 영상의 크기는 100×80 픽셀의 크기로 구성되어 있다. FRGC데이터의 경우 총 40명의 영상에 학습 데이터와 테스트 데이터로 400장을 사용하였다. 각 영상의 크기는 80x60 픽셀의 크기로 구성되었다. 극좌표 변환을 한 영상과 원영상 각각에 대해 위의 개수만큼의 데이터를 사용하였으며, 극좌표 변환을 할 때는 코끝의 위치를 중심으로 반지름을 변화시키면서 실험을 하였다.

실험을 수행한 Yale데이터의 경우 극의 좌표는 (x, y) = (40, 65)로 정했으며, FRGC데이터의 경우 (x, y) = (30, 40)을 극의 위치로 정했다. 히스토그램 평활화를 한 영상에 대해서도 같은 방법으로 실험을 수행하였다.

나. 정규화

정규화는 얼굴 인식에서 외부 환경 요인에 의해 인식률을 저하시키는 바이어스를 제거시키는 과정으로, 무의미하고 중복되는 정보에 의한 인식률 저하를 막을 수 있다[10]. LDA의 경우 정규화를 해줌으로써 평균을 0으로 만들어 정확한 고유 벡터의 방향을 찾을 수 있다. 각 행이 하나의 사람에 대한 정보를 담고 있는 데이터 행렬 Tr에 대해서 각 행을 i로 표현하고 열을 j로 표현하면 아래와 같은 방법으로 정규화를 수행한다. 여기서 행렬 Tr은 세로는 특정 사람을 의미하며 가로는 픽셀을 의미한다. 즉 행렬 Tr의 한 열에 한 사람의 픽셀 정보가 담겨 있으며, 세로는 사람 수 만큼의 크기를 갖게 된다. 식 (6)에서 $mean(j)$는 j번째 픽셀값의 평균, $std(j)$는 j번째 픽셀값의 표준편차를 의미하며 Tr_{normal}은 정규화된 얼굴 데이터 행렬을 나타낸다.

$$Tr_{normal}(i,j) = \frac{Tr(i,j) - mean(j)}{std(j)} \qquad (6)$$

다. PCA

LDA, NLDA를 통한 얼굴 인식을 수행할 때 LDA, NLDA를 바로 적용하면, 학습 집합에 없는 데이터 표본, 학습 집합과 다른 배경을 가진 데이터 표본이 주어졌을 때 성능이 떨어진다. 또한 고차원의 영상 벡터가 입력으로 들어오게 되면 방대한 계산으로 인해 어려움

을 초래한다. 따라서 PCA를 통해 차원을 감소시키고 얼굴인식을 수행한다[11][12]. 본 실험에서는 PCA를 이용하여 100차원으로 차원을 감소시킨 후 얼굴 인식을 수행한다.

라. LDA

PCA를 수행하여 차원을 감소시킨 결과를 이용하여 S_w와 S_b을 계산한다. 열벡터 x로 이루어진 얼굴의 표본 집합에 대해 식 (2)와 식 (3)을 통해 S_w와 S_b을 계산한다.

S_w의 고유값과 고유벡터를 구하고, 고유값이 큰 순서대로 나열하여 각각 그에 해당하는 고유벡터를 각 열에 해당하도록 행렬을 만든다. 즉 첫 번째 행에는 고유값이 가장 큰 값에 해당하는 고유벡터가 들어가고, 두 번째 행에는 고유값이 그 다음으로 큰 값에 해당하는 고유벡터가 들어간다. 이를 Mat_{sorted}로 놓는다.

고유값을 순서가 큰 순서대로 나열하여 대각 성분의 값에 고유 값의 루트 값의 역수를 대입하여 행렬을 만든다. 즉 대각 성분은 고유 값이 큰 순서대로 고유 값의 루트의 역수를 대입한 값이고, 대각 성분 외의 나머지 성분은 0이 되는 행렬을 만들게 된다. 즉 역행렬을 만드는 과정을 거치는 것이고, 이 행렬을 $Mat_{invroot-eig}$로 놓는다.

$$W_1 = Mat_{sorted} * Mat_{invroot-eig} \tag{7}$$

$$S_{w1} = W_1' * S_w * W_1 \tag{8}$$

$$S_{b1} = W_1' * S_b * W_1 \tag{9}$$

S_{b1}에 대해서도 고유값 분해를 수행하여 W_2을 구한다.

$$W_{lda} = W_1 * W_2 \tag{10}$$

마. NLDA

PCA를 수행하여 차원을 감소시킨 결과를 이용하여 식 (5)의 목적함수를 만족하는 W를 구한다[9].

먼저 S_W의 영공간은 식 (11)을 통해 구할 수 있다. 그리고 S_b를 S_w의 영공간으로 사영을 시키면 식 (12)와 같은 결과를 얻을 수 있다. $\widetilde{S}_b$의 고유값 중 가장 큰 값Λ에 대해 고유벡터 U를 구한다.

$$V^T S_w V = 0 \quad s.t \quad V^T V = I \tag{11}$$

$$\widetilde{S}_b = V^T S_b V \tag{12}$$

$$U^T \widetilde{S}_b U = \Lambda \tag{13}$$

최종 목적함수 W는 V와 U의 곱을 통해 구할 수 있다.

$$W_{nlda} = VU \tag{14}$$

바. 사영(Data projection)

최적화된 변환행렬인 W_{lda} 및 W_{nlda}와 PCA를 이용하여 차원을 감소시킨 결과와 곱을 하여 사영을 하고, 클래스 분리를 최대화 하는 벡터의 방향을 찾아 그 방향으로 사영을 시켜 차원을 감소시킨다. 같은 클래스 내의 데이터들은 서로 가깝게 모여 있고, 동시에 다른 클래스의 평균간의 거리는 최대화하는 방향을 찾게 된다.

사. 테스트 데이터의 입력과 정규화

입력된 테스트 데이터에 대하여 정규화를 수행한다. 식 (15)에서 $mean(j)$는 (나)단계에서 구한 학습에 이용된 행렬 Tr의 j번째 픽셀값의 평균, $std(j)$는 행렬 Tr의 j번째 픽셀의 표준편차를 나타내고 $Test$와 $Test_{normal}$은 각각 정규화되기 전후의 테스트 데이터 행렬을 나타낸다.

$$Test_{normal}(i,j) = \frac{Test(i,j) - mean(j)}{std(j)} \tag{15}$$

아. 사영(Data projection)

학습과정에서 구한 weight 값을 이용하여 테스트 데이터와 곱하여 사영을 수행한다.

자. 분류(Classification)

K-최근접 이웃 알고리즘은 기계학습 알고리즘에서 가장 흔하게 사용되는 방법으로서 자신의 이웃하는 데이터들의 대다수에 해당하는 클래스 값으로 분류되는 방법을 말한다. K=1로 된다면 자신과 가장 가까이 위치하는 이웃의 클래스 값으로 자신의 클래스가 결정되게 된다. 실험 과정에서는 $l_2 - norm$을 이용하여 계산을 수행한다.

1. Yale 데이터에 대한 인식률

Yale데이터는 총 165장의 얼굴 영상으로 구성되어 있으며, 15명의 사람에 대해 각각 11장씩 구성되어 있다. 각각의 영상의 크기는 100×80 픽셀의 크기로 구성되어 있다. 표 2와 표 3은 Yale데이터에 극좌표계 변환을 적용한 후 LDA와 NLDA의 인식률을 나타낸 것으로, 특징의 개수는 LDA와 NLDA 모두 14개로 하여 얻은 결과이다. 표 2는 원영상과 극좌표 변환한 영상의 인식률을 나타낸 것이며, 표 3은 히스토그램 평활화를 한 영상과 이 영상을 극좌표 변환한 영상의 인식률을 나타낸 것이다.

이 때 분류는 1-최근접 이웃 알고리즘을 이용하였고, $l_2 - norm$을 이용하여 계산을 수행한다. 각 인식률은 총 15명의 사람의 10장씩의 영상을 학습 영상으로 사용하고, 나머지 1장의 영상을 테스트 영상으로 사용하여 인식률을 확인하고 총 11번의 실험을 통해 얻은 인식률 값을 평균을 하여 나타낸 것으로, 교차검증 방법을 통해 나온 결과이다.

표 2에서 볼 수 있듯이 LDA의 경우 히스토그램 평활화를 하지 않은 영상의 경우 원영상의 경우 인식률이 78.78%임에 비해, 반지름을 변화시키며 측정한 극 좌표계 변환한 영상의 인식률은 84.24% 이상으로 상승하는 것을 볼 수 있다. NLDA의 경우 히스토그램 평활화를 하지 않은 영상의 경우 표 2에서 볼 수 있듯이 원영상에 대한 인식률은 87.88%이며, 반지름이 30일 경우를 제외하고, 반지름을 35이상으로 했을 경우 인식률이 더 상승하는 것을 볼 수 있다.

표 3에서 볼 수 있듯이 LDA의 경우 히스토그램 평활화를 한 영상의 경우 원영상의 경우 인식률이 92.12%

표 2. Yale 데이터에 대한 원영상과 극좌표 변환을 수행한 결과의 인식률

Table 2. The recognition rate for original image and image using polar coordinate transform for Yale database.

no histeq		LDA	NLDA
original		78.78	87.88
polar	r = 30	84.24	87.87
	r = 35	91.51	90.30
	r = 40	92.72	94.54
	r = 45	92.72	93.33
	r = 50	90.90	92.12

표 3. 히스토그램 평활화를 수행한 Yale 데이터에 대한 원영상과 극좌표 변환을 수행한 결과의 인식률

Table 3. The recognition rate for original image and image using polar coordinate transform for histogram equalized Yale database.

histeq		LDA	NLDA
original		92.12	93.33
polar	r = 30	89.09	92.12
	r = 35	92.73	98.79
	r = 40	93.94	97.58
	r = 45	95.76	98.18
	r = 50	96.97	99.39

임에 비해, 반지름을 변화시키며 측정한 극 좌표계 변환한 영상의 인식률은 반지름이 30일 경우를 제외하고, 92.73% 이상으로 상승하는 것을 볼 수 있다. NLDA의 경우 히스토그램 평활화를 한 영상의 경우 표 3에서 볼 수 있듯이 원영상보다 극 좌표계 변환한 영상의 인식률이 93.33%이며, 반지름이 30일 경우를 제외하고, 반지름을 35이상으로 했을 경우 인식률이 더 상승하는 것을 볼 수 있다.

표 2에서는 반지름이 40일 때 인식률이 가장 높은데 이는 Yale 데이터에 얼굴 외곽 주변에 그림자가 진 영상이 포함되어 있기 때문에, 이러한 그림자 부분이 포함되지 않음과 동시에 얼굴의 중요 특징을 모두 포함하고 있는 반지름을 40픽셀로 한 경우가 인식률이 가장 높게 된다. 표 3의 경우 히스토그램 평활화를 하였기 때문에 LDA나 NLDA를 통한 얼굴인식을 할 경우 그림자에 의한 영향이 적어지게 되므로, 반지름을 증가시킴으로써, 얼굴 영상의 특징 정보들이 많이 포함될수록 인식률이 증가되는 경향을 나타냄을 볼 수 있다.

히스토그램 평활화를 수행하기 전의 영상에 대한 실험의 경우에는 반지름이 너무 작게 되면 얼굴 영상의 중요한 특징 정보를 포함하기 어려우며, 너무 크게 되면 표 2의 결과에서 볼 수 있듯이 그림자나 머리카락 등 얼굴 인식에 방해가 되는 요소가 포함될 여지가 있기 때문에 반지름을 알맞게 설정하는 작업이 중요하다.

2. FRGC 데이터에 대한 인식률

FRGC 데이터의 경우 총 40명의 영상에 학습 데이터와 테스트 데이터로 400장을 사용하였다. 각 영상의 크

기는 80x60 픽셀의 크기로 구성되었다. 표 4는 FRGC 데이터에 극 좌표계 변환을 적용한 후 LDA와 NLDA의 인식률을 나타낸 것이다. 이 때 분류는 1-최근접 이웃 알고리즘을 이용하였고, $l_2 - norm$을 이용하여 계산을 수행한다. 표 4와 표 5의 각 인식률은 총 40명의 사람의 10장씩의 영상을 학습 영상으로 사용하고, 10장씩의 영상을 테스트 영상으로 사용하여, 학습영상으로 400장, 테스트 영상으로 400장을 사용하여 인식률을 얻어낸 결과이다. 표 4와 표 5의 LDA와 NLDA의 인식률은 특징을 39개를 사용하여 나타낸 결과이며, 10번의 실험을 통해 서로 다른 학습영상과 테스트 영상에 대해 얻은 인식률을 평균을 하여 나타낸 결과이다.

표 4에서 볼 수 있듯이 LDA의 경우 히스토그램 평활화를 하지 않은 원영상을 사용하는 경우 인식률이 90.71%임에 비해, 반지름을 30픽셀 이상으로 했을 경우 91.48% 이상의 인식률을 보인다. NLDA의 경우 히스토그램 평활화를 하지 않은 영상의 경우 표 4에서 볼 수 있듯이 원영상의 경우 인식률이 94.77%이며, 반지름을 35픽셀이상으로 했을 경우 극좌표 변환의 인식률이 더 높은 것을 볼 수 있다.

표 4. FRGC 데이터에 대한 원영상과 극좌표 변환을 수행한 결과의 인식률

Table 4. The recognition rate for original image and image using polar coordinate transform for FRGC database.

no histeq		LDA	NLDA
original		90.71	94.77
polar	r = 25	89.62	92.30
	r = 30	91.48	94.22
	r = 35	92.40	95.27
	r = 40	92.53	95.53
	r = 45	92.05	95.53

표 5. 히스토그램 평활화를 수행한 FRGC 데이터에 대한 원영상과 극좌표 변환을 수행한 결과의 인식률

Table 5. The recognition rate for original image and image using polar coordinate transform for histogram equalized FRGC database.

histeq		LDA	NLDA
original		94.17	96.92
polar	r = 25	93.4	95.08
	r = 30	94.95	96.48
	r = 35	95.45	97.05
	r = 40	95.43	97.33
	r = 45	95.33	97.30

표 5에서 볼 수 있듯이 LDA의 경우 히스토그램 평활화를 한 영상의 경우 원영상의 경우 인식률이 94.17%임에 비해, 반지름을 변화시키며 측정한 극 좌표계 변환한 영상의 인식률은 반지름이 30픽셀 이상일 경우, 94.95% 이상으로 상승하는 것을 볼 수 있다. NLDA의 경우 히스토그램 평활화를 한 영상의 경우 표 5에서 볼 수 있듯이 LDA의 실험 수행 결과보다도 더 인식률이 높은 것을 볼 수 있다. 표 4와 표 5의 결과에서 볼 수 있듯이 반지름을 40픽셀로 한 경우 인식률이 가장 높은 것을 볼 수 있다.

얼굴 영상의 중요한 특징 정보를 모두 포함하며, 얼굴 외각 부분의 얼굴 인식에 방해가 되는 요소는 포함되지 않도록 반지름을 설정하는 작업이 중요하게 된다.

V. 결 론

극좌표계 변환 방법은 얼굴 영상은 얼굴의 중심부에 가까운 영역에 얼굴을 분류하기 위한 주요 부위가 밀집되어 있다는 사실에 기반 한다. 이는 원 얼굴 영상과 머리카락 같은 얼굴의 외곽부분은 제외하고, 눈, 코, 입과 같은 얼굴 부분만을 포함하게 잘라내어 이를 원 얼굴 영상과의 인식률을 비교해 본 결과를 통해서 이러한 점을 인식할 수 있다. 즉 원영상에 대해 LDA, NLDA를 이용해 얻은 얼굴 인식률과 머리카락 같은 얼굴의 외곽부분을 제거하고 눈, 코, 입 주변 부위만 잘라내어 인식률을 비교해 보면 후자의 경우가 인식률이 더 높은 것을 알 수 있다. 극 좌표계 변환은 얼굴의 중심부분의 한 점을 극으로 삼아 일정길이 만큼 얼굴 영상을 샘플링하여 새로운 얼굴 영상을 제작하고 이를 바탕으로 기존의 특징 추출 방법을 이용하여 얼굴인식의 성능 향상을 도모할 수 있다. 앞으로 조명에 덜 민감하며, 얼굴이 돌아가 있는 경우에도 인식률을 높일 수 있는 방법에 대한 연구가 필요할 것이다. 또한, 코끝의 위치를 자동으로 찾을 수 있는 방법에 대한 연구도 이루어져야 할 것이다.

참 고 문 헌

[1] Martinez and A. Kak, "PCA versus LDA," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, 2001.

[2] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, March 1997.

[3] M. S. Bartlett, J. R. Movellan and T. J. Sejnowski, "Face recognition by independent component analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 6, March 2000.

[4] S. Matschitsch, H. Stogner, M. Tschinder and A. Uhl, "Rotation-invariant iris recognition: Boosting 1D Spatial-Domain Signatures," ICINCO 2008: Proceedings of the 5th International Conference on Informatics in Control, Automation and Robotics, vol. SPSMC, pp. 232-235, 2008.

[5] H. Zhang, D. Wang and K. Song, "Object recognition using polar-exponential grid technique," Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the 3rd International Symposium, vol. 2, pp. 816-819, Sept 2003.

[6] J. Yang, Y. Yu, and W. Kunz, "An efficient LDA algorithm for face recognition," Proceedings of the Sixth International Conference on Control, Automation, Robotics and Vision, vol. 33, no. 10, 2000.

[7] 한학용, 패턴인식개론, 한빛미디어, 2005.

[8] Jieping Ye and Tao Xiong, "Null space versus orthogonal linear discriminant analysis," Preceedings of the 23rd international conference on Machine learning, pp.1073-1080, June 2006.

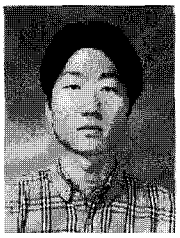
[9] Z. Xiaoxun and J. Yunde, "Symmetrical null space LDA for face and ear recognition," Neurocomputing, vol. 70, pp. 842-848, January 2007.

[10] Ganhua Li, Xuanping Cai and Xianshuai Li, "An Efficient Face Normalization Algorithm Based on Eyes Detection," Proceedings of the 2006 IEEE/RSJ, pp. 3843-3848, Oct 2006.

[11] 홍은혜, 고병철, 변혜란, "PCA와 LDA를 이용한 실시간 얼굴 검출 및 검증기법," 정보과학회논문지 : 소프트웨어 및 응용, 제31권 (2), pp. 213-223, Feb 2004.

[12] M. Turk and A. Pentland, "Face recognition using eigenfaces," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.

저 자 소 개



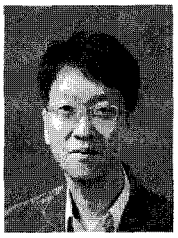
오 재 현(학생회원)

2006년 아주대학교 전자공학부 학사 졸업.

2008년 아주대학교 전자공학과 석사 졸업.

2008년~현재 아주대학교 전자공학과 박사과정.

<주관심분야 : 컴퓨터 비전, 신호처리>



곽 노 준(정회원)

1997년 서울대학교 전기공학부 졸업.

1999년 서울대학교 전기 컴퓨터 공학부석사 졸업.

2003년 서울대학교 전기 컴퓨터 공학부박사 졸업.

2003년 3월~2006년 8월 삼성전자 정보통신총괄 통신연구소 표준연구팀 책임연구원.

2006년 9월~2007년 2월 서울대학교 전기컴퓨터 공학부 BK조교수.

2007년 3월~현재 아주대학교 전자공학부 조교수.

<주관심분야 : 컴퓨터 비전, 영상 처리>