

극좌표계에서 회전에 강인한 LDAR을 이용한 얼굴 인식

오재현 광노준

아주대학교 전자공학과

hyunsa@ajou.ac.kr

Rotation invariant face recognition in a polar coordinate system using LDAR

Oh, Jae-Hyun Kwak, Nojun

Division of Electrical & Computer Engineering, Ajou University

요약

본 논문은 기존 평행좌표를 이용하는 얼굴 영상 대신 극좌표계 변환을 이용한 얼굴 영상을 이용하여 회전에 강인한 얼굴 인식 방법을 제안한다. 극좌표계 변환 방법은 얼굴의 중심부분의 한 점을 극으로 삼아 이 점을 기준으로 360도 각 방향으로 일정 길이만큼 얼굴 영상을 샘플링 하여 새로운 얼굴 영상을 제작하는 방법이다. 이 극좌표계 변환 방법을 이용해 재구성된 영상에 대해 회귀(regression)문제 해결을 위해 변형된 LDA인 LDAR(LDA for regression)을 이용하여 얼굴의 중심부분의 한 점인 극을 중심으로 임의의 각도로 회전된 영상의 회전 정도를 추정하여 이를 정규화 시키는 방법을 통해 얼굴 인식의 인식률을 향상시키고자 한다. LDAR은 LDA의 기본개념인 각 클래스 간 떨어진 정도를 최대화하는 것이 목적으로 클래스간 분산과 클래스내 분산의 비율을 최대화 하는 방법을 응용하여 이를 회귀문제에 적용할 수 있게 변형을 한 것이다. 즉, LDAR은 목표값(target)의 차이가 큰 샘플들과 목표값의 차이가 작은 샘플들 간의 거리의 비율을 최대화 하는 것을 목적으로 하게 된다. 제안된 방법을 Yale데이터에 적용하여 임의의 각도로 회전시킨 영상에 대해 회전 각도를 정확히 찾아내는 것을 확인할 수 있었다.

1. 서론

얼굴인식 분야에서 패턴은 입력 변수의 집합으로 표현될 수 있으며, 일반적으로 얼굴인식에서는 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 [1] 같은 문제를 해결하는 효과를 얻을 수 있다. 얼굴 인식 분야에서 사용되는 차원 감소 방법에는 아공간(subspace) 방법들이 많이 사용되며 이들 중 대표적인 것으로 LDA (Linear Discriminant Analysis) [2], NLDA (null space based LDA) [3] 등의 방법이 있다. 이 중 LDA는 각 클래스 간 떨어진 정도를 최대화하는 것이 목적으로 특징 벡터의 차원을 감소시키는 기법 중의 하나이며, 클래스간 분산과 클래스내 분산의 비율을 최대화 하는 방법을 이용하여 데이터에 대한 특징 벡터의 차원을 축소한다.

본 논문에서 적용하는 알고리즘인 LDAR은 LDA의 기본개념인 각 클래스 간 떨어진 정도를 최대화하는 것이 목적으로 클래스간 분산과 클래스내 분산의 비율을 최대화 하는 방법을 응용하여 이를 회귀문제에 적용할 수 있게 변형을 한 것이다[4].

본 논문에서는 기존의 연구 결과[5][6]인 극좌표계 변환 방법을 이용하여 조명을 보정하는 방법으로 구성된 극좌표계 영상을 사용하여 실험을 수행한다. 극좌표계 변환 방법은 얼굴을 분류하기 위한 중요한 특징들의 대부분은 얼굴의 중심부에 위치한다는 사실에 기반하며, 극좌표계를 이용하면 얼굴을 분류하는데 중요한 특징들이 평행좌표계를 이용한 원영상을 그대로 사용할 때보다 많이 고려되어, 아공간 방법을 사용한 얼굴인식을 수행할 때 인식률의 향상을 도모할 수 있다.

본 논문에서는 먼저, 극좌표계 변환 방법과 회전에 강인한 LDAR을 이용한 얼굴인식에 대해 설명하며, LDAR의 기본개념을 서술한다. 그리고 실험 결과 및 분석을 통해 LDAR을 통해 임의로 회전된 영상에 대한 회전각 추정 정도를 판단하며, 마지막으로 결론을 도출한다.

2. 극좌표계 변환 방법과 회전에 강인한 LDAR을 이용한 얼굴인식

가. 극좌표계 변환 방법

얼굴 인식 분야에서 통상적으로 사용하는 기존의 평행 좌표계(원 영상)와 달리 얼굴을 분류하는데 중요한 역할을 하는 눈, 코, 입 등 얼굴 인식에서 중요하게 사용되는 특징이 밀집한 부분을 보다 세밀하게 표현할 수 있는 극좌표계 변환방법을 이용한다. 얼굴 영상을 분류하기 위해서는 분류를 하기 위해 각 얼굴의 중요한 정보를 담고 있는 특징들을 뽑아내는 작업이 중요하다. LDA나 NLDA를 통한 얼굴 인식에서 일반적인 얼굴 영상을 그대로 사용하게 되는 경우 영상의 모든 픽셀을 동일한 중요도를 갖고 차원벡터를 줄이게 된다. 이 때 머리 스타일의 변화가 있거나 모자를 쓴 경우와 같이 얼굴의 중심부분 외의 다른 부분이 바뀔 때에도 인식률이 크게 저하될 수 있는 문제점이 발생한다. 이러한 문제점을 해결하기 위해 본 연구에서는 극좌표계를 이용한 얼굴인식 방법을 제시한다.

극좌표계 변환은 얼굴의 중심부를 기준점으로 일정한 반지름의 원을 그리며 360도 방향으로 각각 원의 중심에서부터 반지름의 끝 부분까지 샘플링을 하여 얼굴 데이터를 재구성한다. 눈, 코, 입 등 얼굴의 주요 특징들이 얼굴의 중심부분에 집중적으로 모여 있으므로, 만일 극좌표계와 평행좌표계를 이용하는 영상이 동일한 수의 픽셀로 이루어

저 있다면 극좌표계에서 얼굴의 주요 특징점들이 평행좌표계보다 더 많이 샘플링되는 효과가 있다. 따라서 극좌표계로 변환한 영상에 LDA 나 NLDA와 같은 기존의 아공간 방법을 사용해 얼굴인식을 수행한다면 인식률의 향상을 도모할 수 있다. 그림 1(위)은 Yale데이터의 원 영상을 나타낸 것이고, 그림1(아래)은 기존의 연구 결과인 [5]를 이용하여 원 영상에 대해 반지름을 40픽셀로 극좌표계 변환을 통해 얻은 영상을 대칭성을 이용하여 조명 보정을 한 영상이다. 본 논문에서는 대칭성을 이용해 조명 보정을 한 극좌표계 변환을 통해 얻은 영상을 이용하여 조명의 영향을 줄인 영상에 대해 임의의 각도로 회전을 시킨 후 회전 정도를 추정한다.

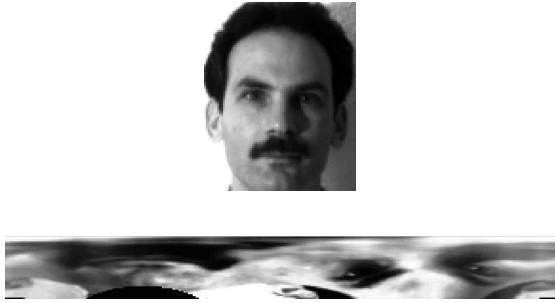


그림 1. Yale데이터 원 영상(위)과 원 영상에 대해 반지름을 40픽셀로 극좌표계 변환을 통해 얻은 영상을 대칭성을 이용해 조명 보정을 한 영상(아래)

나. 회전에 강인한 LDAR을 이용한 얼굴인식

LDA는 얼굴 인식 분야에서 사용되는 차원 감소 방법 중 대표적으로 많이 사용되는 방법이다. LDA는 각 클래스 간 떨어진 정도를 최대화하는 것이 목적으로 특징 벡터의 차원을 감소시키는 방법 중 하나이며, 클래스간 분산과 클래스내 분산의 비율을 최대화 하는 방법을 이용해 데이터에 대한 특징 벡터의 차원을 축소하게 된다. 이 방법을 회귀 문제에 적용하기 위해 변형된 LDA의 형태인 새로운 특징 추출을 하고자 하는 것이 LDAR의 목적이다[4].

LDAR에서는 먼저 입력과 목적값의 쌍을 $\{x_i, y_i\}_{i=1}^n$ 로 놓는다. 회귀 문제를 풀기 위해서는, 목적값 y 에 대해 많은 정보가 포함되어 있는, x 의 선형 변환 형태의 특징들($=w_i^T x$) 찾는 것을 목적으로 한다. 이러한 목적에 맞게 Fisher's criterion을 회귀 문제에 맞게 변형한다. 먼저 분류문제를 풀기위한 LDA를 회귀문제를 풀기 위한 형태로 변형하기 위해 이산 값인 y 를 클래스로 생각한다. 그리고 기존 LDA에서의 클래스간 분산과 클래스내 분산을 각각 식(1)과 (2)와 같이 변형하여 정의한다.

$$S'_b = \frac{1}{n_b} \sum_{(i,j) \in A_b} f(y_i - y_j)(x_i - x_j)(x_i - x_j)^T \quad (1)$$

$$S'_w = \frac{1}{n_w} \sum_{(i,j) \in A_w} f(y_i - y_j)(x_i - x_j)(x_i - x_j)^T \quad (2)$$

여기서 A_b 와 A_w 는 각각 멀리 떨어져 있는 샘플들 (x_i, x_j) 의 인덱스 쌍과 가까운 샘플들 (x_i, x_j) 의 인덱스 쌍의 집합을 나타낸다. 또

한 식(1)과 식(2)에서 $f(\cdot)$ 는 가중치 함수로서 양수 값을 갖게 된다. 목표값이 이산 값으로 표현되는 분류문제와 달리 회귀 문제에서는 목표값이 연속적인 분포를 나타내므로 클래스내 분산과 클래스간 분산을 정의하는 것이 어렵다. 따라서 회귀문제에 적용하기 위해 분류문제에서 y 값을 연속적인 분포를 갖는 변수로 보고, soft class라는 개념을 도입한다. 식(1)과 식(2)의 A_b 와 A_w 를 soft class의 개념을 도입하여 다음과 같이 정의한다.

$$A_b = \{(i,j) \mid |y_i - y_j| \geq \tau, i < j\} \quad (3)$$

$$A_w = \{(i,j) \mid |y_i - y_j| < \tau, i < j\} \quad (4)$$

식(3)과 식(4)에서 τ 는 A_b 와 A_w 를 구분하는 문턱값이며, 식(1)과 식(2)에서 $n_b = |A_b|$, $n_w = |A_w|$ 이 된다. 즉, 목표값의 차이가 문턱값보다 작은 샘플들의 경우 같은 클래스에 속하고, 그렇지 않은 경우 다른 클래스에 속하게 된다. 이렇게 변형된 A_b 와 A_w 를 이용하여 Fisher's criterion을 구해보면, 식(5)와 같이 얻을 수 있다.

$$W = \operatorname{argmax} \frac{|W^T S'_b W|}{|W^T S'_w W|} \quad (5)$$

식(5)의 Fisher's criterion을 최대화 하는 것은 식(6)의 고유값 문제를 풀게 됨으로써 동일한 결과를 얻을 수 있다.

$$S_w^{-1} S_b w_k = \lambda_k w_k \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \quad (6)$$

LDAR에서 문턱값 τ 를 결정하는 것이 중요하게 되는데, $\tau = \alpha \sigma_y$ 와 같이 목표값 y 의 표준편차 σ_y 에 상수 α 를 곱한 값으로 표현 가능하게 된다. 통상적인 α 의 범위는 0.1과 1.0 사이로 정할 수 있으며, 본 논문에서는 α 값을 0.3으로 실험을 수행하여 결과를 얻었다. 또한 가중치 함수 $f(\cdot)$ 로는 $\sqrt{|x_i - y_j|}$ 를 사용하였다.

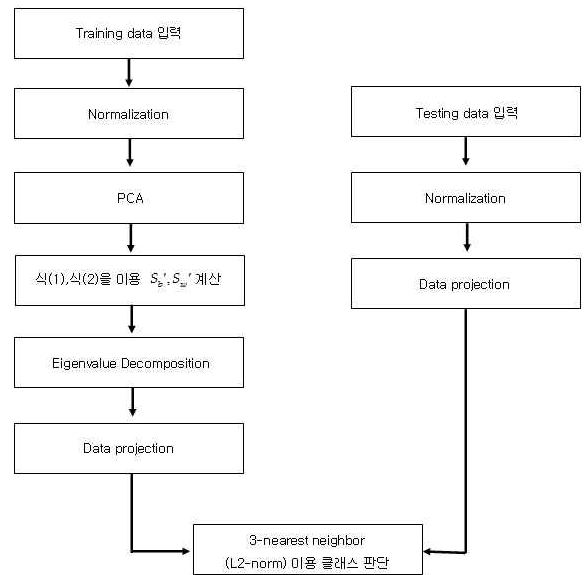


그림 2. LDAR을 이용한 얼굴인식의 흐름도

본 논문에서는 LDAR을 임의의 각도로 얼굴이 회전되어 있는 극좌표계로 변환된 영상을 이용하여 회전된 각도를 추정한다. 본 실험에서 목표값은 회전각도이며, 학습 영상과 테스트 영상으로 나누어 LDAR을 수행하여 정확도를 판단한다. 그림 2는 LDAR을 이용한 얼굴 인식의 흐름도를 나타낸 것이다. 먼저 학습 데이터를 정규화 한 후에 PCA를 이용하여 차원을 감소시킨 후 식(1)과 식(2)을 이용하여 클래스간 분산과 클래스내 분산을 계산하고 고유값 문제를 풀어 학습 데이터를 사영을 시킨다. 또한 테스트 영상에 대해서도 정규화를 시킨 후 학습 과정에서 LDAR을 통해 얻어냈던 최적화된 행렬에 사영을 시킨다. 테스트 영상에 대해서 먼저 구해놓은 학습 영상들과의 l_2 -norm을 계산하고 3-최근접 이웃 알고리즘(3-nearest neighbor)을 이용하여 얼굴 영상의 회전 각도를 판단한다.

3. 실험 결과 및 분석

Yale데이터는 총 165장의 얼굴 영상으로 구성되어 있으며, 15명의 사람에 대해 각각 11장씩 구성되어 있다. 각각의 영상의 크기는 100×80 픽셀의 크기로 구성되어 있다. 각 영상을 극좌표계 변환을 통하여 40×360 크기의 극좌표계 변환된 영상을 만든다. 그림 3은 극좌표계 변환된 영상을 -30°부터 +30°의 범위 내에서 임의의 각도로 회전시킨 영상들 중 일부를 나타낸 것이다. 이렇게 임의의 각도로 회전된 영상을 총 165장의 영상들에 대해 한 장에 10번씩 임의의 각도로 회전을 시킨 영상을 만들어 총 1650장의 영상을 이용하여 이 중 770장은 학습 영상으로 사용하고, 880장은 테스트 영상으로 사용하여 회전각도 추정의 정확도를 확인한다. 이 때 회전각도 추정을 위해 3-최근접 이웃 알고리즘을 이용하였고, l_2 -norm을 이용하여 계산을 수행한다.

표 1은 -20°부터 +20°의 범위, -30°부터 +30°의 범위, -40°부터 +40°의 범위에서 임의의 각도로 회전시키고 각각 학습 영상과 테스트 영상에 대한 추정된 회전각을 실제 회전각과 비교하여 rms error를 계산한 것이다. 각도의 변화 정도가 클수록 학습 영상의 rms error는 증가하는 것을 볼 수 있고, 테스트 영상의 경우 -30°부터 +30°의 범위 내에서 각도를 임의로 변화시켰을 때 rms error가 1.421°로 가장 작게 나타나는 것을 볼 수 있다. 임의의 테스트 영상이 들어오게 될 경우 2.458° 범위 내의 오차 이내로 회전각을 추정하는 것을 볼 수 있다.

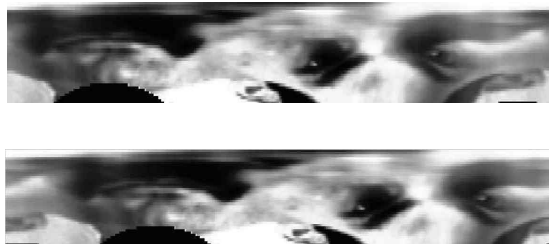


그림 3. 대칭성 이용해 조명 보정을 한 극좌표계 변환의 결과를 시계방향으로 22° 회전한 영상(위)과 반시계방향으로 9° 회전한 영상(아래)

Rotation(°)	± 20 °	± 30 °	± 40 °
tr rms error(°)	1.358 e-15	0.030	0.057
test rms error(°)	2.458	1.421	1.659

표 1. 임의의 각도로 회전한 극좌표계 영상에 대해 LDAR을 이용하여 추정된 rotation 정도의 rms error값

4. 결론

극좌표계 변환 방법은 얼굴을 분류하기 위한 중요한 특징들의 대부분은 얼굴의 중심부의 주변에 위치한다는 사실에 기반 한다. 극좌표계 변환 방법은 얼굴의 중심부분의 한 점을 극으로 삼아 이 점을 기준으로 360도 각 방향으로 일정 길이만큼 얼굴 영상을 샘플링 하여 새로운 얼굴 영상을 제작하는 방법이다. 이 극좌표계 변환 방법을 이용해 재구성된 영상에 대해 회귀문제 해결을 위해 변형된 LDA인 LDAR을 이용하여 얼굴의 중심부분의 한 점인 극을 중심으로 임의의 각도로 회전된 영상의 회전 정도를 추정하여 이를 정규화 시키는 방법을 통해 얼굴 인식의 인식률을 높일 수 있다. 제안된 방법을 Yale데이터에 적용하여 임의의 각도로 회전시킨 영상에 대해 회전 각도를 추정할 수 있다.

참고 문헌

- [1] 한학용, 패턴인식개론, 한빛미디어, 2005.
- [2] Martinez and A. Kak, "PCA versus LDA," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, 2001.
- [3] Jieping Ye and Tao Xiong, "Null space versus orthogonal linear discriminant analysis," Proceedings of the 23rd international conference on Machine learning, pp. 1073-1080, 2006.
- [4] Nojun Kwak, Jung-Won Lee, "Feature Extraction based on Subspace Methods for Regression Problems," Neurocomputing, vol. 73, issues 10-12, pp. 1740-1751, 2010.
- [5] 오재현, 광노준, "극좌표계를 이용한 조명 보정과 얼굴 인식," 제 22회 영상처리 및 이해에 관한 워크샵 논문집, IPIU 2010 워크샵, 2010.
- [6] 오재현, 광노준, "극좌표계 변환에 기반한 얼굴 인식 방법," 전자공학회논문지, 제 47권, 제 1호, pp. 44-52, 2010.