

# 얼굴 인식을 위한 PCA, LDA, 및 정합기법의 비교연구

이동훈<sup>0</sup> 박영태

경희대학교 전자공학과

dhlee@kvision.khu.ac.kr, ytpark@khu.ac.kr

## A comparative study of PCA, LDA, and Matching Methods for Face Recognition

Donghun Lee<sup>0</sup> Youngtae Park

Dept. of Electronic Engineering, Kyunghee University

### 요 약

본 논문에서는 얼굴 영상의 변화를 보상할 수 있는 전처리 과정으로서 기하학적 특징에 기반한 순수 얼굴 영역 검출 기법을 도입하고 후처리 과정으로서 간단한 정합 기법을 사용한 얼굴 인식 기법을 제안한다. 실험결과를 보면 제안한 기법은 PCA와 LDA 기법에 비해 영상의 변화에 민감하지 않고 높은 인식률을 가진다.

### 1. 서 론

널리 알려진 얼굴 인식 기법은 크게 기하학적 특징에 기반한 템플릿 정합 기법, PCA (principal component analysis) 그리고 LDA (linear discriminant analysis) mapping 기법으로 나눌 수 있다. PCA 기법은 클래스 정보를 사용하지 않고 특징 공간에서의 데이터 분포의 특성을 유지하면서 영상의 차원을 줄이는 목적[1]이 있으므로 인식을 위한 분류의 목적으로 사용하기에는 적합하지 않다. LDA 방법은 클래스 정보를 사용하여 동일 클래스에 속한 데이터는 서로 모이고 다른 클래스에 속한 데이터는 되도록 분리 하는데 목적[2]이 있기 때문에 클래스간의 데이터들을 비교적 정확하게 분류한다. 대표적인 예로서는 FLD (Fisher linear discriminant)[3] 기법을 들 수 있다.

얼굴 영상의 크기가 작을수록 유사 얼굴 형상에 대한 변별력이 낮아 지며, 클 경우 얼굴 영역의 회전, 이동, 표정, 조명등의 변화에 따라 인식특성이 민감하여 eigenface나 fisherface의 적용이 어렵다. 또한 학습영상이 충분하지 않을 경우, 영상의 미세한 변화에 매우 민감하여 fisherface의 계산이 어려운 단점도 있다. 템플릿정합[4]도 영상의 변화에 민감한 특성을 가진다. 따라서 신뢰도가 높은 얼굴 인식 기법의 구현을 위해선 순수한 얼굴 영역의 위치를 정확히 추출하는 전처리 과정이 필수적으로 요구된다.

본 논문에서는 얼굴 인식 시스템을 구현함에 있어 기존의 기법을 보완할 수 있는 전처리 과정을 수행한 후의 mapping과 템플릿 정합에 의한 인식 특성을 비교하여 상용화가 가능한 궁극적인 얼굴 인식 기법을 제안한다. 또한 PCA와 LDA 기법을 결합함으로써 높은 인식률을 유지하면서 고차원의 영상 데이터에서의 계산상의 어려움을 극복할 수 있는 것도 보인다. 얼굴 영역의 다양한 변화를 자동으로 보상할 수 있는 전처리 과정으로서 기하학적 특징에 기반한 순수 얼굴 검출 기법[5]을 적용한다. 전처리 과정에서 산출된 순수 얼굴 영역으로부터 대표적인 템플릿 영상을 산출하는 기법을 제시하고, 간단한 템플릿 정합에 의해 PCA나 LDA 기법에 비해 월등한 성능을 보장할 수 있는 얼굴 인식 알고리즘을 구현할 수 있음을 보인다.

## 2. PCA, LDA, PCA+LDA 에 의한 mapping 기법

### 2.1 PCA

학습 영상의 집합을  $\Phi = [\phi_1, \phi_2, \dots, \phi_N]$  로 표기하고 각 영상으로부터 평균영상벡터  $\bar{\phi}$  를 뺀 영상의 집합을  $X = [x_1, x_2, \dots, x_N]$  로 표기하면 covariance matrix 는 다음과 같이 주어진다.

$$\Sigma_x = \sum_{i=1}^N x_i x_i' = XX' \quad (1)$$

$\Sigma_x$ 의 eigenvalue 와 eigenvector 행렬  $\Lambda$  와  $\Psi$  는 다음과 같은 관계를 갖는다.

$$(XX')\Psi = \Psi\Lambda \quad (2)$$

$\Sigma_x$ 를 효율적으로 계산하기 위하여 행렬  $X'X$ 의 eigenvalue 와 eigenvector matrix를  $V$ 와  $\Lambda'$ 라 하자.

$$(X'X)V = V\Lambda' \quad (3)$$

식(3)의 양변에  $X$ 를 곱하면 다음과 같다.

$$X(X'X)V = (XX')(XV) = (XV)\Lambda' \quad (4)$$

식(4)와 식(2)를 비교하면,  $\Lambda'$ 와  $XV$ 는  $\Sigma_x$ 의 eigenvalue 와 eigenvector matrix임을 알 수 있다. eigenvalue의 크기가 큰 순으로 배열했을 때  $p$ 개의 eigenvector로 구성된 matrix를 다음과 같이 구할 수 있다.

$$W_{PCA} = [Xv_1, Xv_2, \dots, Xv_p] \quad (5)$$

$m \cdot n$  크기의 영상벡터  $x_i$ 는  $p$  차원의 새로운 벡터  $y_i$ 로 다음 식에 의해 재구성된다.

$$y_i = W_{PCA}' x_i \quad (6)$$

### 2.2 LDA

FLD 기법에서 within-class matrix와 between-class matrix는 다음과 같이 계산된다.

$$S_w = \sum_{i=1}^N \Pr(C_i) E[(V_i - \mu_i)(V_i - \mu_i)'] \quad (7)$$

$$S_b = \sum_{i=1}^N \Pr(C_i) (\mu - \mu_i)(\mu - \mu_i)' \quad (8)$$

$\mu$ 는 전체 평균벡터를,  $\mu_i$ 는  $i$ 번째 클래스의 평균벡터를 나타낸다. FLD 기법에서는  $|W'S_bW|/|W'S_wW|$ 를 최대로 하는 프로제션 matrix  $W$ 를 사용한다. 이는 eigenvalue 문제로서 다음식에 의해 구할 수 있다.

$$S_b u_i = \lambda_i S_w u_i \quad (9)$$

eigenvalue의 크기가 큰 순으로 배열했을 때  $g$ 개의 eigenvector로 구성된 matrix를 다음과 같이 구한다.

$$W_{LDA} = [u_1, u_2, \dots, u_q] \quad (10)$$

$m \cdot n$  크기의 영상벡터  $x_i$  는  $q$  차원의 새로운 벡터  $z_i$  로 다음 식에 의해 재 구성된다.

$$z_i = W_{LDA}^T x_i \quad (11)$$

### 2.3 PCA+LDA

LDA 기법의 패턴인식 시스템의 분류 바이어스(classification bias)가 커지는 것과 학습 영상의 수가 적을 경우 식(9)의 eigenvalue와 eigenvector를 계산하는 알고리즘이 수렴하지 않는 문제점을 극복하기 위해 PCA와 LDA를 결합한 PCA+LDA 기법을 다음과 같이 적용할 수 있다.  $m \cdot n$  크기의 영상벡터  $x_i$  는  $q$  차원의 새로운 벡터  $\tau_i$  로 다음 식에 의해 재 구성된다.

$$\tau_i = W_{LDA}^T W_{PCA}^T x_i \quad (12)$$

### 2.4 분류기법 (classification)

클래스의 개수를  $N_c$  라 하면  $i$ 번째 특징벡터를  $F_i = (f_{i1}, f_{i2}, \dots, f_{i, N_c-1})^T$  로 표기하고 각각의 특징량을 정규화한 특징벡터는

$$\bar{f}_{ij} = (f_{ij} - \mu_j) / \sigma_j \quad (13)$$

로 구성된다.  $\mu_j$  와  $\sigma_j$  는 각각  $j$ 번째 특징량의 평균과 표준편차를 나타낸다. 클래스 별로 계산된 클러스터 센터를  $\{\eta_1, \eta_2, \dots, \eta_{N_c}\}$  로 표기하면 입력 얼굴영상은 다음과 같이 분류된다.

$$classification = \arg \min_k \| \bar{F} - \eta_k \| \quad (14)$$

### 3. 템플릿 정합기법

기하학적 특징에 기반한 순수 얼굴영역 검출기법[5]에 의해 회전과 이동 변화를 보정한 얼굴영역을 산출할 수 있으므로 간단한 템플릿 정합으로도 높은 정확도를 보장할 수 있는 얼굴 인식기법을 구현할 수 있다. 얼굴 표정의 다양함을 고려하여 템플릿 영상의 자동 생성기법을 다음과 같이 제안한다.  $k$ 번째 클래스에 속한 학습영상의 집합을  $I^k = \{I_1^k, I_2^k, \dots, I_n^k\}$ 로 표기한다.

#### 템플릿 생성기법:

Step 1:  $m=1$  ( $m$ 은 템플릿 개수를 의미한다.)

Step 2: 모든 화소의 좌표  $(x, y)$ 에 대해 화소값들  $\{I_1^k(x, y), I_2^k(x, y), \dots, I_n^k(x, y)\}$ 의 median 값을 산출하여  $m$ 번째 템플릿 영상  $T_m$ 을 생성한다.

Step 3:  $T_m$ 과 각 영상  $I^k$ 의 거리(distance)를 다음과 같이 계산한다.

$$\delta_i = \frac{1}{N_\Delta} \min_{\Delta} \sum_{(x,y)} |T_m(x, y) - I_i(x, y + \Delta)| \quad (15)$$

여기서  $N_\Delta$ 는  $T_m$ 과  $I_i$  영상에서의 중첩된 영역에 속한 화소수를 의미한다.

Step 4: 템플릿 영상의 생성에 참여한 모든 영상의  $\delta$  값이  $Q$ 보다 작다면 템플릿 생성과정을 마친다.

Step 5: 일부 영상의  $\delta$  값이  $Q$ 보다 작다면, 이들 영상의 다음 단계의 템플릿 생성과정에서  $m$ 을 증가시키고 Step 2의 과정을 되풀이한다.

Step 6:  $\delta$  값이  $Q$ 보다 작은 영상이 존재하지 않을 경우,  $\delta$  값이 가장 큰 영상으로  $T_m$ 을 대체한다.  $m$ 을 증가시키

고 Step 2의 과정을 되풀이한다.

식(15)에서는 수직 좌표축의 변화의 가능성을 고려하여  $\delta_i$ 의 계산과정에서 수직축으로 일정 구간만큼 이동시키면서 최소의  $\delta_i$  값을 계산한다. 또한  $Q$  값에 따라 생성된 템플릿 개수가 결정된다. 템플릿 영상의 생성 후, 입력 영상은 거리  $\delta$ 가 가장 작은 클래스로 분류된다.

### 4. 실험결과 및 고찰

#### 4.1 영상 데이터베이스와 전처리 과정

실험에서는 AR-face 데이터베이스[6]를 사용하였다. 기하학적 특징에 기반한 순수 얼굴영역 검출기법을 적용하여 각 영상으로부터 회전과 이동 변화가 보정된 순수 얼굴영역을 추출하였다. 그림 1에 그 예를 보인다. 조명 변화를 보정하기 위해 DC-notch filter를 사용하였다. 얼굴영역 검출 기법에 관한 자세한 내용은 [5]에서 다루고 있다.



그림 1. 검출된 순수 얼굴영역의 예

#### 4.2 저해상도 영상에서의 실험결과(16x16)

각 클래스마다 10개의 순수 얼굴영상을 추출하여 16x16 (혹은 30x30)크기로 정규화 하였다. 어려움을 측정하기 위해 leave-one-out 기법을 사용하였다. 7명의 인물 데이터에 대해 식(11)에 의해 FLD공간으로 mapping하고 첫 번째와 두 번째의 principal axis로 투영한 결과와 세 번째와 네 번째의 axis로 투영한 결과를 그림2의(a)와 (b)에 보인다. 학습 영상과 테스트 영상을 구별하기 위해 테스트 영상은 화살표로 표현하였다.

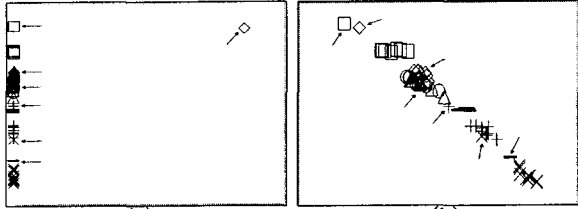


그림 2. FLD 공간에서의 데이터 분포 (data set 1)

테스트 영상의 위치는 소속된 클래스로부터 상당히 멀리 떨어져 분포하여 어려움이 필연적으로 높아진다. 이는 영상 데이터의 차원에 비해 너무 적은 수의 영상을 학습과정에 사용했기 때문이다. PCA에 의해 차원을 줄인 후 LDA기법을 적용할 경우 위에서 언급한 LDA의 단점을 극복할 수 있다. 그림 2의 (a),(b)와 같은 axis로 PCA+LDA 기법에 의해 투영된 예를 그림3의 (a)와 (b)에 보인다.

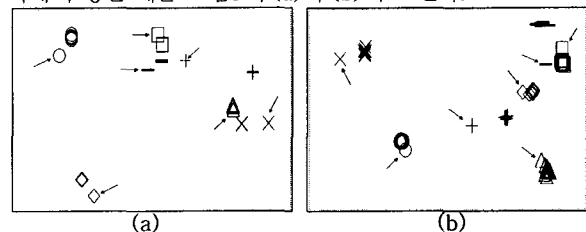


그림 3. PCA+LDA 공간에서의 데이터 분포 (data set 1)

테스트 영상의 위치는 소속 클래스의 위치에 가까이 분포함을 알 수 있다.

표 1에서 제시한 바와 같이 클래스의 수  $N_c$ 가 증가할수록 많은 영상이 학습과정에서 사용되므로 LDA 기법에서의

인식률이 향상된다. PCA+LDA 기법에서는  $N_c$ 에 무관하게 높은 인식률을 보인다. 따라서 LDA에 비해 우수한 인식 성능을 가진다. 템플릿 정합은 PCA+LDA 기법과 거의 유사한 인식 성능을 보인다. 이는 전처리 과정에서 얼굴 영상의 변화가 이미 정확히 보상되었기 때문이다.

표 1. LDA, PCA+LDA, 템플릿 정합의 Leave-One-Out 에러율 (16 x 16)

( $N_c$  = Number of classes,  $N_s$  = Total number of samples,  $N_f$  = Number of features)

Data Set	$N_c$	$N_s$	$N_f$	LDA	PCA+LDA	Template Matching
1	7	70	6	57(81.43%)	0(0%)	0(0%)
2	11	110	10	100(90.91%)	2(1.82%)	2(1.82%)
3	26	260	25	250(96.15%)	4(1.54%)	11(4.23%)
4	36	360	35	40(11.11%)	3(0.83%)	13(3.61%)
5	47	470	46	27(5.74%)	13(2.77%)	24(5.11%)
6	56	560	55	33(5.89%)	25(4.46%)	34(6.07%)
7	68	680	67	36(5.29%)	32(4.71%)	48(7.06%)
8	75	750	74	39(5.2%)	35(4.67%)	52(6.93%)

4.3 고해상도 영상에서의 실험결과(30x30)

PCA+LDA 기법과 템플릿 정합의 성능을 비교하기 위해 leave-one-out 방식과 hold-out 방식을 사용해 에러율을 측정하였고 그 결과를 표 2에 보인다. 두 가지 에러 측정 방식에서 모두 템플릿 정합의 성능이 월등함을 알 수 있다. leave-one-out 방식에서는 99%의 정인식률을 보이며 특히 적은 수의 학습 영상을 사용한 hold-out 방식에서도 클래스의 수에 무관하게 높은 인식률(96% 이상)을 얻을 수 있었다. 이와 같은 이유는 템플릿 정합의 해상도가 높아 유사한 클래스 영상에서의 변별력이 높기 때문이다. 이와 같은 템플릿 정합의 우월성은 ARENA[7] 알고리즘에서도 보고되었다.

표 2. PCA+LDA, 템플릿 정합의 에러율 (30x30 영상)

Data Set	Leave-One-Out Error Rate (%)		Hold-Out Error Rate (%)	
	PCA+LDA	Template Matching	PCA+LDA	Template Matching
1	3(4.29%)	0(0%)	2(5.71%)	0(0%)
2	3(2.73%)	0(0%)	6(10.91%)	0(0%)
3	4(1.54%)	0(0%)	12(9.23%)	3(2.31%)
4	4(1.11%)	1(0.28%)	9(5%)	3(1.67%)
5	10(2.13%)	2(0.43%)	11(4.68%)	4(1.7%)
6	19(3.39%)	3(0.54%)	23(8.21%)	8(2.86%)
7	31(4.56%)	4(0.59%)	29(8.53%)	12(3.53%)
8	36(4.8%)	4(0.53%)	40(10.67%)	14(3.73%)

4.4 영상의 변화에 따른 실험결과

회전, 이동변화에 따른 인식 알고리즘의 강건성을 유추하기 위해, 전처리 과정에서 회전, 이동변화의 보상기능을 제거하여 실험하였다. Data set 2에 대한 실험 결과를 표 3에 제시한다. 표 3에서 제시한 바와 같이 고 정확도의 얼굴인식 기법을 구현하기 위해서는 영상의 미세한 변화를 보상할 수 있는 전처리 과정의 선택이 필수적이다. 영상의 회전에 따른 효과를 명확히 보이기 위해 Data set 1에 속한 샘플 영상 1개를 선택하여 3°에서 15°까지 3° 간격으로 회전하고 학습된 PCA+LDA 공간에 투영하였다. 그림 4에 보인 바와 같이 투영된 점들은 소속된 클래스로부터 상당히 이탈되어 나타난다. 이는 mapping 기법에서의 인식 성능이 영상의 미세한 회전에 매우 민감함을 의미한다.

표 3. PCA+LDA와 템플릿 정합의 에러율 비교

Method	Leave-One-Out Error Rate (%)		Hold-Out Error Rate (%)	
	보상기능	보상기능 제거	보상기능	보상기능 제거
PCA+LDA	2(1.82%)	25(22.73%)	11(20%)	25(45.45%)
Template Matching	2(1.82%)	17(15.45%)	3(5.45%)	17(30.91%)

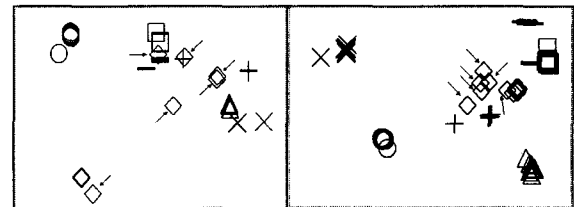


그림 4. 회전된 얼굴영상에 대한 PCA+LDA 공간에서의 데이터 분포. (a) 첫 번째와 두 번째의 principal axis로 투영한 결과 (b) 세 번째와 네 번째의 axis로 투영한 결과

5. 결론

얼굴 인식기법에서 널리 사용되어 온 PCA, LDA 등의 mapping 기법과 템플릿 정합 기법 모두 학습 영상의 수가 적을 경우, 얼굴 영역의 미세한 회전과 이동 변화에 민감하게 반응하여 높은 인식율을 유지하기 어려움을 실험을 통해 확인하였다. 특히 LDA 기법을 직접 영상에 적용하면 분류 바이어스(classification bias)가 커지고 학습 영상의 수가 적을 경우 fisherface의 계산이 불가능한 단점을 가진다. PCA+LDA는 이러한 문제점을 극복할 수 있는 기법이다. 기하학적 특징에 기반한 순수 얼굴영역 검출기법과 같이 회전 및 이동 변화를 보상할 수 있는 강건한 전처리 과정을 도입할 경우, 본 논문에서 제안한 간단한 템플릿 정합기법이 mapping에 의한 기법에 비해 월등한 성능을 가짐을 확인하였다. 또한 템플릿 정합기법은 클래스의 수가 증가할 경우 새로운 클래스의 템플릿 영상만을 추가함으로써 새로운 인식 시스템을 구성할 수 있는 장점을 가진다. 따라서 본 논문에서 제안한 기법은 높은 인식률을 보일 수 있는 궁극적인 얼굴 인식기법으로 활용될 수 있다.

6. 참고문헌

[1] M. A. Turk, A. P. Pentland, Face recognition using eigenfaces, Int. Conf. on Pattern Recognition, (1991) 586-591.  
 [2] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 19, (1997) 711-720.  
 [3] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, 2nd Edition, John Wiley & Sons, (2001).  
 [4] R. Brunelli, T. Poggio, Face recognition: Features versus templates, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 15, (1983) 1042-1052.  
 [5] 이대호, 박영태, 기하학적 특징에 기반한 순수 얼굴영역 검출기법, 한국정보과학회 논문집 (2002).  
 [6] A.M. Martinez, R. Benavente, "The AR Face Database," CVC Technical Report #24, June (1998).  
 [7] T. Sim, R. Sukthankar, M. D. Mullin, S. Baluja, High-performance memory-based recognition for visitor identification, Tech. Report JPRC-TR-1999-001-1, Just Research, (1999).