

## 얼굴 인식을 위한 지역적·전역적 특징 분석

이용진<sup>o</sup> 이경희 반성범  
 한국전자통신연구원 생체인식기술연구팀  
 {solarone<sup>o</sup>, uniromi, sbpan}@etri.re.kr

### Local and Global Feature Analysis for Face Recognition

Yongjin Lee<sup>o</sup> Kyunghee Lee Sungbum Pan  
 Biometrics Technology Research Team  
 Electronics and Telecommunications Research Institute

#### 요 약

Local Feature Analysis(LFA)는 눈, 코, 턱, 그리고 볼과 같은 얼굴의 지역적 특징을 잘 추출하는 것으로 알려져 있으나, 얼굴 인식에 이용하기에는 몇 가지 문제점이 있다. 본 논문에서는 LFA의 문제점을 개선하여 인식에 적합한 새로운 얼굴 특징 추출 방법을 제안한다. 제안 방법은 kernel 생성, 선택 그리고 중첩의 3 단계로 이루어진다. 첫번째 단계에서 얼굴의 지역적 특징을 검출할 수 있는 kernel을 생성하고, 두번째 단계에서 인식에 적합한 kernel을 선택한다. 마지막으로 선택된 kernel을 중첩시켜 적은 개수의 조밀한 형태의 kernel로 재표현한다. 실험을 통하여 제안 방법이 적은 개수의 특징을 이용하여 좋은 인식율을 보임을 확인하였다.

#### 1. 서 론

얼굴 인식에서 가장 기본적이면서 중요한 단계는 얼굴 특징 추출이다. 얼굴 특징 추출 방법은 얼굴의 전역적인 특징을 추출하는 방법과 지역적인 특징을 추출하는 두 가지 방법으로 나눌 수 있다.

전역적인 특징을 추출하는 대표적인 방법으로 Principal Component Analysis(PCA)를 이용한 eigenface가 있다[1]. Eigenface에서의 basis는 눈, 코, 턱 그리고 볼처럼 얼굴을 이루는 지역적인 구조보다 얼굴의 전체적인 형태를 기술한다. 그리고 조명과 얼굴 방향 같은 외부 변화나 전체적인 형태 변화에 따라 인식율에 많은 차이를 나타낸다.

반면 얼굴의 지역적 특징을 이용하는 방법은 환경 변화에 비교적 영향을 받지 않는 것으로 알려져 있으며, 대표적인 방법으로 Local Feature Analysis(LFA)가 있다[2]. LFA는 눈, 코 같은 지역적 구조의 특징을 잘 추출하지만, 인식에 사용하기에는 몇 가지 문제점이 있다 [3].

본 논문에서는 인식과 관련해 LFA의 문제점을 지적하고 개선하여, 새로운 얼굴 특징 추출 방법을 제안한다. 2, 3장에서 관련 배경 이론으로 LFA와 Fisher Score를 살펴보고, 4장에서 새로운 특징 추출 방법을 제안한다. 그리고 실험을 통해 제안 방법이 효과적임을 확인하고 마지막으로 결론을 내린다.

#### 2. Local Feature Analysis

LFA는 얼굴 영상으로부터 눈, 코, 턱과 같은 얼굴의 특징 구조에 잘 반응하는 kernel을 생성하여, eigenface의 eigenvector 처럼 특징 추출을 위한 basis로 사용한다. 그러나 (그림 1)에서 보듯이, 얼굴의 전반적인 형태를 기술하는 eigenface와는 달리, LFA의 kernel은 얼굴의 지역적인 형태를 기술한다.

Kernel은 얼굴 입력 영상의 covariance 행렬의 eigenvector 를 이용하여 생성된다. Eigenface와 마찬가지로 PCA에 기반하

지만, eigenvector 사이에 topology를 설정하여 입력 영상과 마찬가지로 위상 정보를 가지게 된다.

$n$ 개의  $\nu$  차원 얼굴 영상  $\{\phi_i\}_{i=1}^n$ 으로부터 얻은 covariance 행렬의  $N$ 개의 eigenvector  $\{\Psi_r\}_{r=1}^N$ 와 eigenvalue  $\{\lambda_r\}_{r=1}^N$ 가 있다고 할 때, kernel은 다음과 같이 정의 된다.

$$K(x, y) = \sum_{r=1}^N \Psi_r(x) \frac{1}{\sqrt{\lambda_r}} \Psi_r(y) \quad (1)$$

행렬  $K$ 는 대칭이기 때문에, 행 또는 열 벡터가 kernel을 나타낸다. 그리고  $t$ 번째 입력영상에 대한 kernel의 출력은 다음과 같다.

$$O_t(x) = \sum_{y=1}^{\nu} K(x, y) \phi_t(y) = \sum_{r=1}^N \frac{A_r(r)}{\sqrt{\lambda_r}} \Psi_r(x) \quad (2)$$

괄호 안의  $x, y$ 는 행렬과 벡터의 원소들의 인덱스를 나타내며,  $A_r(r)$ 은 다음과 같이 정의된다.

$$A_r(r) = \sum_{y=1}^{\nu} \Psi_r(y) \phi_t(y) \quad (3)$$

여기서 주목해야 할 것은 생성된 kernel의 개수가 입력 영상의 차원과 동일하다는 것이다. 따라서 출력의 차원 또한 입력 영상의 차원과 같다. 출력의 차원을 줄이기 위해 다음 식과 같이 복원 오차값이 가장 큰 출력의 kernel을 순차적으로 선택한다.

$$\arg \max_x \left\| \mathbf{O}_t(x) - \mathbf{O}_t^{rec}(x) \right\|^2 \quad (4)$$

Kernel이 위상 정보를 가지고 있기 때문에, kernel의 출력을 선택하는 것은 kernel을 선택하는 것과 동일하다.  $O_i(x)$ 에 대한 복원치  $O_i^{rec}$ 는 다음과 같이 정의된다.

$$O_i^{rec}(x) = \sum_{m=1}^{|M|} a_m O_i(x_m) \quad (5)$$

하지만, 계수  $a_m$ 을 구하기 위해서는 역행렬 연산이 필요하기 때문에, eigenvector의 개수( $N$ )보다 많은 kernel을 선택하기 어렵다는 문제점이 있다.

### 3. Fisher Score

Fisher Score는 하나의 class가 다른 나머지 class와 얼마나 잘 구분되는지를 나타내는 판별력 측정치로서, between-class scatter  $S_B$ 와 within-class scatter  $S_W$ 의 비율로 계산된다[4].  $c$  class에 대해 전체  $n$ 개의 data,  $x_1, \dots, x_n$ 가 있고, 각 class에는  $n_i$ 개의 data가 있다고 할 때, Fisher Score는 다음과 같다.

$$J = \frac{|S_B|}{|S_W|} \quad (6)$$

그리고  $S_B$ 와  $S_W$ 는 다음과 같이 정의된다.

$$S_B = \sum_{i=1}^c n_i (\mathbf{m} - \mathbf{m}_i)(\mathbf{m} - \mathbf{m}_i)^T \quad (7)$$

$$S_W = \sum_{i=1}^c \sum_{x \in c_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \quad (8)$$

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{x \in c_i} \mathbf{x}, \quad \mathbf{m} = \frac{1}{n} \sum_{i=1}^c n_i \mathbf{m}_i \quad (9)$$

### 4. 제안 방법

복원 오차값을 이용하여 LFA의 kernel을 선택할 경우, kernel을 생성할 때 사용된 eigenvector의 개수( $N$ )보다 많은 kernel을 선택할 수 없다. 또한 얼굴 영상 자체를 얼마나 잘 복원할 수 있는가를 기준으로 삼기 때문에, 인식에 적합한 kernel을 선택하는데 유용한 방법이라고 볼 수 없다.

(그림 2)의 (a)는 복원 오차값(식 4)을 이용하여 선택된 kernel의 위치와 선택된 순서를 얼굴 평균 영상 위에 표시한 것이다. 순서는 25번째 kernel까지만 표시하였다. Kernel을 생성하기 위해 100개의 eigenvector를 이용하였고, 생성된 kernel 중에서 100개를 선택하였다. 그림에서 보듯이 얼굴 윤곽과 옷깃에 해당되는 영역의 kernel들이 많이 선택되었다. 인식에 적합한 얼굴의 특징을 선택하기 보다는 영상 전체를 잘 기술할 수 있는 kernel들이 선택된 것을 알 수 있다. 그리고, (그림 2)의 (b)는 kernel의 Fisher Score(식 6)를 계산하여, 각각의 kernel 위치에 픽셀 값으로 표시한 것이다. 즉 흰색으로 표시된 영역이 높은 Fisher Score 값을 의미한다. Kernel의 Fisher Score는 kernel의 출력을 이용하여 계산하였다. 눈썹, 코, 볼, 턱처럼 인식에 적합한 영역의 kernel이 높은 값을 가지는 것을 알 수 있다. 그러나 인식에 적합한 kernel이 얼굴 전체에 넓게 퍼져 있

고, Fisher Score가 kernel 사이의 상관관계를 고려하지 않기 때문에 많은 수의 kernel이 선택된다. 하지만 여러 개의 kernel을 중첩시켜 조밀하게 표현함으로써 kernel의 개수를 줄일 수 있다.

Fisher Score를 이용하여 선택된 kernel 집합을  $I$ 라고 할 때, 중첩된 kernel  $G$ 를 선형 조합 가중치  $w_i$ 를 이용하여 다음과 같이 쓸 수 있다.

$$G = \sum_{i=1}^{|I|} w_i K(:, x_i) \quad (10)$$

여기서  $K(:, x_i)$ 는 행렬  $K$ 의  $x_i$ 번째 열 벡터, 즉  $x_i$ 번째 kernel을 의미한다.

따라서  $i$ 번째 입력 영상  $\phi_i$ 에 대한 최종 출력은 다음과 같다.

$$s_i = G^T \phi_i = \sum_{i=1}^{|I|} w_i O_i(x_i) \quad (11)$$

$G$ 를 만들기 위해서는 선형 조합 가중치  $w_i$ 를 결정해야 한다.  $w_i$ 를 결정하기 위해 여러 가지 방법이 사용될 수 있으나 문제를 간단히 하기 위해 본 논문에서는 entropy 최대화 방법을 이용한다 [5]. 즉, kernel을 중첩 시키면서 손실 되는 정보를 최소화하기 위하여  $G$ 의 출력  $s_i$ 의 entropy가 최대가 되도록  $w_i$ 를 결정한다.  $s_i$ 의 분포를 다음과 같은 정규 분포로 가정한다.

$$p(s) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{s^2}{2\sigma^2}\right] \quad (12)$$

그러면,  $p(s)$ 의 entropy를 최대로 하는 것은 variance  $\sigma^2$ 을 최대로 하는 것과 동일하고, 다음과 같이 정리할 수 있다.

$$\begin{aligned} \sigma^2 &= \frac{1}{n} \sum_{i=1}^n s_i^2 \\ &= \sum_{i=1}^{|I|} \sum_{j=1}^{|I|} w_i \mathbf{P}(x_i, x_j) w_j \\ &= \mathbf{w}^T \mathbf{P}(I, I) \mathbf{w} \end{aligned} \quad (13)$$

여기서  $\mathbf{P}(I, I)$ 는 kernel 집합  $I$ 의 출력의 covariance 행렬이고,  $\mathbf{w}^T = [w_1 \dots w_{|I|}]$ 이다.

$\mathbf{P}(I, I)$ 는 대칭 행렬이기 때문에  $\mathbf{w}^T \mathbf{w} = 1$ 의 제약 조건을 두면, (식 13)을 최대로 하는  $\mathbf{w}$ 를 eigen 문제로서 간단히 구할 수 있다. 즉, 집합  $I$ 의 출력에 대해 PCA를 적용하는 것과 같다. 그리고  $\mathbf{P}(I, I)$ 는 처음 kernel을 만들 때 사용한 얼굴 영상의 covariance 행렬의 eigenvector에 기반하기 때문에, 최대  $N$ 개의 중첩된 kernel만이 유용하다는 것을 알 수 있다.

(그림 1)은 eigenface, LFA의 kernel, 중첩된 kernel을 나타낸 것이다. Eigenface는 전체적인 얼굴 형태를 나타내는 반면, LFA의 kernel은 눈, 코와 같은 얼굴의 개별적인 구조를 강조하고 있다. 제안한 방법의 중첩된 kernel의 경우 두 방법의 중간적인 형태를 띤다. 즉 지역적이면서 전역적인 특징을 찾아내는 것

을 알 수 있다. 이는 눈, 코, 턱과 같은 지역 구조를 기술하는 여러 개의 kernel을 중첩시켜서 표현하였기 때문이다.

5. 실험 및 절차

Eigenface, LFA, Fisher Score를 이용하여 kernel을 선택한 경우, 그리고 중첩된 kernel을 이용한 4 가지 경우에 대해서 인식(identification) 실험을 하였다. 사용한 얼굴 영상 데이터는 ETRI 직원 55명으로 구성되어 있으며 영상의 크기는 64x64이다. 한 사람마다 20장의 얼굴 영상을 취득하여, 일인당 15 장은 특징 추출 및 등록에 사용하였고, 나머지 5장은 테스트에 사용하였다. 등록된 데이터와 테스트 데이터는 Euclidean 거리를 계산하여 비교하였다.

LFA의 kernel을 생성하기 위해 입력 영상으로부터 얻은 covariance 행렬의 eigenvector 100개를 사용하였다. 그리고 Fisher Score를 이용하여 1000개의 kernel을 선택하여 중첩된 kernel을 만들었다.

(표 1)의 실험결과는 rank에 따른 인식율을 나타낸 것이다. 첫번째 행의 숫자는 사용한 특징의 개수로, rank 1 인식율을 기준으로 최고의 성능을 나타낸 최소의 특징 개수이다. 제안한 중첩된 kernel을 이용한 방법이 가장 적은 특징 개수를 이용하여 가장 좋은 인식율을 보여주었다. 중첩 과정을 통해 Fisher Score로 선택한 많은 kernel이 성공적으로 줄어든 것을 알 수 있다.

6. 결론

본 논문에서는 LFA를 개선하여 새로운 얼굴 특징 방법을 제안하였다. 제안한 중첩된 kernel은 LFA kernel과 eigenface의 중간적인 형태를 지니고 있으며, 지역적이면서 전역적인 얼굴 특징을 추출한다. 그리고 비교 실험을 통해 제안한 방법을 이용하여 얻은 특징이 얼굴 인식에 효율적임을 확인하였다.

제안한 방법은 kernel 생성, 선택, 중첩의 3단계를 거친다. 각 단계에서 결정해야 되는 변수는 eigenvector의 개수, Fisher Score에 의한 kernel의 개수, 그리고 최종적인 특징 개수인 중첩된 kernel의 개수이다. 마지막 단계의 중첩된 kernel의 개수는 kernel 생성 단계에서 사용한 eigenvector의 개수에 제한되기 때문에 쉽게 결정할 수 있다. 그리고 Fisher Score를 이용하여 선택해야 되는 kernel의 개수의 경우, 특정 범위 안에서는 최종 인식율에 별다른 영향을 주지 않았다. 그리고 적절한 kernel의 개수는 (그림 2)의 (b)를 통해 쉽게 추측할 수 있다. 즉, 전체 단계에 걸쳐 주의 깊게 결정해야 하는 변수는 eigenface와 마찬가지로 얼굴 영상으로부터 얻은 covariance 행렬의 eigenvector의 개수이다. 따라서 제안 방법이 3단계로 이루어져 있지만 비교적 간단한 특징 추출 방법이라고 할 수 있다.

Rank	Eigenface (특징수: 67)	LFA (특징수: 97)	Fisher Score (특징수: 1000)	중첩된 kernel (특징수: 53)
1	86.18	69.45	94.18	94.18
2	90.91	76.36	97.45	96.36
3	95.27	77.82	97.82	97.82
4	96.36	81.45	97.82	97.82
5	97.09	85.09	98.18	98.18
10	98.18	90.55	98.55	98.55
20	98.91	92.00	100.00	100.00

표1. Rank에 따른 인식율 (%)

참고 문헌

[1] Matthew A. Turk, Alex P. Pentland, *Face Recognition using Eigenfaces*. Computer Vision and Pattern Recognition, 1991  
 [2] P. Penev, J. Atick, *Local Feature Analysis, A General Statistical Theory for Object Representation*. Network: Computation in Neural Systems 7(3), pp. 477-500, 1996  
 [3] M. Bartlett, *Face Image Analysis by Unsupervised Learning*, Kluwer Academic Publisher, Boston, 2001  
 [4] Richard O. Duda, Peter E. Hard, David G. Strok, *Pattern Classification*, John Wiley and Sons, 2001  
 [5] J. C. Principe, D. Xu, J. W. Fisher III, *Information-Theoretic Learning*, In S. Haykin, editor, Unsupervised Adaptive Filtering: Blind Source Separation, John Wiley and Sons, 2000

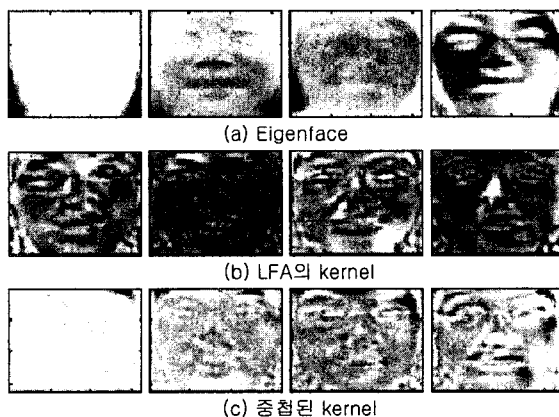


그림 1. 특징 추출 방법에 따른 basis 비교

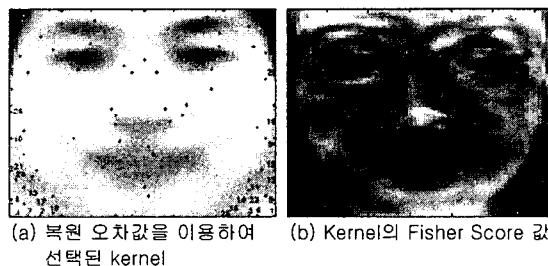


그림 2. 복원 오차값을 이용하여 선택된 kernel과 kernel의 Fisher Score 값