

# 효과적인 얼굴 영상 분류를 위한 컬러 고유 공간 분석

## (Color Eigen-Space Analysis for Efficient Face Image Classification)

김경수, 최형일

(Kyeong-Su Kim, Hyung-Il Choi)

승실대학교 컴퓨터학부

(School of Computing, Soongsil University)

### 요약

영상을 분류한다거나 물체를 인식하는 방법들은 대부분 흑백 영상에 대한 것이다. 그 이유는 기존의 분류 방법에 어떻게 컬러 정보를 결합시킬 것인가 하는 문제를 쉽게 해결하지 못하거나 처리하는데 훨씬 많은 시간이 소요되기 때문이다. 본 연구에서는 컬러 영상들을 분류하기 위하여 기존의 고유 벡터를 컬러 공간에 이용할 수 있는 방법을 제안하고, 이 고유 벡터를 이용하여 컬러 얼굴 영상에 대한 분류 실험을 통해 여러 가지 특징에 대한 고유 벡터를 영상 분류에 이용할 수 있음을 보였다.

### 1. 서론

기계로 인간의 얼굴을 분석하기 위한 연구는 과거 20년 전부터 여러 분야에서 이루어져 왔다. 또한, 최근 6년 동안에 이루어진 연구들은 얼굴의 구성 요소들을 분리하는 방법[8,9,10,12]이나 얼굴의 여러 가지 특징을 추출하는 분야[6,7,11]에서 많은 진전이 있었으며, Karhunen-Loeve(KL)변환에 기반한 통계적인 방법[1,2,3,4,5]이나 신경망을 이용한 얼굴 인식 분야[17]에도 많은 발전이 있었다.

또한, 이러한 얼굴 영상 인식에 사용되는 대상 영상도 흑백 영상에서 컬러 영상으로 전환되고 있으며, 정지 영상과 동영상에 이르기까지 어

떤 분야에서도 사용될 수 있는 방법이 고안되고 있다.

최근, 시각 학습 방법(visual learning method)은 얼굴 인식이나 표정 인식 분야에서 큰 성과를 보이고 있다[1,2,3,4,5]. KL 변환에 기반하는 이 방법은 인식 등에는 뛰어난 성능을 가지고 있지만, 형태 정합(pattern matching)이 가지고 있는 일반적인 문제점인 형판과 정합 대상이 같은 크기를 가지고 있어야 한다는 문제를 가지고 있으며, 이를 위한 탐색에 소요되는 비용을 감안하지 않을 수 없다. 또한, 학습 시의 환경과 인식 시의 환경이 동일해야 한다는 즉, 환경에 적응적인 인식이 어렵다는 문제점을 가지고 있다[16]. 마지막으로, KL 변환에 사용되는 입력 데이터는 하나의 벡터로 표현되기 때문에 여러 가지 혼합적인 특징을 사

용하기 어렵다는 문제를 가지고 있다.

본 논문에서는 이러한 문제점들 중에서 마지막의 문제점을 극복하는 한 방안으로 컬러 고유 공간을 표현하는 방법을 제안하고, 기존의 방법과 비교하여 영상을 분류하여 제안하는 방법의 우수성을 보이려고 한다.

2 장에서는 컬러 고유 공간을 표현하는 방법을 설명하고, 3 장에서는 컬러 고유 공간을 이용하여 영상을 분류하는 방법을 설명하며, 4 장에서 실험 결과를 보이고, 5 장에서 결론을 맺는다.

## 2. 컬러 고유 공간 표현 방법

### 2.1. 기존의 고유 공간 표현 방법

얼굴 영상  $I(x,y)$ 를 밝기 값의 2차원 ( $M \times M$ ) 배열이라고 하자. 하나의 영상은  $N^2$  차원의 벡터로 볼 수 있다. 그래서  $256 \times 256$  크기의 영상은 65,536 차원의 벡터가 되며, 결국 65,536 차원 공간의 한 점으로 표현될 수 있는 것이다. 영상들의 집합은 결국 거대한 차원의 공간에서의 많은 점들의 집합이 된다.

얼굴 영상들은 전체적인 구성을 동일하지만 거대한 차원의 공간에서 무작위로 분산되어 있지는 않게 된다. 그래서, 다른 영상과 구분에 영향을 주지 않거나 영향을 주지 않는 한도 내에서 차원을 제거한다면 비교적 적은 차원의 부공간으로 얼굴 영상들을 표현할 수 있다. Principal component analysis(PCA 또는 Karhunen-Loeve expansion)의 주된 아이디어는 영상 공간 전체에서 얼굴 영상의 분포를 가장 잘 표현해 주는 벡터들을 찾는 것이다. 이들 벡터들은 얼굴 영상의 부공간을 정의하는 것이 되며, 이 부공간을 “얼굴 공간(face space)”라고 하자.  $N^2$  길이의 각 벡터는  $M \times M$  영상을 표현하는 것이며, 원래 영상들과의

선형 조합(linear combination)을 한 것이다. 이들 벡터들은 원래 얼굴 영상들에 대한 공분산 행렬(covariance matrix)의 고유 벡터(eigenvector)들이고, 그들이 나타나는 형태가 얼굴과 유사하기 때문에 이를 “고유 얼굴(eigenface)”들이라 칭한다.

얼굴 영상들의 혼련 집합을  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_M$  이라 하자. 이 집합의 평균 얼굴을  $\bar{\mathbf{z}} = \frac{1}{M} \sum_{n=1}^M \mathbf{z}_n$  으로 정의한다. 그런 다음 각 얼굴과 평균 얼굴로부터 차를 구하여 얼굴 영상  $\mathbf{Q}_i = \mathbf{z}_i - \bar{\mathbf{z}}$  를 구한다. 이러한 매우 큰 벡터들의 집합에 대한 PCA 문제를 해결하기 위해 다음과 같은 사항을 고려해 보자. 여기에서 구하고자 하는 벡터들은 데이터의 분산을 가장 잘 기술해주는  $M$  개의 orthonormal 벡터들( $\mathbf{u}_n$ )이다.  $k$  번째 벡터( $\mathbf{u}_k$ )에 대하여

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (\mathbf{u}_k^T \mathbf{Q}_n)^2 \dots\dots\dots (1)$$

이 최대가 되는  $\mathbf{u}_k$ 를 선택하게 되는데, 여기에서  $\mathbf{u}_k$ 는 orthonormal 하기 때문에 다음과 같이 된다.

$$\mathbf{u}_l^T \mathbf{u}_k = \delta_{lk} = \begin{cases} 1, & \text{if } l = k \\ 0, & \text{otherwise} \end{cases} \dots\dots\dots (2)$$

여기에서 벡터들  $\mathbf{u}_k$ 와 스칼라들  $\lambda_k$ 는 각각 공분산 행렬  $\mathbf{C}$ 의 고유 벡터들과 고유값들이다.

$$\mathbf{C} = \frac{1}{M} \sum_{n=1}^M \mathbf{Q}_n \mathbf{Q}_n^T \dots\dots\dots (3) \\ = \mathbf{A} \mathbf{A}^T$$

여기에서 행렬  $\mathbf{A} = [\mathbf{Q}_1 \mathbf{Q}_2 \dots \mathbf{Q}_M]$ 이다. 그러나, 행렬  $\mathbf{C}$ 는  $N^2 \times N^2$ 이고,  $N^2$  개의 고유 벡터들

과 고유값들이 결정되게 되는데, 이를 위한 계산은 영상의 크기가 일반적으로 매우 큰 행렬로 나타나기 때문에 매우 다루기 힘든 문제가 된다. 그래서, 이러한 고유 벡터들을 찾기 위한 계산적으로 가능한 방법이 필요하다.

만일, 영상 공간에서 데이터 점들의 수가 그 공간의 차원보다 작다면( $M < N^2$ ), 의미 있는 고유 벡터들의 수는  $N^2$  개가 아닌  $M-1$  개 뿐일 것이다. 나머지 고유 벡터들은 그에 대응되는 고유값들이 거의 0의 값을 가지게 될 것이다. 다행히도  $M \times M$  행렬(즉,  $16,384 \times 16,384$  행렬이 아닌  $16 \times 16$  행렬)의 고유 벡터들을 구하고, 얼굴 영상들  $Q_i$ 들과 선형 조합하여,  $N^2$  차원의 고유 벡터들을 구하기 위한 방법이 존재한다. 이를 수학적으로 증명하기 위하여 다음과 같은  $A^T A$ 의 고유 벡터들인  $v_i$ 를 고려해 보자.

$$A^T A v_i = \mu_i v_i \dots\dots\dots (4)$$

양변의 좌측에  $A$ 를 곱하면 다음과 같다.

$$A A^T A v_i = \mu_i A v_i \dots\dots\dots (5)$$

여기에서  $A v_i$ 는  $C' = A A^T$ 의 고유 벡터들이다.

이러한 분석에 따라서,  $M \times M$  행렬  $C' = A A^T$ (여기에서  $C'_{mn} = Q_m^T Q_n$ )을 만들고,  $C'$ 의  $M$ 개의 고유 벡터들인  $v_i$ 를 찾는다. 이 벡터들에서 고유 얼굴들  $u_i$ 를 구하기 위해 다음과 같이  $M$ 개의 훈련 집합의 얼굴 영상들과 선형 조합을 하게 된다.

$$u_i = \sum_{k=1}^M v_{ik} Q_k \dots\dots\dots (6)$$

이렇게 함으로써  $N^2$  차원의 고유 벡터를 구하는 문제에서  $M^2$  차원의 고유 벡터를 구하는 문제로 축소되었다. 고유 벡터를 계산하는데 필요한 산술 연산이 매우 많기 때문에  $N^2$  차원에서  $M^2$  차원에서의 차원 축소는 계산 속도를 크게 절감할 수 있게 한다.

## 2.2. 컬러 고유 공간 표현 방법

고유 벡터를 구하기 위한 절차의 입력 영상으로는 같은 성질의 값들(예를 들어, 밝기 값들)의 배열로 된 것을 사용한다. 그러나, 이러한 한 가지의 값만을 이용하는 것보다는 여러 가지의 값(즉, 컬러)들 또는 특징들을 이용하는 것이 영상 분류에 효과적이라는 것은 모두가 인식하고 있는 바이다. 그렇다면, 컬러를 표시하는 값들에 대한 컬러 고유 벡터를 구하는 방법이 문제가 된다. 이를 해결함으로써 컬러 영상 뿐만 아니라 여러 가지 특징으로 만들어진 영상에 대한 고유 벡터를 구할 수 있게 된다.

영상에서 사용하는 컬러 공간은 RGB, YIQ, HSL 등 많은 종류의 컬러 공간이 있다. 특히 컬러를 컴퓨터 화면에 표시하는 경우 일반적으로 RGB를 사용하는데, 이럴 경우 R과 G와 B 요소가 24비트로 된 하나의 값으로 표현된다. 즉, 세 개의 값이 서로 연결되어 하나의 값으로 되는 것이다. 이러한 값은 연속적이지 않기 때문에, 이러한 값으로 된 컬러 영상을 그대로 고유 벡터를 구하는데 사용할 수는 없다. 그래서, 본 연구에서는 식 (7)과 같이 RGB 컬러 공간에서 R의 영상, G의 영상, 그리고 B의 영상을 일렬로 배치함으로써 컬러 영상에 대한 고유 벡터를 구할 수 있도록 설계하였다. 이렇게 했을 때, 이러한 영상들로부터 만들어지게 되는 공분산 행렬은 각각의

영상 내에서의 관계뿐만 아니라 서로 간의 관계까지도 고려할 수 있게 된다.

$$\mathbf{z}' = [\mathbf{z}_R \mathbf{z}_G \mathbf{z}_B] \dots \dots \dots (7)$$

### 3. 고유 공간을 이용한 영상 분류

#### 3.1. 영상 분류를 위한 고유 공간의 표현

고유 벡터가 영상의 많은 변화를 표현하는데 매우 유용하다는 것은 이미 널리 알려져 있기 때문에, 본 연구에서는 얼굴 영상을 분류하기 위하여 얼굴에 대한 고유 벡터로 만들어지는 고유 얼굴 공간에 얼굴 영상을 투영하여(즉, 그 영상이 각 고유 얼굴에 미치는 기여도를 계산하여) 구해지는 기여 벡터를 분류함으로써 얼굴 영상을 분류한다.

식 (8)은 입력 영상  $z'_m$ 이  $k$ 번째 고유 얼굴이 미치는 기여도를 계산하는 식이다. 식 (9)는 식 (8)에서 얻어지는 값들의 벡터로 고유 벡터를 표시한다.

$$\omega_k = \mathbf{u}_k^T (\mathbf{z}'_m - \bar{\mathbf{Z}}) \dots \dots \dots (8)$$

$$\Omega^T = [\omega_1 \omega_2 \dots \omega_M] \dots \dots \dots (9)$$

이들은  $M$ 차원의 벡터 공간을 가지며, 하나의 영상은  $M$ 차원 벡터 공간 안의 한 점으로 표현이 된다.

기여 벡터를 구하는데는 많은 양의 기억 공간과 많은 계산량을 필요로 한다. 특히, 훈련을 위한 얼굴 영상이 많을수록 영상을 분류하는데 드는 비용은 급격히 증가한다. 그러나 다행히도 좀 더 적은 수의 고유 얼굴만을 가지고도 얼굴 영상 분류를 할 수 있는 방법이 존재한다.

고유값은 고유 벡터의 중요도를 결정짓는 것

이다. 즉, 고유값이 크면 클수록 그에 대응되는 고유 벡터는 고유 벡터를 구하는데 사용되어진 벡터를 분류하는데 중요한 역할을 하는 것이다. 본 연구에서는 이러한 고유 값의 임계값을 결정하기 위하여 다음과 같은 방법을 사용하였다.

$$W_k = \frac{\sum_j^k \lambda_j}{\sum_i^M \lambda_i} \geq T_s \dots \dots \dots (10)$$

이 식은  $W_k$ 가 미리 정해진 임계치인  $T_s$ 보다 큰 가장 작은  $k$ 번째까지의 고유 값에 해당되는 고유 벡터만을 사용하고자 하는 것이다. 여기에서 구해진 가장 큰  $k$ 값은  $M'$ 개의 고유 벡터로 이루어지는 고유 얼굴 공간을 사용한다. 이 고유 공간은 얼굴 영상을 완벽하게 복구하지는 못하지만 영상 분류에는 큰 영향을 미치지 않는다는 사실은 많은 문헌들에서 설명하고 있다. 이렇게  $M'$ 개의 고유 얼굴로부터 만들어진 기여 벡터는  $M'$ 차원이 되고, 얼굴 영상을 분류하는데 소요되는 비용은 극히 감소하게 된다.

#### 3.2. 얼굴 영상 분류

사람의 얼굴의 생김새는 안경을 쓴 사람, 콧수염이 있는 사람, 얼굴이 넓직한 사람, 얼굴이 긴 사람, 턱이 뾰족한 사람, 이마가 넓은 사람 등 그 형태가 매우 다양하다. 이러한 얼굴의 형태 분류를 위해서는 일반적으로 얼굴 영상에서 에지를 추출하여 에지의 분포를 분석하여 얼굴의 형태를 결정하였다. 그러나, 에지를 사용하는 분석 방법은 조명이나 영상의 질에 크게 영향을 받으며, 에지를 추출하고 분석하는 과정에 많은 비용을 투자해야 한다는 단점을 가지고 있다.

본 연구에서는 기여 벡터들의 집합에 대하여 여러 개의 군집으로 분류함으로써 간단하게 얼굴

을 분류하였다. 즉, 하나의 기여 벡터는  $M'$  차원의 벡터 공간의 한 점으로 표현되므로 이러한 점들간의 거리를 식 (11)과 같이 계산하여 이를 기준으로 영상을 분류한다.

$$d_c = \|\Omega - \Omega_c\|^2 \dots\dots\dots (11)$$

여기에서  $c$ 는  $c$  번째 군집을 나타낸다.  $\Omega$ 는 입력 영상에 대한 기여 벡터이며,  $\Omega_c$ 는  $c$  번째 군집의 대표치(무게 중심)이다. 그러므로  $d_c$ 는 두 기여 벡터의 차에 대한  $L_2$  노름이 된다.

#### 4. 실험 결과

본 연구는 Pentium PC®에서 구현되었으며, Metrox Meteor 영상 처리 보드와 컬러 CCD 카메라를 사용하였다. 컴파일러는 Visual C++ 4.0을 사용하였으며, 운영체제는 Window 95를 사용하였다.

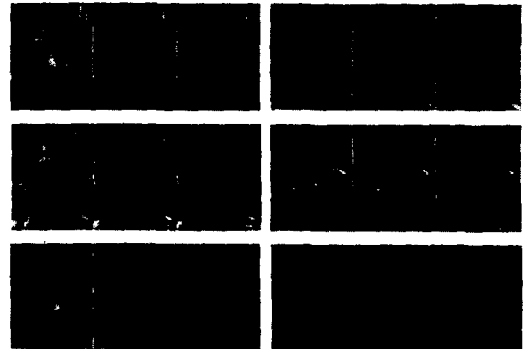
입력 영상으로는 256 컬러의 RGB 영상을 사용하였다. 얼굴 형태를 5 가지 종류로 분류하고, 각각의 형태에 대하여 5 명의 사람에 대한 총 200 개의 영상을 획득하고, 이 중에서 20 개는 고유 얼굴을 구하는데 사용하고, 고유 얼굴을 구하는데 사용된 영상과 80 개의 영상은 군집화 (clustering)를 위해 사용하였으며, 100 개는 분류 시험에 사용하였다. <그림 1>은 사용되어진 영상의 예를 보여준다.

표 1. 분류 결과 비교표

	컬러 영상	흑백 영상
실험 횟수	100	100
정분류	95	91
오분류	5	9
분류율	95 %	91 %

표 1은 실험에 사용된 영상들을 기존의 흑백 영상들로 변환하여 분류한 결과와 본 논문에서 제안하는 방법으로 분류한 결과를 비교하고 있다.

그림 1은 분류를 위해 사용된 얼굴 영상의 예를 보여주고, 그림 2는 평균 얼굴을, 그리고 그림 3은 고유 얼굴들을 보여준다. 여기에서 보여주는 영상은 R, G, 그리고 B에 대한 영상을 흑백으로 표현한 3 개의 영상이 하나의 영상으로 나타나 있다.



<그림 1> 입력 영상의 예



<그림 2> 평균 얼굴

#### 5. 결론

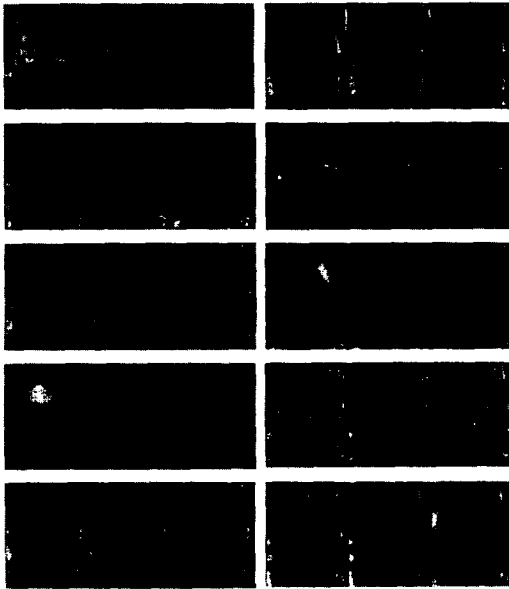
본 논문에서는 고유 벡터를 이용한 인식 분야에 컬러 영상의 특징을 반영하는 시도를 하였다. 즉, 컬러라는 특징 값들에 의해 고유 벡터를 산출함으로써, 더욱 정확한 분류와 정합을 할 수 있음을 보였다.

컬러 고유 벡터는 현 시대에 맞는 컬러 영상을 다룰 수 있는 방법을 제시하였으며, 좀 더 많은 특징을 사용할 수 있다는 가능성을 보여준 것

이다.

앞으로 본 연구를 실제 환경에 적용함에 있어서 몇 가지 문제점을 해결해야 한다. 첫째는 탐색에 드는 비용을 최소화해야 하며, 둘째는 더욱 정확한 정합을 위한 후 처리가 필요하다.

또한, 이러한 정합에서 만들어지는 결과를 활용하는 방안을 모색함으로써, 어떠한 응용 분야에도 사용할 수 있도록 연구를 계속할 계획이다.



<그림 3> 가장 좋은 10 개의 고유 얼굴들

### 참고 문헌

[1] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience* Vol.3, No.1, 1991.

[2] Pentland, A., Starner, T., Etcoff, N., Masoiu, A., Oliyide, O., and Turk, M., "Experiments with Eigenfaces," *IJCAI'93*, MIT Media Lab. TR#-194, 1992.

[3] Matthew Turk and Alex Pentland, "Face Recognition without Features," *IAPR Workshop on Machine Vision Application(MVA'90)*, pp.267-270, 1990.

[4] M. Kirby and L. Sirovich, "Application of Karhunen-Loeve Procedure for the Characterization

of Human Faces," *PAMI* Vol.12, No.1, pp.103-108, 1990.

[5] Baback Moghaddam and Alex Pentland, "Probabilistic Visual Learning for Object Detection," *5<sup>th</sup> ICCV*, 1995, MIT Media Lab. TR#-326, 1995.

[6] Z. Hong, "Algebraic feature extraction of image for recognition," *Pattern Recognition*, Vol.24, pp.211-219, 1991.

[7] Alan L. Yuille, Peter W. Hallinan, and David S. Cohen, "Feature Extraction from Face Using Deformable Templates," *ICCV'92*, pp.99-111, 1992.

[8] B. S. Manjunath, R.Chellappa, and C. v.d. Malsburg, "A feature based approach to face recognition," *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.373-378, 1992.

[9] R. Brunelli and T. Poggio, "Face Recognition through Geometrical Features," *Proc. ECCV'92*, pp.792-800, <http://hera.itc.it:3003/~brunelli/Papers.html>, 1992.

[10] Y. Cheng, K. Liu, J. Yang and H. Wang, "A robust algebraic method for human face recognition," *ICPR'92*, pp.221-224, 1992.

[11] I. Craw, H. Ellis, and J. Lishman, "Automatic extraction of face features," *Pattern Recognition Letters*, Vol.5, pp.183-187, 1987.

[12] Y. Yacoob and L. S. Davis, "Computing spatio-temporal representations of human faces," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.70-75, 1994.

[13] G. Yang and T. S. Huang, "Human Face Detection in a Scene," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.453-458, 1993.

[14] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade, "Human Face Detection in Visual Scenes," *CMU-CS-95-158R*, 1995.

[15] Ingemar J. Cox, Joumana Ghosn, and Peter N. Yianilos, "Feature-Based Face Recognition Using Mixture-Distance," *©NEC Research Institute Technical Report 95-09*, 1995.

[16] Kohtaro Ohba and Katsushi Ikeuchi, "Recognition of the Multi Specularity Objects using the Eigen-Window," *CMU CS Technical Report CMU-CS-96-105*, 1996.

[17] Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, Andrew D. Back, "Face Recognition: A Hybrid Neural Network Approach," *Technical Report UMIACS-TR-96-16 and CS-TR-3608*, University of Maryland, 1996.