

# 최대고유성분해석을 이용한 인간의 얼굴표정인식

김영익\* · 정광민\*\* · 조석제\*\*\*

## Human Facial Expression Recognition Using Principal Component Analysis

*Young-Ik Kim, Kwang-Min Jeong, Seok-Je Cho*

### Abstract

A probabilistic approach for the recognition of human facial expression is proposed. The algorithms we developed utilize K-L expansion computation to extract the principal components of facial expression and use the input face image as a whole pattern for generating a new pattern of features of it. Recognition of four facial expression is experimented and evaluated.

### 1. 서 론

인간은 얼굴표정을 통해 언어만으로는 표현할 수 없는 무수한 정보를 나누고 있다. 그 예로 그림 1에서는 종족이나 사회를 초월한 인간의 공통적인 몇 가지 얼굴표정(기쁨, 무표정, 놀람, 그리고 웅크)들을 보여주고 있다. 이러한 표정들에 대한 연구는 이미 오래 전부터 인지과학자(cognitive scientist)들에 의해 이루어져 왔으며, 근래에는 컴퓨터 비전(computer vision)이나 그래픽(graphics)분야에서 인지 컴퓨팅(perceptual computing)이란 주제로 논의가 되어 왔다. 또한 최근에 들어 와서는 인간과 컴퓨터의 상호작용(HCI : human computer interaction)이라는 요구에 맞추어, 컴퓨터에게 인간의 표정을 자동으로 인식할 수 있는 능력을 부여 하고자 하는 연구가 활발하게 진행되고 있다.

\* 해양해양대학교 대학원

\*\* 경남정보대학 전자정보통신과 전임강사

\*\*\* 한국해양대학교 이공대학, 자동화정보공학부 부교수

기존의 얼굴표정 인식(facial expression recognition)에 관한 연구방법은 크게 두 가지로 나누어 볼 수 있는데, 그 하나는, 얼굴에서의 국부적인 특징부분들(눈, 코, 입 등)을 서로 비교함으로써 그 변화 정도에 의해 얼굴 표정을 인식하고자 하는 것이다<sup>11)2)</sup>. 이를 위해서는 얼굴영상에서 정확하게 특징부분을 추출하는 일이 선행되어야 하며, 실제 영상에서는 조명의 차이라든가 사람간의 개인적인 차이를 극복하는 데 있어 어려움이 있었다.

또 하나의 연구방향으로는 표정의 변화에 따른 얼굴근육의 움직임 정보(optical flow)를 사용하여 얼굴표정을 구분한다<sup>13)~15)</sup>. 이 방법에서는 순차적으로 입력되는 얼굴영상으로부터 움직임을 찾아내고, 이를 분석하여 얼굴표정이 가지는 특징요소를 찾아낸다. 하지만, 얼굴표정 인식의 성능을 좌우하고 있는 움직임 정보를 일반적인 환경 속에서 정확하게 구해내기 힘들뿐더러, 이

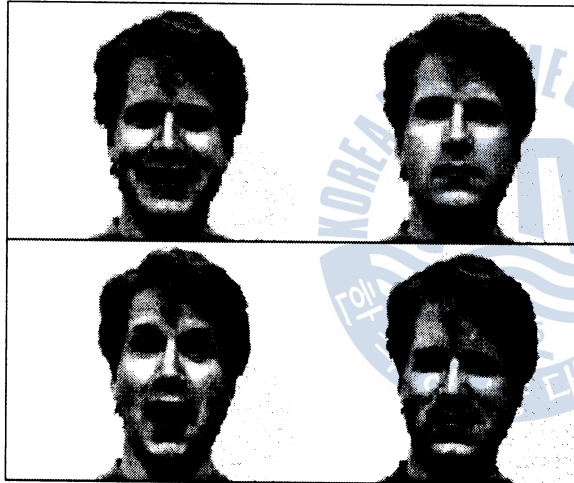


그림 1. 인간의 4가지 얼굴표정

러한 방법에서는 단일영상에 대한 표정인식에 대해서는 생각하지 않고 있다.

본 연구에서는 Kirby와 Sirovich가 제안하였고<sup>19)</sup>, Turk와 Pentland에 의해 고유얼굴(eigenface)이라는 개념<sup>110)</sup>으로 얼굴 인식과정에서 많이 쓰이고 있는 확률 통계적 접근방법에 기반을 둔다. 이 방법에서는 K-L전개를 통해, 각 얼굴 표정마다 지니고 있는 고유의 성분만을 그 특징 값으로 찾아낸다. 그리고 실제 얼굴표정의 인식단계에서는 전처리 과정을 통해 정규화 된 입력 얼굴 영상을 K-L전개를 통해 구해진 최대

고유 얼굴표정 공간(eigen facial expression space)에 단순히 투사(projection)함으로써 입력된 영상의 얼굴이 기존의 모집단으로부터 형성된 고유 얼굴표정과 가지는 유사 정도를 쉽게 그 특징 값으로 찾아내고 있다.

실험으로는 그림 1에서 보여주고 있는 4가지 종류의 얼굴 표정에 대해 각 10명의 모집단을 구성한 다음, 표정마다 5개의 최대 고유얼굴표정(EFE : eigen facial expression)을 만들어 냈다. 그리고 새로운 25개의 얼굴표정에 대해 테스트 해 본 결과 92%의 만족스러운 인식률을 얻을 수 있었다. 이는 인간이 사람의 얼굴표정을 인식해내고 있는 능력과 비교할 수는 없겠으나, 비교적 소규모 모집단으로부터 고유의 얼굴표정의 특징을 찾아낼 수 있었다는 점에서 본 연구의 접근방법이 인간이 행하고 있는 인식방법과 많은 유사성을 가지고 있다는 것을 보여주고 있다.

## 2. 고유얼굴표정을 이용한 얼굴표정인식 시스템

얼굴표정 인식시스템의 전체적인 구성은 크게 전처리 단계, 고유 얼굴표정 만들기 단계, 인식 단계의 세 가지로 나누어 볼 수 있다.

### 2.1 전처리 단계

일반적인 시스템 구성을 위해서는 가장 먼저, 입력된 영상에서 얼굴부분을 찾아내는 과정이 수반되어야 한다. 그러나 본 논문에서는 그 주제의 복잡성으로 인해, 일단 영상에서 얼굴의 위치가 찾아졌다고 가정하고, 영상 속에서 얼굴을 찾아내는 방법에 대해서는 다른 연구 결과를 참고하기로 한다<sup>[6]-[8]</sup>.

#### 2.1.1 얼굴영상 크기의 정규화

입력되는 얼굴영상이 다소 각도의 차이나 거리의 차이를 보이게 되더라도 사람의 얼굴이 지니는 좌우 대칭성과 얼굴내부에 있는 눈, 코, 입 등이 상대적으로 균등한 위치를 갖고 있다는 성질을 이용하여, 그 특징부분의 크기를 정규화 시킬 수 있다.

그림 2는 실제 시스템으로의 구현 가능성을 고려하여 본 연구에서 사용한 방법으로, 입력된 영상에서 얼굴의 특징부분만을 선택한 뒤, 그 크기를 특정 크기(70×100 픽셀)로 정규화 하는 한 예를 보여주고 있다.

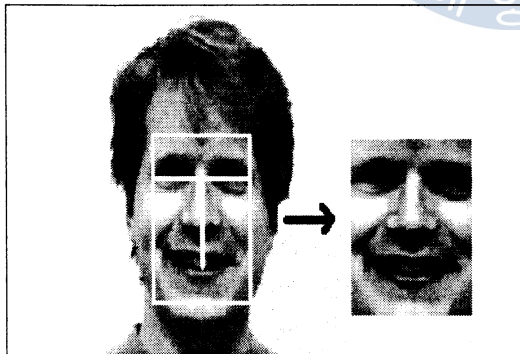


그림 2. 얼굴영역의 크기 정규화

- 정규화를 위한 세부적 절차는 다음과 같다.
- 1) 왼쪽 눈의 왼쪽 끝점(LEE : left eye ending point), 오른쪽 눈의 오른쪽 끝점(REE : right eye ending point), 입의 중심(COM : center of mouth)을 각각 찾아낸다.
  - 2) 양쪽 눈 끝점의 중앙값을 얼굴의 중심(COF : center of face)으로 잡는다.
  - 3) COF와 COM간의 거리 차 ( $dx$ ,  $dy$ )를 계산한다.
  - 4) COM으로부터 얼굴표정의 변화영역에 속하는 이마부분과 턱 부분을 포함하기 위해 얼굴특징영역을 COF과 COM으로부터  $dy/2$ 만큼 상하영역을 확장한다.
  - 5) LEE, REE와  $COF+dy/2$ ,  $COM-dy/2$ 로 정의되는 얼굴영역을 그 x축과 y축간의 길이 비율에 상관없이 70×100 픽셀로 정규화 시킨다.
- 상기의 과정으로 비교적 간단하게 얼굴상의 특징 점들(LEE, REE, COM)이 동일한 위치를 가지게 되는 정규화 된 크기의 얼굴특징영역을 구해 낼 수가 있다.

### 2.1.2 히스토그램의 정규화

입력된 영상들의 얼굴간에는 조명이나, 주인공의 차이에 따라 얼굴영상의 히스토그램 분포

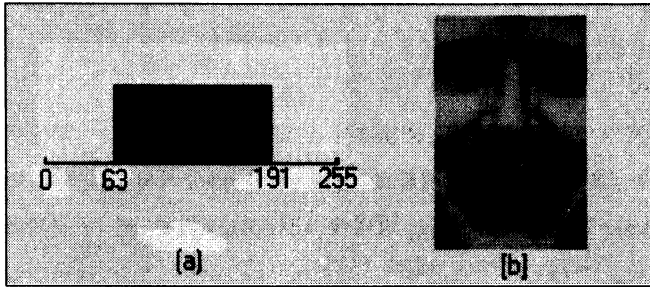


그림 3. 히스토그램 정규화  
(a)특정 분포 (b)변화된 얼굴영역

가 상당한 차이가 난다. 이러한 점은 얼굴의 주인공을 찾아내는 데에는 유용하게 사용될 수 있으나, 사람들의 얼굴표정간 공통점을 찾기 위해서는 그 차이를 일정하게 줄여줄 필요가 있다.

그림 3에서는 본 연구에서 사용한 특정 히스토그램 분포와 이를 통하여 정규화 된 얼굴 특징영역의 히스토그램을 보여주고 있다.

### 2.2 고유 얼굴표정 만들기 단계

얼굴과 같은 영상은 모든 인간을 그 모집단으로 하였을 때, 그 특징적인 요소들(눈, 코, 입등)의 전체적인 구성이 대단히 유사하기 때문에, 이들로부터 저차원의 최고 고유성분들을 얻어낸다면 유용한 정보로 사용할 수가 있다. 본 논문에서는 인간의 얼굴 표정들간의 공통적인 특징을 취하기

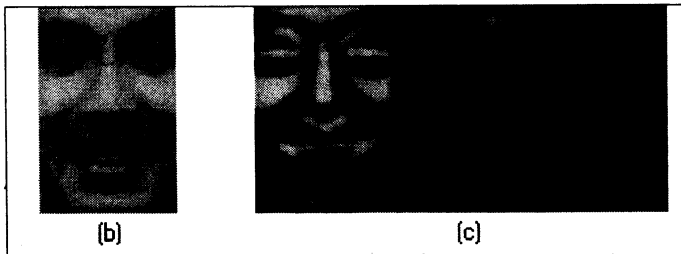
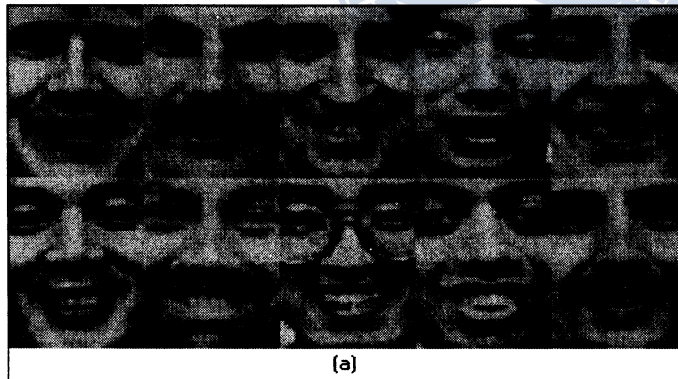


그림 4. 고유얼굴표정 만들기  
(a)웃음 얼굴표정 모집단 (b)평균 웃음 얼굴표정  
(c)구해진 3장의 최대고유 얼굴표정들

위해 각 표정들 마다 M개의 모 집단을 구성하고, 이들 속에서 F개의 최고고유성분을 구해내어 고유 얼굴표정으로 삼았다.

그림 4에서는 본 연구에서 사용한 웃는 얼굴표정(HFE:happy FE)의 모집단(M=10) 및 그들의 평균 웃는 얼굴표정(MHFE : mean happy FE)과 구해진 3장(F=3)의 고유 웃는 얼굴표정을 보여주고 있다.

세부적인 절차는 다음과 같다.

1)  $X \times Y$  픽셀크기의 M장의 모 집단 얼굴표정영상 ( $\Gamma_1, \Gamma_2, \dots, \Gamma_M$ )으로부터 픽셀의 위치에 따른 M차원의 패턴 벡터  $\chi$ 를

생각한다.

$$x = [x_1, x_2, \dots, x_M]^T \quad (1)$$

2) 패턴벡터  $x$ 공간의 전체  $X \times Y$ 개의 모집단 벡터들에 대해 평균값 벡터  $m_x$ 를 구한다.

$$m_x = E\{x\} \quad (2)$$

3) 전체 모집단 벡터들의  $m_x$ 에 대한 공분산 행렬  $C$ 를 구한다.

$$C = E\{(x - m_x)(x - m_x)^T\} \quad (3)$$

4) 행렬  $C$ 로부터  $F$ 개의 최대 고유치를 가지는 벡터들 ( $v_1, v_2, \dots, v_F$ )을 계산해 낸다.

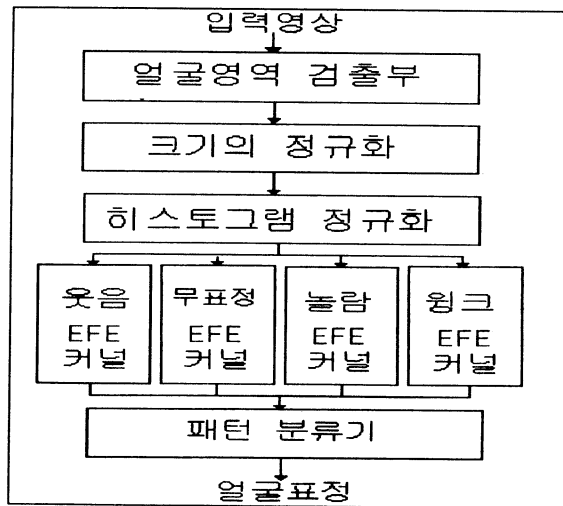
5) 식(4)에 의해 각 픽셀들의 고유값을 구함으로써 해서  $F$ 장의 고유얼굴표정 ( $u_1, u_2, \dots, u_F$ )을 구한다

$$u_l = \sum v_l x \quad l=1,2,\dots,F \quad (4)$$

### 2.3 얼굴표정 인식 단계

새로운 입력영상이 주어졌을 때, 이로부터 얼굴표정을 판단하기까지의 전체적인 블록 다이어그램을 그림 5에서 도시하였다. 여기에서 정규화 된 입력영상  $I$ 가 식(5)로 표현되는 고유얼굴표정의 커널에 투사되어  $F$ 개의 특징 값( $\omega_k$ )을 구하고 있다.

$$\omega_k = u_k^T (I - \Psi) \quad (5)$$



이때 구해진  $\Omega^T = [\omega_1, \omega_2, \dots, \omega_k]$ 의 가중치 특징벡터는 일반적으로 쓰이고 있는 패턴인식 알고리즘에 의해 직접적으로 사용되어지며, 이로부터 쉽게 입력영상의 얼굴표정을 구분해 낼 수 있다.

본 연구에서는 패턴인식 알고리즘으로 가장 간단하다고 할 수 있는, 최소거리 분류기에 의해 식(6)의 유클리디언 거리 ( $\epsilon_k$ )를 최소화하는  $k$ 번째 부류의 얼굴표정으로 패턴을 분류하도록 하였다.

$$\epsilon_k = \|(\Omega - \Omega_k)\|^2 \quad (6)$$

그림 5. 얼굴표정인식 시스템의 흐름도



### 3. 실험 및 고찰

본 연구에 의해 제안된 얼굴표정인식 시스템의 얼굴표정인식능력을 평가해 보기 위하여 다음의 두 가지 실험집단을 구성하였다. 그림 6은 미국 Yale대학에서 실험용으로 제작한 얼굴영상들로서 고유얼굴표정의 모집단에 속하지 않은 영상들로 구성하였다. 그리고 그림 7에서는 실제 영상에의 적용 예를 보이기 위해 일간 신문에서의 영상을 스캐너로 읽어 들인 것이다.

실험과정에서 입력얼굴영상에 대한 얼굴영상의 검출부분은 직접 PaintShop Pro 4.0에서 작업하였으며, 70×100픽셀로의 크기정규화 부분 역시 PaintShop Pro 4.0의 Resize기능을 이용하였다. 다음으로 히스토그램 정규화부분과 고유얼굴표정(EFE)의 커널부분은 Sun- Solaris2.5환경에서 GNU C\C++ 1.09에 의해 제작되었다.

히스토그램 정규화에 사용된 히스토그램 분포는 그림 3과 같고, 각 EFE 커널은 5개의 최대 고유성분을 가지고 입력영상과의 가중치를 구해내었다. 하나의 입력영상에 대하여 4개의 EFE 커널로부터 구해진 총 20개의 가중치 값이 하나의 패턴 벡터로 최소거리 분류기의 입력으로 사용되었다.

실험결과 집단 A에 대해서 얼굴표정인식 시스템은 20개의 입력영상에 대해 18개의 올바른 결과를 보여 주었으며, 집단 B에 대해서는 5개의 입력영상에 대해 모두 올바른 결과를 나타냈다.

EFE 커널의 출력패턴을 분석해본 결과, 웃는 얼굴표정에 대해서는 뚜렷한 유클리디언 거리

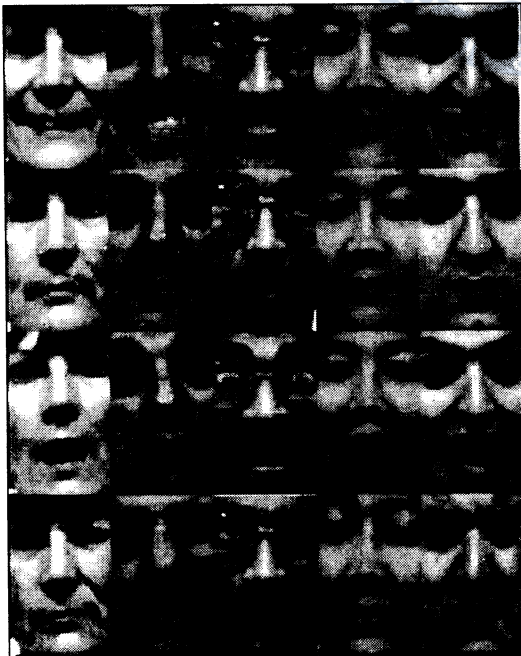


그림 6. 실험집단 A (미국 yale대학의 Face Database에서 가져온 것임)



그림 7. 실험집단 B (동아일보 97년 11월 27일자에서 스캐너로 읽은 영상)

차이를 보여주었으며, 링크에 대해서는 다소 작은 거리차이가 나타났다. 이는 웃는 얼굴이 링크 보다는 많은 사람들 사이에서 더 많은 공통요소를 지니고 있음을 보여주고 있고, 또한 각 EFE 커널에서 출력된 5개의 가중치 값들을 검토해본 결과 3개의 최대고유성분만으로도 충분히 얼굴 표정을 구분해 낼 수가 있음을 알 수 있었다. 그리고 더욱 정교한 얼굴표정의 차이를 검출하기 위해서는 각 개인의 얼굴표정에 대한 데이터베이스화된 정보가 필요하다.

#### 4. 결론 및 향후과제

본 연구에서는 인간의 얼굴표정을 인식하기 위한 접근방법으로서 확률 통계적인 K-L전개에 의한 최대고유성분 해석법을 사용하였다. 얼굴전체에 대해서 특징 값을 찾아냄으로써, 얼굴표정을 좌우하는 특징요소를 찾아야 하는 부담이 없으며, K-L전개의 특성으로 인해 많은 얼굴표정들이 가지는 최대의 공통성분만을 사용하기 때문에 정보량에 비해 그 계산량이 상당히 작아, 실시간 얼굴표정 인식시스템을 구현하고자 할 때에는 상당한 매력을 지니고 있다. 실제로 본 연구의 실험에서는 한 장의 입력영상에 대해서 얼굴을 찾아내는 과정을 생략한 총 계산 시간이 Ultra Sun-Sparc상에서 0.1초 미만이었다.

향후 연구되어야 할 부분들로는 일반적인 환경 하에서 사용될 수 있는 강건한 시스템을 위해 다양한 조명환경을 극복할 수 있도록 히스토그램 정규화과정을 개선하는 것과 입력영상의 얼굴이 어느 정도의 각도를 가지게 될 때에 이를 정면얼굴로 매핑시킬 수 있는 정규화 과정 등을 들 수가 있다. 또한 본 연구에서는 수작업에 의해 처리된 바 있는 정확한 얼굴영역을 검출할 수 있는 알고리즘을 구해내는 데에 관한 연구가 필요하다.

그리고 장기적으로는 대규모의 실험집단에 적용되었을 때 최대 고유 얼굴표정에서 발생되는 다양한 패턴벡터로부터 유용한 판단을 내릴 수 있는 인식구조로서 신경망 등을 이용하는 방법과, 또한 각 개개인마다 발생하는 표정의 정보를 데이터베이스화하여 보다 정교한 얼굴표정 인식시스템을 구현해 나가는 것 등에 관한 연구가 요구된다.

#### 참 고 문 헌

- [1] Terzopoulos and K. Waters, "Analysis and Synthesis of Facial Image Using Physical and Anatomical Models", *IEEE trans. Patt. Anal. Machine Intell.*, Vol. 15(6), pp. 569~579, 1993
- [2] M. A. Shackleton and W. J. Welsh, "Classification of Facial Feature for Recognition", *Proc. CVPR*, pp. 573~579, 1991
- [3] Y. Yacoob and L. Davis, "Computing Spatio-temporal Representations of Human Faces", *Proc. CVPR*, pp. 70~75, 1994
- [4] I. A. Essa and A. Pentland, "Facial Expression Recognition using a Dynamic Model and Motion Energy", *Proc. of 5th ICCV*, pp. 360~367, 1995
- [5] Satoshi KIMURA and Masahiko YACHIDA, "Facial Expression Recognition and Its Degree

- Estimation", *Proc. CVRP*, pp. 295~300, 1997
- [6] Q. Chen, H. Wu, and M. Yachida, "Face Detection by fuzzy pattern matching", *Int. Conf. on Computer Vision*, pp. 591~596, 1995
- [7] T. Leung, M. Burl, and P. Perona, "Finding faces in cluttered scenes using labeled random graph matching.", *Int. Conf. on Computer Vision*, pp. 637~644, 1995
- [8] Moghaddam and Pentland, "Probabilistic visual learning for object detection.", *Int. Conf. on Computer Vision*, pp. 786~793, 1995
- [9] Kirby. M and Sirovich. L, "Application of the Karhunen-Loeve procedure for the characterization of human faces", *IEEE trans. Patt. Anal. Machine Intell.*, Vol. 12(1), pp. 103~108, 1990
- [10] M. Turk and A. Pentland, "Eigenfaces for Recognition", *J. of Cognitive Neuroscience*, Vol. 3(1), pp. 71~86, 1991

