

# 주요고유성분해석에 의한 인간의 얼굴정보 인식

김 영 익\* · 조 석 제\*\*

## Human Facial Information Recognition Using Principal Component Analysis

Young-ik Kim\* · Seok-Je Cho\*\*

### Abstract

In this paper, we present a computer vision algorithm to extract facial informations from human faces image using principal component analysis. Three types of facial informations are considered for recognition: facial expression, glasses on/off information, and gender information of the face image.

The proposed algorithm has three main stage for the facial information recognition: preprocessing, reference pattern image generation and classification. At the preprocessing stage, we choose different sizes and histograms for face image normalization for robust recognition of each facial information. Then at the reference pattern image generation stage, a set of principal component images for each facial information is constructed from the normalized face images using the principal component analysis method. Each facial information has its own principal component images and these images are used for the recognition of the facial information. At the classification stage, the input face image is projected to the principal component images of each facial information and outputs distance vectors for each facial information. The output distance vectors are used for the classification of the facial information.

The performance of the proposed algorithm is evaluated for three test face image

---

\* 한국해양대학교 제어계측공학과 석사과정

\*\* 한국해양대학교 자동화·정보공학부 교수

sets. And the experimental result gives a recognition rate of 93% on average, which shows the effectiveness of the proposed method.

## 1. 서 론

최근 들어 컴퓨터를 이용하여 얼굴영상에서 표정이나 안경 착용상태 등의 얼굴 시각 정보를 자동으로 인식하는 연구들이 많이 이루어지고 있다<sup>[1-6]</sup>. 이 분야의 연구는 얼굴 영상이 누구인지를 판별해 내는 주인공 인식연구<sup>[7,8]</sup>와 함께 멀티미디어 데이터베이스(multimedia database)에서 얼굴영상 검색, 얼굴영상 데이터의 전송을 위한 데이터 압축(data compression), 감성공학(sensitivity engineering) 등 여러 분야에서 매우 중요한 요소로 이용된다.

얼굴영상에서 시각정보를 얻고자 하는 기존의 연구들은 대부분 얼굴정보 중 빠른 정보인 정서적 정보(슬픔, 기쁨, 놀람 등)의 표정 인식에 중요성을 두고 있으며 다른 시각정보에 대해서는 고려하지 않고 있다<sup>[3, 6]</sup>. 그리고 이러한 기존의 연구들은 인식을 위해 사용하는 얼굴영상의 종류에 따라 동영상을 이용한 접근방법과 정지영상을 이용한 접근방법으로 나눌 수 있다.

Yacoob 등<sup>[4]</sup>과 Essa 등<sup>[5]</sup>은 동영상에서 얼굴근육의 움직임 정보를 추출해 낸 뒤 이를 이용하여 각종 시각정보를 얻어내는 방법을 제안하였다. 이 방법은 정지영상에 의한 방법보다 인식률이 우수하지만, 동영상에서 얼굴근육의 움직임정보를 추출하기가 어려워 실제적인 구현이 힘들고, 정지영상에 대해서는 정보추출이 불가능하다는 단점이 있다. 정지영상에서 얼굴정보를 얻는 방법은 두 가지가 있다. 그 하나는 특징부분 해석방법으로 얼굴의 특징부분에 해당하는 눈, 코, 입 및 얼굴 윤곽선 등을 추출해 낸 뒤 이들에서 원하는 정보를 얻는 방법이다<sup>[2,3]</sup>. 그리고 다른 하나는 얼굴의 전체영역에 대한 해석방법으로 전체 얼굴 영상 속에서 원하는 정보를 얻어내고자 하는 방법이다<sup>[6-12]</sup>. 얼굴의 특징부분 해석방법과 전체영역 해석방법은 상호 보완의 관계에 있다고 할 수 있으며, 최근에는 이 두 가지 방법을 병행하여 보다 나은 인식 결과를 보이고 있다<sup>[7,8]</sup>. 또 하나의 연구방법으로 Turk 등<sup>[10]</sup>은 얼굴영상에 주요고유성분해석(principal component analysis) 방법을 적용하여 고유얼굴(eigenface)에 의한 주인공 인식방법을 제안하였다. 고유얼굴에 의한 얼굴인식은 비교적 알고리즘이 간단하고 그 처리속도가 빨라서 최근 가장 활발한 연구가 이루어지고 있으며, 어느 정도의 성과를 거두고 있다<sup>[9-12]</sup>.

본 논문에서는 기존의 연구들과는 달리 얼굴영상에서 주인공 정보와 표정 정보뿐만 아니라 성별 정보 및 안경착용 정보를 포함하는 다양한 얼굴의 시각정보를 주요고유

성분해석 방법을 이용하여 인식하고자 한다. 인식을 위한 전처리 단계에서는 구하고자 하는 얼굴 시각정보의 종류에 맞게 얼굴영상의 크기 및 히스토그램을 정규화 한다. 정규화 된 얼굴영상들로 모집단을 구성한 다음 주요고유성분해석 방법으로 해당 얼굴정보의 고유성분들을 특징으로 구한다. 그리고 입력 얼굴영상을 얼굴정보의 각 고유성분들에 투사하여 발생하는 특징값 벡터를 최소거리 분류 방법(minimum-distance classification method)에 따라 해당 얼굴정보로 분류한다<sup>13)</sup>.

제안한 알고리즘은 전처리단계에서 구하고자 하는 얼굴정보에 따라 각각 다른 정규화과정을 적용함으로써 얼굴영상에서 조명이나 개인간의 차이(피부색, 콧수염 등)에 따른 영향을 줄이는 한편 얼굴정보에 적합한 주요고유성분을 만들고 있다. 그리고 구해진 주요고유성분을 이용하여 비교적 단순하고 높은 인식률의 얼굴정보 인식과정을 유출하고 있다. 또한 각각의 얼굴정보들을 서로 독립적, 병렬적으로 구할 수 있다는 장점은 실제적인 시스템 구현시 많은 매력을 가진다.

## 2. 얼굴영상의 정규화

본 절에서는 얼굴영상에서 원하는 정보를 추출해 내기 위한 전처리단계로 얼굴영역의 크기 정규화 과정과 히스토그램을 특정 분포로 정규화 하는 과정을 설명한다. 정규화 과정에서 얼굴영상은 구하고자 하는 얼굴정보의 종류에 따라 각각 다른 얼굴영역 크기와 히스토그램 분포를 갖도록 정규화 된다.

### 2.1 크기의 정규화

얼굴영상에서 얼굴영역이 다소 각도나 거리의 차이를 보이게 되더라도 사람의 얼굴이 지니는 좌우 대칭성과 얼굴내부의 특징점(눈, 코, 입 등)들이 얼굴 내부에서 상대적으로 균등한 위치를 갖고 있다는 점을 이용하여 얼굴영상에서 얼굴영역의 크기를 정규화 한다. 정규화과정에서 사용한 얼굴의 특징점은 자동 검출시의 용이함을 고려하여 그림 2.1에서와 같이 양쪽 눈의 끝점과 입의 중심점으로 하였다.

본 논문에서는 미리 구해진 3개의 특징점(양쪽 눈의 좌우 끝점과 입의 중심점)으로부터 얼굴영역을 상하, 좌우의 상대적인 크기만을 정규화 시키는 방법을 사용한다. 정규화 되는 얼굴영역의 크기는 필요로 하는 얼굴정보의 종류에 따라 달리 하고 있는데, 그림 2.1에서는 얼굴의 세 곳의 특징점으로부터 두 가지 크기의 얼굴영역을 획득하는 과정을 보여주고 있다.

얼굴영상에서 표정 정보 및 안경 착용상태정보를 위해서는 그림 2.1(b)처럼 얼굴내

부의 주요 특징부분만을 포함하는  $80 \times 120$ 화소 크기의 정규화 영상을 사용하였고, 주인공정보 및 성별정보를 추출하고자 할 때는 그림 2.1(c)의 얼굴 윤곽선이 모두 포함된  $140 \times 140$ 화소 크기의 정규화 영상을 사용하였다.

## 2.2 히스토그램의 정규화

얼굴영상은 다양한 조명이나 개인간의 피부색 차이 등으로 얼굴영역에서 명도의 히스토그램 분포가 다르게 나타난다. 이러한 명도 분포의 차이는 인식을 어렵게 하므로 그 차이를 줄이기 위한 히스토그램의 정규화 과정이 필요하다.

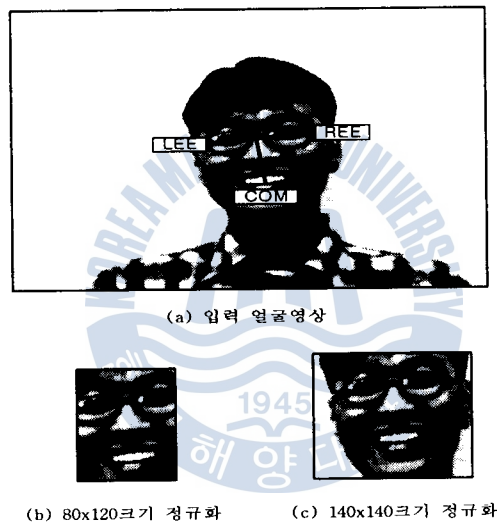


그림 2.1 크기 정규화

그리고 이러한 분포는 얼굴영상에서 구해내고자 하는 정보의 종류에 따라 각기 다르게 정규화 될 필요가 있다. 예를 들면, 얼굴영상에서 주인공정보를 얻고자 할 때는 피부색도 하나의 정보가 될 수 있기 때문에 영상의 명도 정보를 보존해야할 필요가 있는 반면, 표정 정보를 얻고자 할 때에는 가능한 한 정규화 과정에서 명도차이를 없애 주는 것이 바람직하다.

본 논문에서는 얼굴영상에서 구해내고자 하는 정보의 종류에 따라 두 가지 형태의 히스토그램 정규화과정을 적용한다. 그림 2.2에서는 이들 두 가지 형태의 히스토그램 정규화 분포와 얼굴영상에의 적용 결과를 각각 보여주고 있다. 먼저 그림 2.2(a)의 명도분포 A는  $140 \times 140$ 화소 크기의 영상에 적용함으로써 얼굴영상의 주인공 정보 및 성별 정보 추출에 사용하였다. 이는 명도 분포 A와 같은 히스토그램 균일화 방법

(histogram equalization method)에 의해 얼굴영상의 명도 대비를 높여 줌으로써 얼굴의 피부색이나 버리카락 등의 정보를 살려 내고자 함이다. 그리고 그림 2.2(c)의 명도 분포 B는  $80 \times 120$ 화소 크기의 얼굴영상에 적용하여 얼굴영상의 표정 정보 및 안경착용상태 정보를 추출하는데 사용하였다. 명도 분포 B는 인위적으로 영상의 히스토그램 명도 분포가 중앙에 모이도록 하여 피부색, 조명등에 의한 얼굴영상간의 명도 분포의 차이를 줄인다.

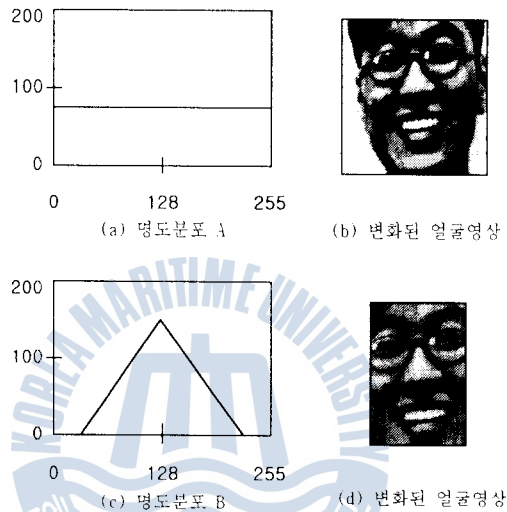


그림 2.2 히스토그램 정규화

### 3. 주요고유성분해석에 의한 얼굴정보 인식

얼굴영상에서 주요고유성분해석(principal component analysis) 방법을 이용해 얼굴영상의 고유성분을 구한 다음 그 선형조합에 의해 얼굴영상을 효율적으로 표현할 수 있다는 생각을 기반으로, Turk 등은 고유얼굴(eigenface) 개념을 도입한 얼굴인식시스템을 소개하였다. 그들은 이 방법을 통해 비교적 대규모의 얼굴영상 모집단으로부터 고유얼굴이라 불리는 얼굴의 고유성분들을 찾아낸 다음, 고유얼굴들의 공간에서 모집단 및 새로운 실험집단 얼굴영상의 주인공을 인식하고 있다. 본 논문은 이러한 고유얼굴 개념을 확장하여 얼굴영상에서 표정, 성별 및 안경착용정보 등의 다양한 얼굴정보를 추출해 내는 방법을 제안한다.

본 장에서는 주요고유성분해석 방법을 이용하여 얼굴영상에서 각종 얼굴정보의 고유성분을 추출해 내는 과정과 전반적인 얼굴정보 인식시스템의 구성에 대해 설명한다.

### 3.1 얼굴정보의 고유성분 추출

얼굴영상들로 구성된 모집단을 생각하자. 모집단에 속한 임의의 얼굴영상을 하나의 벡터  $\mathbf{X}$ 로 생각하면 식 (3-1)과 같이 나타낼 수 있다.

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad (3-1)$$

여기에서  $n$ 은 얼굴영상의 화소 수에 해당한다. 그림 3.1은 표정정보의 하나인 웃음정보를 추출하기 위한 웃음 얼굴영상들의 모집단이다. 모집단으로부터 얼굴영상  $\mathbf{X}$ 의 평균은 식 (3-2)과 같이 구해진다.

$$\mathbf{M}_X = E\{\mathbf{X}\} \quad (3-2)$$

모집단에 속한 얼굴영상들의 공분산 행렬(covariance matrix,  $\mathbf{C}_X$ )은 식 (3-3)과 같이 정의할 수 있다.

$$\mathbf{C}_X = E\{(\mathbf{X} - \mathbf{M}_X)(\mathbf{X} - \mathbf{M}_X)^T\} \quad (3-3)$$

여기에서  $\mathbf{X}$ 는  $n$ 차원의 벡터이므로,  $\mathbf{C}_X$ 는  $n \times n$ 차원의 행렬이다. 또한  $\mathbf{C}_X$ 의  $i \times i$ 번째 요소인  $c_{ii}$ 는 모집단 속에서  $\mathbf{X}$ 벡터의  $i$ 번째 요소의 분산에 해당하며,  $c_{ij}$ 는 모집단 속에서의  $\mathbf{X}$ 벡터의  $i$ 번째 요소와  $j$ 번째 요소사이의 공분산에 해당한다. 그리고  $\mathbf{C}_X$ 행렬은 일반적으로 실수의 정방행렬(square matrix)이다. 그리고  $\mathbf{X}$ 벡터의  $i$ 번째 요소와  $j$ 번째 요소사이의 상호 연관성이 없다면 그들에 의한 공분산은 0이 되므로  $c_{ij} = c_{ji} = 0$ 이 된다.

얼굴과 같은 2차원 영상데이터를 하나의 벡터로 생각하게 될 때, 그 차원이 너무 커지게 되어 현실적으로 고유벡터를 구해내는 계산이 불가능하게 되는 계산상의 문제점이 발생한다. 예를 들어 본 논문에서 다루고 있는  $80 \times 120$ 화소 크기의 얼굴영상의 경우, 이를 하나의 벡터로 생각할 때에는 9600차원이 되고 만다. 그리고 이들을 모집단으

로 구해지는 공 분산 행렬은 그 크기가  $9600 \times 9600$ 차원이 되어 현실적으로 고유벡터의 계산이 불가능하다. 이러한 고차원 벡터들에서 그 고유성분을 구해내기 위한 방법으로는 일반적으로 고차원의 벡터를 여러 개의 저차원 벡터들로 나누어 계산하는 방법이 있다<sup>[11]</sup>. 본 논문에서는  $80 \times 120$ 화소 크기의 얼굴영상의 경우 한 장의 얼굴영상을  $10 \times 10$ 화소 크기의 조각 얼굴영상 96개로 나누어 생각하였다.

각각의  $10 \times 10$ 화소 크기 조각 얼굴영상에서 구해진 공분산 행렬  $C_X$ 는 실수의 정방 행렬이기 때문에 이로부터 각 100개의 독립된 고유벡터와 고유값을 구해낼 수 있다. 여기에서 구해진 고유벡터를 고유 얼굴영상 벡터( $E_i, i = 1, 2, \dots, 100$ )라 하고, 그 고유값의 크기에 따라 식 (3-4)의 관계를 만족하도록 고유 얼굴영상 벡터를 재배열한다. 이때 각각의 조각 얼굴영상들로부터 구해진 고유 얼굴영상 벡터들은 서로 독립적으로 계산된다. 그림 3.2에서는 고유값의 크기 순서로 재배열된 상위 10장의 고유 얼굴영상 벡터를 보여주고 있으며, 이들을 웃음얼굴의 주요고유성분이라 한다.

$$\lambda_j \geq \lambda_{j+1}, \quad j = 1, 2, \dots, n-1 \quad (3-4)$$

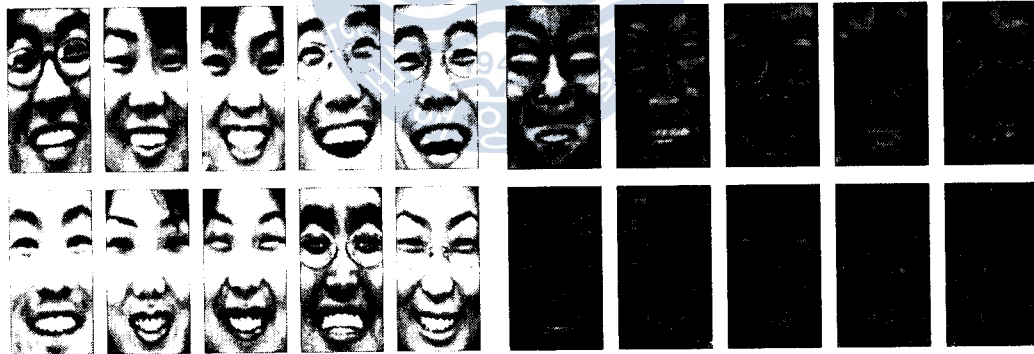


그림 3.1 웃음정보의 모집단

그림 3.2 주요고유성분 10장

### 3.2 인식시스템 구성

얼굴정보 인식을 위한 전반적인 시스템 구성을 그림 3.3에 나타내었다. 주어진 입력 얼굴영상( $X$ )에 대해 먼저 전처리단계로 정규화과정을 거친다. 그리고 분류하고자 하는 각 얼굴정보의 평균영상( $M_k$ )과의 차를 낸 다음  $l$ 개의 고유 얼굴영상 벡터( $E_k$ )에 각각 투사함으로써  $l$ 개의 거리 특징값( $w_k$ )을 식 (3-5)와 같이 구할 수 있다.



$$\omega_k = \mathbf{E}_k^T (\mathbf{X} - \mathbf{M}_X) \quad (3-5)$$

여기에서  $k=1, 2, \dots, l$  이다. 그리고 이때 구해진 입력영상의 거리 특징값들로 주어진 입력 얼굴영상의 특징값 벡터( $\mathbf{Q}^T$ )를 식 (3-6)과 같이 만들 수 있다.

$$\mathbf{Q}^T = [\omega_1, \omega_2, \dots, \omega_l] \quad (3-6)$$

그리고 주어진 입력 얼굴영상에서 그 특징값 벡터를 구하는 계산과정은 얼굴영상들간의 차와 곱이 전부이므로 실시간 인식시스템의 구현도 고려해 볼 수 있다.

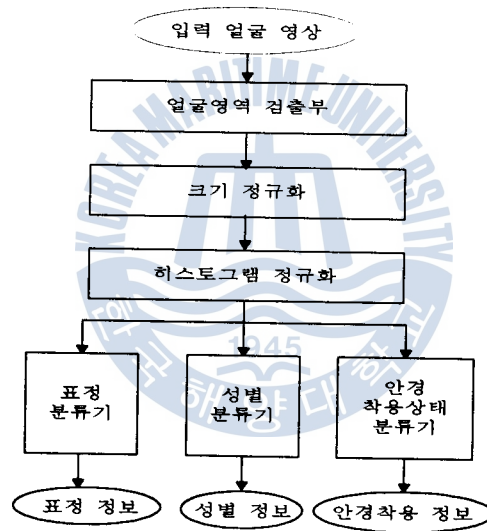


그림 3.3 얼굴 정보 인식시스템의 구성

인식과정에서는 일반적으로 패턴분류 알고리즘으로 많이 사용되고 있는 최소거리 분류방법<sup>[13]</sup>을 사용한다. 이 방법으로 입력영상을 식(3-7)과 같이 정의되는 유클리디언 거리( $\epsilon$ )를 최소화하는  $i$ 번째 부류의 집단으로 분류한다.

$$\epsilon = \| (\mathbf{Q} - \mathbf{Q}_i) \|^2 \quad (3-7)$$

이때 분류하고자 하는 얼굴의 정보는 얼굴의 주인공 정보와 함께 병렬처리로 구해지도록 하였다. 이는 인간의 얼굴정보 인식과정과 유사하다고 할 수 있는데, 심리학자들



의 연구에 따르면 인간은 얼굴의 주인공 정보 및 표정 정보를 별개의 과정으로 획득하고 있음을 보고하고 있다. Bruce 등<sup>[14]</sup>의 몇몇 학자들은 다음의 실험을 통해 그러한 과정을 설명하고 있다. 두뇌의 어떤 특정부분이 손상된 Prosopagnosia라는 병명의 환자들은 얼굴로부터 표정 정보는 일반인들과 동일하게 획득할 수가 있는 반면 주인공 정보는 얻어내지 못하고 있다. 반대로 다른 몇몇 병종의 환자들은 얼굴로부터 주인공 정보를 무난히 획득하는 반면, 표정 정보는 얻어 낼 수 없다고 한다. 이러한 사실에서 인간의 두뇌 속에서는 얼굴영상으로부터 주인공 정보와 표정 정보를 획득하는 부분이 병렬로 존재함을 알 수 있다.

## 4. 결 론

본 논문에서는 주요고유성분해석 방법을 이용하여 인간의 얼굴영상으로부터 각종 얼굴정보를 추출하는 알고리즘을 제안하였다. 기본적으로 얼굴영상에서 세 곳의 특징점, 즉 양쪽 눈의 끝점과 입의 중심점을 찾아내는 과정을 제외한 전 과정을 알고리즘화함으로써 차후 실제적인 얼굴정보 추출 및 인식시스템의 구현에 용이하도록 하였다.

인식을 위한 전처리단계에서는 구하고자 하는 얼굴정보의 종류에 따라 각각 다른 크기와 명도 히스토그램 분포를 갖도록 정규화 했으며, 특징추출단계에서는 주요고유성분해석 방법을 이용해 각 얼굴정보의 고유성분들을 구했다. 이때 고유성분의 계산과정에서 얼굴영상은  $10 \times 10$ 화소 크기의 조각들로 나누어 계산하였다. 분류단계에서는 입력된 얼굴영상을 전 단계에서 구해진 각 얼굴정보의 고유성분들에 투사하여 입력 얼굴영상이 고유성분들과 가지는 거리벡터를 구한 다음, 최소거리 분류 방법으로 입력 얼굴영상의 거리벡터를 해당 얼굴정보로 분류했다.

제안된 인식 알고리즘의 구현결과 특정한 얼굴정보를 구해내는데 소요되는 시간은 얼굴영상으로부터 세 곳의 특징점을 추출하는 과정을 제외하였을 때 Ultra Sun-Sparc 컴퓨터 상에서 1초 미만이었다. 이는 시스템이 인간의 인식과정의 병렬성을 충분히 살려 구현한 결과에 해당하며, 또한 실시간 시스템으로의 구현 가능성을 보여준다. 그리고 세가지 실험집단에 대한 인식실험 결과 평균 90% 이상의 비교적 높은 인식률을 보였다.

향후 연구되어야 할 부분으로는 영상에서 얼굴 및 그 특징점들을 자동으로 찾아내기 위한 연구를 들 수 있으며, 입력영상의 정규화 단계에서 얼굴이 어느 정도 3차원 회전량을 가지더라도 이를 정면얼굴로 매핑 시키는 과정이 필요하다. 그리고 장기적으로 대규모 실험집단에 적용되었을 때 주요고유성분 해석에 의해 발생하는 다양한 패턴 벡터로부터 유용한 판단을 내릴 수 있는 다양한 인식구조에 대한 연구 및 실험이 필요하다.

## 참 고 문 헌

- [1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *IEEE Proc.* vol. 82(5), pp. 705-740, 1995.
- [2] 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.
- [3] M. A. Shackleton, and W. J. Welsh, "Classification of Facial Feature for Recognition," *Proc. CVPR*, pp. 573-579, 1991.
- [4] Y. Yacoob, and L. Davis, "Computing Spatio-Temporal Representations of Human Faces", *Proc. CVPR*, pp. 70-75, 1994.
- [5] I. A. Essa, and A. Pentland, "Facial Expression Recognition Using a Dynamic Model and Motion Energy," *Proc. of 5th ICCV*, pp. 360-367, 1995.
- [6] S. KIMURA, and M. YACHIDA, "Facial Expression Recognition and Its Degree Estimation", *Proc. CVRP*, pp. 295-300, 1997.
- [7] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 19(7), pp. 743-756, 1997.
- [8] K. M. Lam, and H. Yan, "An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View," *IEEE Trans. Patt. Anal. Machine Intell.*, vol 20(7), pp. 673-686, 1998.
- [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.
- [11] 김선종, 구탁모, 성효경, 최홍문, "영상의 위상차를 이용한 얼굴인식," 전자공학회 논문지, vol. 35(6), pp. 81-87, 1998.
- [12] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *ECCV*, 1996.
- [13] R. O. Duda, and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, 1973.
- [14] V. Bruce, "Recognizing Faces," *London: Erlbaum*, 1988.

- [15] <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>
- [16] M. Kass, A. Witkin, and D. Terzopoulos, "Snake: Active Contour Models," *Int'l J. Computer Vision*, vol. 1, pp. 321-331, 1987.
- [17] K.M. Lam, and H. Yan, "Locating and Extracting the Eye in Human Face Images," *Pattern Recognition*, vol. 29(5), pp. 771-779, 1996.
- [18] 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.
- [19] Moghaddam, and Pentland, "Probabilistic Visual Learning for Object Detection," *Int. Conf. on Computer Vision*, pp. 786-793, 1995.
- [20] 김영익, 조석재, "K-L전개를 이용한 인간의 얼굴표정인식," 한국정보처리학회 춘계학술발표대회논문집, vol. 9, 1998.



