

# PCA와 LDA를 이용한 실시간 얼굴 검출 및 검증 기법

## (Real-time Face Detection and Verification Method using PCA and LDA)

홍은혜<sup>†</sup>      고병철<sup>\*\*</sup>      변혜란<sup>\*\*\*</sup>  
 (Eun-Hye Hong)    (Byoung-Chul Ko)    (Hyeran Byun)

**요약** 본 논문에서는 실시간 응용을 위해 형판 정합 방법을 기반으로 하면서 동시에 외형 기반 (appearance\_based) 방법에서 제시하는 학습 모델을 이용한 새로운 얼굴 검출 방법을 제안한다. 우선, 빛이나 조명의 영향에 의한 오류를 방지하기 위한 효과적인 전처리 과정으로 최소-최대 정규화(Min-max Normalization) 방법과 히스토그램 정규화 방법을 적용시킨다. 그런 뒤에 입력 영상과 형판을 PCA 변환하여 각각의 주성분(PC : Principal Component)을 생성하고 이를 LDA 변환한다. PCA 및 LDA 변환된 형판을 이용하여 입력 영상과의 거리 값을 구한 후 거리 값이 가장 작은 영역을 얼굴 영역으로 선택하고, 선택된 영역은 SVM을 이용하여 얼굴인지 아닌지를 검증하는 과정을 거친다. 또한, 본 논문에서는 실시간 얼굴 검출 방법을 위해 전체 영역이 아닌  $\pm 12$  화소 크기의 탐색 윈도우를 이용하여 시스템의 속도 및 정확도를 고려하도록 하였다. 실제 환경과 같은 6개 부류의 동영상 중심을 중심으로 실험한 결과, 본 논문에서 제안하는 방법이 기존의 PCA 변환만을 이용한 방법보다 좋은 성능을 보여줄 수 있었고, 또한 SVM을 이용한 얼굴 검증 과정을 추가한 방법이 PCA 변환과 LDA 변환을 사용한 방법보다 좋은 성능을 보여줄 수 있었다.

**키워드** : 얼굴 검출, 얼굴 검증, 형판 정합

**Abstract** In this paper, we propose a new face detection method for real-time applications. It is based on the template-matching and appearance-based method. At first, we apply Min-max normalization with histogram equalization to the input image according to the variation of intensity. By applying the PCA transform to both the input image and template, PC components are obtained and they are applied to the LDA transform. Then, we estimate the distances between the input image and template, and we select one region which has the smallest distance. SVM is used for final decision whether the candidate face region is a real face or not.

Since we detect a face region not the full region but within the  $\pm 12$  search window, our method shows a good speed and detection rate. Through the experiments with 6 category input videos, our algorithm shows the better performance than the existing methods that use only the PCA transform, and the PCA and LDA transform.

**Key words** : face detection, face verification, template-matching, PCA, LDA, SVM

### 1. 서론

지난 수년간, 심리학, 신경과학 및 여러 공학 분야에

서 얼굴 및 표정 인식 분야에 대한 연구가 수행되어왔고, 최근에는 신분 증명, 보안 시스템, 인사 자료, 회원 관리와 같은 실생활에 이를 이용하기 위한 연구가 활발히 진행중이다. 이러한 얼굴 및 표정 인식 연구의 첫 단계는 영상에서 얼굴이 어디에 위치하고 있는지를 찾아내는 얼굴 검출 단계가 필수적이고, 만약 영상에서 얼굴을 찾지 못한다면, 얼굴 인식 또한 가능하지 않게 된다. 얼굴은 얼굴 크기의 변화, 좌우의 회전 및 위 아래로 회전, 옆 얼굴과 정면 얼굴, 얼굴 표정, 그리고 빛 조건 등

· 본 논문은 학술진흥재단의 지원(2002-005-H20001)으로 연구되었음

† 비회원 : LG전자 Mobile Multimedia 연구소 연구원  
 carrot@lge.com

\*\* 비회원 : 연세대학교 컴퓨터과학과  
 soccer1@aipiri.yonsei.ac.kr

\*\*\* 종신회원 : 연세대학교 컴퓨터과학과 교수  
 hrbyun@aipiri.yonsei.ac.kr

논문접수 : 2003년 1월 7일

심사완료 : 2003년 9월 9일

에 따라 다양하게 나타날 수 있기 때문에 얼굴 검출 연구는 많은 어려움을 가진다. 따라서, 얼굴 검출 연구는 문제의 어려움과 중요성으로 인해서 얼굴 인식 분야의 전처리와 같이 수행되는 것이 아닌 하나의 독립적인 분야로 연구가 진행되고 있다[1].

얼굴 검출이란 영상에서 얼굴의 유무를 결정하고, 만약 영상 내에 얼굴이 존재한다면 얼굴의 위치와 크기를 찾아내는 것을 의미한다. 기존 얼굴 검출 방법은 지식 기반(knowledge-based) 방법, 특징 기반(feature-based) 방법, 형판 정합(template-matching) 방법, 외형 기반(appearance-based) 방법으로 나눌 수 있다.

본 논문에서는 얼굴 검출 연구에서 가장 흔히 쓰이는 형판 정합 방법을 기반으로 하면서 외형 기반 방법에서 제시하는 학습 모델을 사용한 얼굴 검출 시스템을 제안한다. 이는 각각의 방법이 가지는 근본적인 문제점을 상호 보완하여, 얼굴 검출 시 문제가 되는 몇 가지 조건에 대해 민감하지 않은 얼굴 검출 시스템을 제안하기 위함이다.

본 논문의 구성은 다음과 같다. 제2장에서는 주성분의 LDA 변환에 대해서 기술한다. 제3장에서는 얼굴 영역 검출의 전반적인 과정에 대해 기술한다. 제4장에서는 SVM 을 이용한 얼굴 검증 과정을 기술한다. 제5장에서는 실생활에서 발생할 수 있는 여러 가지 조건의 입력 영상에 대한 얼굴 검출 실험 결과를 제시하며, 끝으로 제6장에서는 결론을 맺고 향후 연구 과제에 대해 살펴본다. 그림 1은 본 논문에서 제안하는 얼굴 검출 시스템의 전체적인 개략도이다.

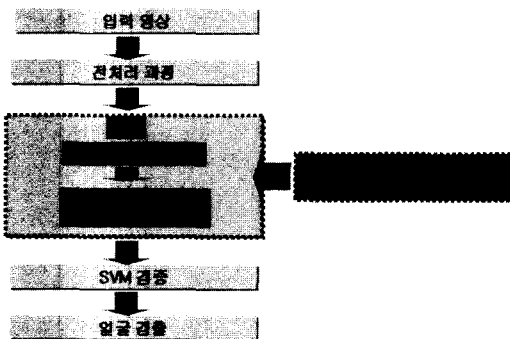


그림 1 얼굴 검출 및 검증에 대한 전체적인 개략도

## 2. 주성분의 LDA 변환

### 2.1 PCA (Principal Component Analysis)

주성분 분석(PCA : Principal Component Analysis) 변환은 신호 처리 분야, 통계학 그리고 신경망 관련 연구에서 주로 쓰인다. 이는 Karhunen-Loeve 변환 또는

Hotelling 변환이라고 불리우기도 한다. PCA 변환이란 고 차원 입력 벡터를 저 차원의 형태로 표현하는 방법인데, 이는 전체 영상의 데이터를 가지고 그것들의 분산이 큰 몇 개의 고유 방향에 대한 축으로 선형 투사시켜서 차원을 줄이는 방법을 말한다[2,3]. PCA 변환을 통해 입력 벡터의 차원을 감소하는 것은 데이터 분포에 대한 정보는 그대로 유지하지만, 계산상의 부하를 감소하고, 노이즈를 제거하며, 데이터 압축과 같은 효과를 가질 수 있도록 한다[4]. 하지만, 각 집단간의 관계를 고려하지 않고, 전체 데이터의 분산을 최대로 하는 방향으로만 선형 투사하기 때문에 집단간의 겹침이 발생하는 단점을 가진다[5]. 그림 2는 2차원 데이터를 PCA 변환을 통해서 1차원으로 감소한 것을 보여준다.

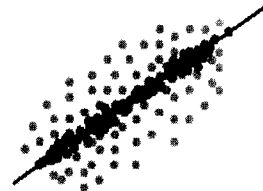


그림 2 PCA 변환

### 2.2 LDA (Linear Discriminant Analysis)

선형 판별 분석(LDA: Linear Discriminant Analysis) 변환은 집단 내 데이터 분산을 나타내는 행렬 (within class scatter matrix:  $S_w$ )과 집단 간 데이터 분산을 나타내는 행렬(between class scatter matrix:  $S_b$ )을 고려하는 방법으로, 이는 데이터에 대해 평균 근처에서 집중 발생하고, 평균 근처에서부터 주위로 감소되는 밀도를 가진다는 특정한 가정 하에 집단의 타당한 분리를 보증하는 방법이다[6,7]

이 변환 방법은 집단간의 편차는 최대로 해주면서, 집단 내 편차를 최소로 하여, 데이터를 쉽게 나눌 수 있고, 집단 또한 쉽게 분리 할 수 있도록 한다.

식 (1)은 집단 내 데이터 분산을 나타내는 식으로 여기에서  $\mu$ 는 모든 클래스의 평균값을 의미하고  $\mu_j$ 는  $j$  클래스의 평균값을 의미하며  $C$ 는 클래스의 개수를 의미한다.

$$S_b = \sum_j (\mu_j - \mu)(\mu_j - \mu)^T \quad (1)$$

식 (2)는 집단 간 데이터 분산을 나타내는 식으로 여기에서  $x_i^j$ 는  $j$  클래스의  $i$  번째 샘플을 의미하고,  $N_j$ 는  $j$  클래스의 샘플의 개수를 의미한다.

$$S_w = \sum_j \sum_i (x_i^j - \mu_j)(x_i^j - \mu_j)^T \quad (2)$$

PCA 변환이 입력 데이터를 다른 공간으로 변환할 때,

입력 데이터의 모양과 위치가 변화하고, 집단간의 겹침이 발생하는 단점을 가지는 반면에, LDA 변환은 입력 데이터의 위치가 변하는 것이 아니라, 주어진 클래스가 확실하게 구별 될 수 있도록 판별 선을 그려주기 때문에 집단간의 겹침이 발생하는 것을 막아줄 수 있다[5].

### 2.3 주성분의 LDA 변환

LDA 변환은 각 집단의 특징을 정확히 분류해주는 특성을 가지고 있기 때문에, 얼굴 관련 연구 중 특히 인식 관련 연구에서 많이 사용되어지고 있다. 하지만, 학습 집합에 없는 테스트 표본, 학습 집합과 현저히 다른 테스트 표본, 학습 집합에 있는 표본과 다른 배경을 가진 테스트 표본과 같은 경우에 대해서는 LDA 변환이 그리 좋지 않은 결과를 가짐을 알 수 있다[8]. 또한, LDA 변환은 고차원의 영상 벡터가 입력으로 들어오면 방대한 계산으로 여러 가지 어려움을 초래하는 단점을 가진다. 따라서, 이러한 단점을 해결하기 위해서 PCA 변환으로 차원을 감소한 다음 LDA 변환을 이용하는 것은 LDA 변환 만 이용하는 것보다 더 나은 결과를 가질 수 있도록 한다[8-10].

학습 데이터로부터 PCA 변환을 통해 선형 투영(linear projection) 벡터  $\phi$ 를 구하고 수식 (3)을 이용하여 입력 영상  $x$ 를 얼굴 공간  $y$ 로 매핑한다.

$$y = \phi(x - m) \quad (3)$$

구해진  $y$ 는 수식 (4)의 선형 분리 변환(linear discriminant transform :  $W_y$ )을 이용하여 분리 공간(discrimination space)  $z$ 로 다시 변환한다.

$$z = W_y^T y \quad (4)$$

본 연구에서는 얼굴 검출에 PCA 변환과 LDA 변환을 같이 사용하는 방법을 적용하였다. 이는 PCA 변환을 통해 학습 집단의 고차원 벡터의 데이터를 저차원 벡터의 형태로 변환해주고, 이 변환을 통해서 생성된 주성분을 LDA 변환하여, 원하는 수의 집단을 만들어 준다.

## 3. 얼굴 영역 검출

### 3.1 기존 연구 방법

기존의 얼굴 검출 방법은 지식 기반(knowledge-based)방법, 특징 기반(feature-based)방법, 형판 정합(template-matching)방법, 외형 기반(appearance-based)방법 등이 있다. 지식 기반 방법이란 사람의 얼굴은 두 개의 눈, 한 개의 코와 입으로 구성되어 있고, 여기의 얼굴 요소들은 일정한 거리와 위치를 가지고 있다는 것을 전제로 하여 이것들 간의 관계를 고려해서 얼굴을 검출하는 방법이다. 이는 정면 얼굴을 가진 영상에서는 얼굴 요소들의 위치와 크기를 이용하는 규칙을 이용해서 쉽게 얼굴을 찾을 수 있지만, 얼굴의 다양한 포즈나

얼굴에 기울어짐이 있는 영상에는 많은 어려움이 따른다는 단점을 가지고 있다. 특징 기반 방법이란 얼굴 요소, 색깔, 모양, 크기와 같은 얼굴 고유의 특징을 이용해서 얼굴 크기 및 위치를 추론하여 얼굴을 검출하고, 얼굴 요소의 거리나 위치 등을 통해 얼굴을 검출하는 방법이다. 이 방법은 처리 시간이 빠르고 쉽게 얼굴을 찾을 수 있는 장점을 가지고 있지만, 조명, 기울어짐, 카메라의 위치, 잡음에 따라 상당히 민감하다는 단점을 가지고 있다. 형판 정합 방법이란 얼굴에 대한 표준 형판을 만든 후, 입력 영상과 비교하여 얼굴을 검출하는 것이다. 이는 복잡한 배경에서도 수행이 가능하다는 장점을 가지고 있어서 일반적으로 많이 사용되지만, 얼굴마다 개인 차이가 크기 때문에 이를 대표할 수 있는 형판의 생성이 어렵다는 단점을 가진다. 외형 기반 방법이란 학습 영상 집합에 의해 학습된 모델을 이용해서 얼굴을 검출하는 방법이다. 이 방법은 고유 얼굴을 이용하는 방법, 신경망을 이용하는 방법, 서포트 벡터 머신을 이용하는 방법 등이 있다. 이 방법은 복잡한 영상에서 얼굴 부분을 검출하기 위해 통계학적인 수치를 이용하는 것으로서 기존의 얼굴과 얼굴이 아닌 부분의 학습 데이터 집단을 이용하여 얼굴을 찾는 것이다.

본 논문에서는 얼굴 검출 연구에서 가장 흔히 쓰이는 형판 정합 방법을 기반으로 하면서 외형 기반 방법에서 제시하는 학습 모델을 사용한 얼굴 검출 시스템을 제안한다.

### 3.2 전처리

조명의 변화는 얼굴 영역 검출 성능을 좌우하는 중요한 요인 중 하나이기 때문에, 정확한 얼굴 검출을 위해서 전처리 과정을 통해 조명을 보정하는 것이 반드시 필요하다.

본 논문에서는 전처리 과정으로 식 (5)와 같은 최소-최대 정규화(Min\_max normalization) 방법을 사용한다. 최소-최대 정규화 방법이란 입력 영상을 새롭게 정의된 데이터 범위로 변형 시켜주는 방법으로, 아주 어둡거나 밝은 입력 영상을 특정 밝기 영역 안에 존재할 수 있도록 변화시켜 줄 수 있다.

$$\hat{y} = \left( \frac{y - \min_1}{\max_1 - \min_1} \right) (\max_2 - \min_2) + \min_2 \quad (5)$$

식 (5)에서,  $\min_1$ 과  $\max_1$ 은 입력 영상의 최소, 최대 밝기 값이며,  $\min_2$ 와  $\max_2$ 은 변형하고자 하는 데이터 범위의 최소, 최대 밝기 값이다. 최소-최대 정규화를 위해, 본 연구에서는 영상에서 빛 성분의 상위 약 11.7%를 reference white(226-256)로 가정하고, 반대로 하위 약 11.7%까지를 reference black(0-30)으로 간주하였다. 여기에서 사용한 11.7%라는 수치는 실험을 통해 가장 좋은 성능을 보여준 값으로 정한 것이다. 결국, 입력 영

상의 최소, 최대 밝기 값에 상관없이 모든 영상이 31-225 범위의 밝기 값을 갖도록 하였다. 하지만, 입력 영상 밝기 값의 평균이 70 이하, 혹은 180 이상으로 영상이 너무 어둡거나, 밝은 예외적인 상황에서는 전처리로 히스토그램 평활화를 사용하였다.

본 연구에서 입력 영상은 실제 환경에서 흔히 볼 수 있는 조명의 변화를 고려하여, 일정한 밝기의 분포를 갖는 영상, 정면 광 영상, 측면 광 영상, 역광 영상, 전체적으로 높은 밝기를 갖는 영상, 전체적으로 어두운 밝기를 갖는 영상의 여섯 가지 부류로 나누었다. 실험에 사용된 영상은 <http://vip.yonsei.ac.kr/facetracking>에서 다운받을 수 있다.

그림 3은 위에서 언급한 여섯 가지 부류에 대해 최소-최대 정규화와 히스토그램 평활화를 이용해서 입력 영상의 조명을 보정한 것이다. 이 그림에서  $\mu_o$ 는 입력 영상의 밝기 평균값을 의미하고,  $\mu_p$ 는 전처리 과정을 거친 영상의 밝기 평균값을 의미한다.

표 1은 위와 같은 입력 영상에 대해 전처리 과정을 실행했을 때와 그렇지 않았을 때의 얼굴 검출율을 비교한 것으로, 전자의 경우에 얼굴 검출율은 93%이고, 후자의 경우에 얼굴 검출율은 81%로 전처리 과정을 수행하는 것의 얼굴 검출율이 높음을 알 수 있다.

**3.3 형판생성**

본 논문에서 제안하는 얼굴 검출 방법은 얼굴의 전체

표 1 전처리 유무에 따른 얼굴 검출율

|     | Non-Preprocessing(%) | Preprocessing(%) |
|-----|----------------------|------------------|
| (a) | 100                  | 100              |
| (b) | 100                  | 100              |
| (c) | 100                  | 100              |
| (d) | 78                   | 76               |
| (e) | 84                   | 89               |
| (f) | 0                    | 100              |
| 평균  | 81                   | 93               |

적인 특성을 이용하는 형판 정합 방법을 기반으로 한다. 형판은 영상을 얼굴 영역 중 눈썹부터 입술까지 40x40 크기로 잘라내어 만들어 준다. 본 연구에서 사용하는 형판(그림 4-(d))은 연세대학교 컴퓨터과학과에서 수집한 얼굴 데이터베이스(<http://vip.yonsei.ac.kr/facetracking>)를 이용하여 만들었다. 형판 생성 시, 전처리 과정으로 각 얼굴 영상에 히스토그램 평활화를 적용하였는데, 이는 어두운 부분은 더욱 어둡고 밝은 부분은 더욱 밝게 만들어 주기 때문에, 얼굴의 명암차이에 의한 변화를 줄여주고 얼굴 요소간의 특징을 부각시킬 수 있다(그림 4(b)). 또한, 본 연구에서는 실시간 영상의 얼굴 검출에 적합한 방법을 제안하기 때문에 형판을 그대로 사용하지 않고, (그림 4(c))와 같은 4x4 모자의 영상으로 변환하는 과정을 추가하였다. 이 과정은 입력 영상과 형판의 화소들이 1대1 정합할 경우 수행 속도가 느려지는 단점

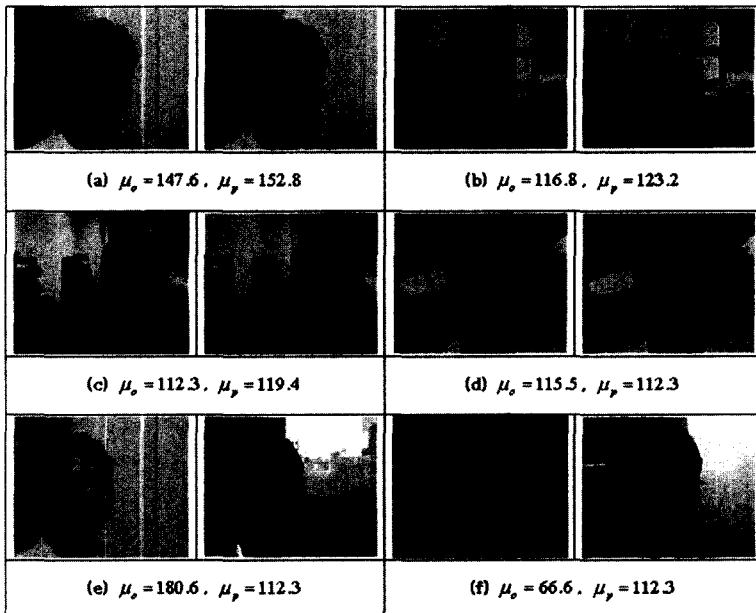


그림 3 전처리 과정에 의한 조명 보정

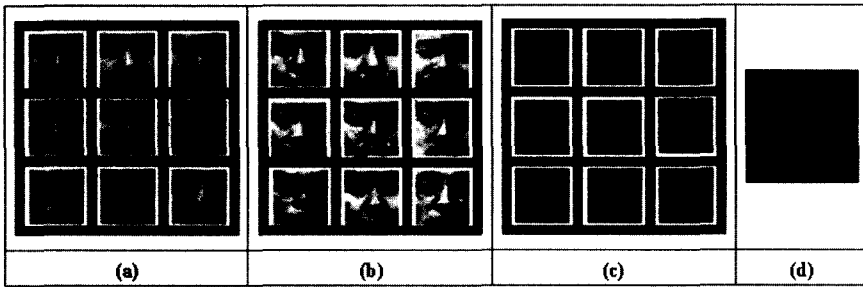


그림 4 얼굴 형판 생성 과정 : (a) 얼굴 영상, (b) 히스토그램 평활화를 적용한 영상, (c) 4x4 모자이크 영상, (d) 모자이크 평균 영상

을 보완할 수 있도록 해준다[11].

모자이크 영상이란 일정한 개수의 화소를 하나의 셀로 정의하고, 셀을 구성하는 화소 값을 하나의 값으로 재 생성하는 것을 말한다. 생성되는 셀의 화소 값을 계산하는 방법으로는 셀을 구성하는 화소의 평균값을 구하는 방법과, 셀을 구성하는 화소의 중간 값을 선택하는 방법이 있다. 후자의 방법은 셀의 화소 값을 매번 정렬한 후, 값을 선택해야 하기 때문에 실시간 처리에는 적합하지 못하다. 따라서 본 연구가 실시간 기반 방법임을 고려하여 첫 번째 방법을 사용하였다. 모자이크 영상의 생성은 얼굴의 표정, 얼굴의 개인차, 영상의 잡음, 얼굴의 작은 기울어짐 등에 대한 보정을 해주기 때문에 얼굴의 공통적인 특징을 유지할 수 있다는 장점을 가진다.

3.4 학습과정

PCA 변환에서 고유 벡터 개수는 얼굴 검출 성능에 영향을 미치기 때문에, 본 연구에서는 실험을 통해서 가장 좋은 성능을 보여주는 값으로 고유 벡터 개수를 설정하도록 하였다. 표 2에서 보면 영상의 밝기 평균값이 70이하인 어두운 영상에서는 고유 벡터 개수가 전체 고유 벡터의 40%를 가질 때, 얼굴 검출 정확율이 가장 높다는 것을 알 수 있다. 그리고, 보통 밝기 영상이나 영상의 밝기 평균값이 180이상인 밝은 영상에서는 고유 벡터 개수가 전체 고유 벡터의 30%를 가질 때, 얼굴 검출 정확율이 가장 높다는 것을 알 수 있다.

따라서 본 연구에서는 고유 벡터의 개수를 영상의 밝기 평균값을 이용하여 가변적으로 적용하도록 하였다. 즉, 입력 영상의 밝기 평균값이 70이하일 때는 전체 고

유 벡터의 40%를 가지게 하였고, 그 외의 경우에는 전체 고유 벡터의 30%를 가지게 하였다.

LDA 변환에서 각 집단의 데이터 개수는 2개 혹은 4개일 때, 가장 좋은 결과를 가진다[12]. 본 연구에서는 집단에 적합한 표본 데이터의 개수를 알아내기 위해 그림 3의 6개 부류의 입력 영상을 대상으로 실험하였는데, 표 3에서 보면 알 수 있듯이 표본 데이터의 개수가 2개일 때 가장 높은 얼굴 검출율을 보임을 알 수 있었다. 따라서, 각 집단의 데이터 개수는 2개를 사용하였고, 집단을 구성하는 데이터는 서로 다른 얼굴 영상을 사용하였다.

표 3 클래스의 표본 데이터 수에 따른 얼굴 영역 검출율

| 표본 데이터 수 | 얼굴 영역 검출 정확율(%) |
|----------|-----------------|
| 2        | 83.2            |
| 3        | 72.8            |
| 4        | 62.5            |
| 5        | 52.6            |

테스트 과정에서 얼굴 영상이 입력으로 들어오면, 얼굴 영상으로 구성된 각 집단들과 거리 값이 작으므로, 전체적인 거리 값이 작다. 하지만, 얼굴 영역이 아닌 영상이 들어오면 얼굴 영상으로 구성된 각 집단과의 거리 값이 크므로, 전체적인 거리 값이 커진다. 따라서, 여기에서 구해진 거리 값을 이용하면 얼굴을 검출할 수 있다.

3.5 얼굴 영역 매칭

형판 정합 방법을 기반으로 한 얼굴 검출 방법은 일반적으로 형판의 크기를 변화하는 멀티 스케일(multi-scale)방법과 형판의 크기를 한 개로 고정하고 입력 영상의 크기를 변화하는 다해상도(multi-resolution)방법으로 구분할 수 있다[1]. 전자의 방법을 사용하면, 다양한 크기의 형판에 대해 PCA 변환과 LDA 변환이 필요하고 입력 영상의 밝기 평균값에 따라 다른 고유 벡터

표 2 고유벡터에 따른 얼굴 검출 정확율

| 밝기 평균 \ 개수 값 | Cutoff_30 | Cutoff_40 |
|--------------|-----------|-----------|
| 70 이하        | 45        | 83        |
| 70 이상 80 이하  | 99        | 99        |
| 180 이상       | 82        | 82        |

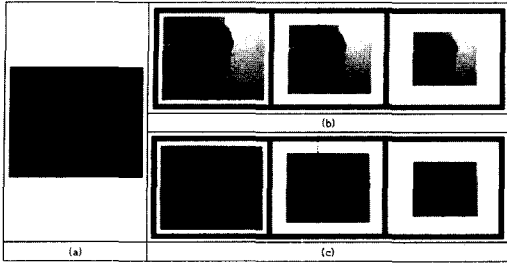


그림 5 입력 영상의 변환

(a) 입력 영상 (b) 다해상도 영상 (c) 4x4 모자이크 영상



그림 6 여러 크기 얼굴 검출



그림 7 얼굴 검출

개수를 가지는 여러 형판에 대해서도 PCA 변환과 LDA 변환이 필요하다. 이와 같은 많은 수의 형판 변환은 수행 시간이 길어지게 되므로 실시간 얼굴 검출 방법으로는 적합하지 못하다. 따라서, 본 연구에서는 후자의 방법을 사용하였다. 즉, 형판의 크기는 고정시키지만 입력 영상의 크기를 여러 단계로 변화시킴으로써 다양

한 크기의 얼굴 검출이 가능하도록 하였다. 형판은 학습 과정에서 구한 고유 백터와 고유 값을 이용해서 PCA 변환 및 LDA 변환한다. 입력 영상은 그림 5와 같이 전처리 과정 후에 크기를 90%, 70%, 50%로 줄이고 이를 모자이크 영상으로 만들어 준다.

그런 후에, 형판 크기인 40x40영역을 선택해서 PCA 변환하고 생성된 주성분을 LDA 변환한다. 형판과 입력 영상을 모두 변환 후 이것들간의 거리 값의 차이를 구한다. 줄어든 크기의 입력 영상마다 형판과의 거리 값의 차이가 가장 작은 영역을 선택한 후, 이 중에서도 가장 거리 값이 작은 영역을 얼굴로 선택하고 입력 영상의 줄어든 크기를 고려하면 실제 얼굴 영역의 크기를 알아낼 수 있다.

본 논문에서 제안한 방법을 사용할 경우, 이론상으로는 얼굴 크기가 45x45이상, 80x80이하인 얼굴을 찾을 수 있는데 형판 정합 방법이 실제 형판의 크기보다 약간 작거나 큰 얼굴에서도 얼굴을 검출할 수 있다는 장점을 가지기 때문에 실제 환경에서 35x35이상, 100x100 이하 크기를 가지는 얼굴에 대해서도 좋은 검출율을 보여줄 수 있다. 그림 6은 다양한 크기의 얼굴 검출을 보여주고 있는데 이는 고정된 크기의 형판을 가지고도 입력 영상의 크기 변화를 통해서 여러 크기의 얼굴이 정확하게 검출되는 것을 볼 수 있다. 그림 7은 그림 3과 같은 여섯 개 부류 입력 영상에 대해 PCA 변환과 LDA 변환을 이용해서 얼굴을 검출한 결과이다.

### 3.6 탐색 윈도우를 이용한 얼굴 검출

실시간 입력 영상에서 이를 구성하는 프레임은 초당 20-30개로 구성되어 있다. 이러한 입력 영상에서 얼굴이 움직임을 갖고 있더라도, 각각의 프레임을 살펴보면 얼굴 움직임의 차이는 그다지 크지 않다. 그러므로, 현재의 얼굴 주위에서 다음 프레임의 얼굴을 찾는다고 해도 이를 놓치는 경우는 많지 않다. 따라서, 얼굴 검출 시 탐색 윈도우를 이용하여 검색 영역을 줄이는 것은 얼굴 검출의 속도 및 정확도를 향상시킬 수 있다.

본 논문에서는 우선 첫 번째 프레임에서 얼굴을 찾은 후 이 영역의 중심을 기준으로 x 축과 y 축으로  $\pm 12$ 화소만큼의 영역을 탐색 윈도우로 정한다. 다음 프레임에서는 탐색 윈도우 내에서 형판과 거리 값의 차이가 가장 작은 영역을 찾는다. 얼굴 영역이 탐색 윈도우를 벗어날 정도의 빠른 움직임을 가질 수도 있으므로, 탐색 윈도우 내에서 형판과 거리 값의 차이가 가장 작은 영역이 확실한 얼굴 영역이라고 단정지을 수 없다. 따라서, 이 영역을 얼굴 영역으로 선택하기 전에 실험에 의해 구한 임계치를 적용한다. 본 논문에서의 임계치는 220이다. 즉, 이 영역과 형판과의 거리 값의 차이가 임계치보다 작으면 이는 탐색 윈도우 내 얼굴이 존재하는

것이기 때문에, 이 영역을 얼굴로 선택할 수 있다. 하지만, 이 영역과 형판과의 거리 값의 차이가 임계치보다 크다면 이는 탐색 윈도우 내에 얼굴 영역이 존재하지 않는 것이라 할 수 있기 때문에, 이때는 프레임의 전체 영역에서 얼굴을 찾아준다[그림 8].

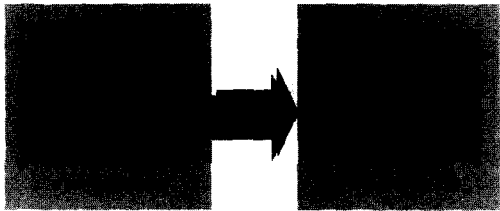


그림 8 탐색 윈도우를 이용한 얼굴 검출

### 3.7 얼굴 크기 변화

형판 정합 방법을 기반으로 얼굴 검출 시, (그림 10(a))에서 보는 것과 같이 형판의 크기가 얼굴 크기와 항상 일치한다고 할 수 없기 때문에, 형판 정합 방법만으로 얼굴의 정확한 크기를 알기란 쉽지 않다.

본 논문에서는 얼굴 영역의 정확한 크기를 알기 위해 얼굴 요소 중, 눈 혹은 눈썹의 위치를 이용한다. 우선 형판 정합 방법으로 찾아진 얼굴 영역에 얼굴 요소만을 남기기 위해 전체 화소들 중 밝기 성분의 하위 20%에 해당하는 값으로 이진화를 해주고 여기에 모폴로지 불림 연산(Morphology Dilation)을 적용하여 해당 영역 내에 남겨진 얼굴 요소를 부각시킨다.

또한, 그림 9와 같이 눈 혹은 눈썹의 위치로 얼굴의 크기를 찾아내기 때문에 본 논문에서는 눈 혹은 눈썹의 위치를 찾기 위해 이들이 가지는 다음과 같은 특징을 이용한다.

- 두 눈 혹은 두 눈썹은 수직 방향 길이 값보다 수평 방향 길이 값이 크다.
- 두 눈 혹은 두 눈썹이 얼굴 요소 중에 가장 큰 높이 값을 가진다.

따라서, 형판 정합으로 찾은 얼굴 영역에 이진화와 모폴로지 연산을 적용한 후 얼굴 요소를 수직 방향과 수평 방향으로 투영해서 수평 성분보다 수직 성분의 값이 큰 요소를 제거한다. 그리고 나서, 레이블링 방법을 이용하면 눈 혹은 눈썹 위에 있는 머리 부분 또한 제거할 수 있다. 이러한 과정을 거친 후, 중심점의 높이 값이 가장 큰 두 개의 요소를 선택하면 이것이 바로 눈 혹은 눈썹 부분 영역이고 이 때 눈 사이의 거리를 이용하면 얼굴의 정확한 크기를 알 수 있다.

그림 10(a)는 형판과 입력 영상을 PCA 변환 및 LDA 변환하여 찾은 얼굴 영역이고, 그림 10(c)는 찾은 얼굴 영역에 있는 두 눈 혹은 눈썹을 가지고 다시 얼굴

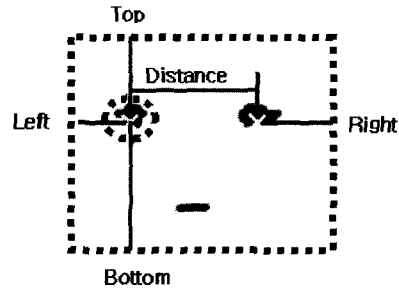


그림 9 얼굴 크기 결정



그림 10 얼굴 크기 변화의 예

- (a) PCA 변환 및 LDA 변환으로 찾은 얼굴 영역
- (b) 얼굴 요소
- (c) 변화된 얼굴 크기

을 찾은 영역이다. 그림 10(a)의 두 번째와 세 번째 얼굴을 살펴보면 얼굴 내 한쪽 눈이 정확하게 검출되지 않지만, 눈을 이용해서 얼굴을 다시 찾아주면 정확한 얼굴을 찾는 것을 알 수 있다. 즉, 눈을 이용해서 얼굴을 다시 찾은 것은 형판 정합 방법만을 이용해서 찾은 부정확한 얼굴 영역 크기를 보정해 줄 수 있으므로 입력 영상에서 정확한 얼굴 영역을 찾을 수 있다. 하지만 이진화, 모폴로지, 얼굴 성분 제거와 같은 여러 가지 과정을 거쳐야만 눈 혹은 눈썹을 찾을 수 있기 때문에 수행 속도가 느려지게 된다. 실제로 탐색윈도우를 적용하지 않은 가장 하에 얼굴 크기 변화를 거치지 않았을 때는 Pentium IV 1.8GHz 환경에서 9frame/sec의 수행속도를 갖지만, 그렇지 않을 때에는 6~7frame/sec의 수행속도를 가짐을 알 수 있다. 또한 그림 10(a)의 네 번째 영상과 같이 얼굴이 옆으로 회전된 영상에서는 눈 혹은

눈썹을 기준으로 얼굴을 다시 찾아주면 얼굴이 아닌 부분까지도 포함되는 것을 볼 수 있다. 따라서, 이러한 얼굴이 있는 입력 영상에서는 눈 혹은 눈썹을 찾아서 얼굴을 다시 찾는 과정을 포함하지 않는 것이 더욱 효과적임을 알 수 있다.

#### 4. 얼굴 검증

본 논문에서는 실시간 영상에서 얼굴을 검출하기 위해 입력 영상의 크기를 90%, 70%, 50%로 줄이고 각 크기의 입력 영상과 형판을 PCA 변환 및 LDA 변환한 후, 형판과의 거리 값의 차이를 이용하는 방법을 제안한다. 하지만, 이 방법은, 얼굴이 아닌데도 단지 형판과의 거리 값이 작아서 잘못 찾는 단점이 발생할 수 있다[그림 11].

따라서, 본 논문에서는 얼굴 검출 방법의 성능을 향상시키기 위해서 위와 같은 오류를 없애기 위해 SVM을 이용한 얼굴 검증 부분을 추가하였다.

##### 4.1 SVM (Support Vector Machine)

Vapnik이 제안한 서포트 벡터 머신(SVM : Support

Vector Machine)은 선형 머신으로 간단한 수식만을 가지고서도 복잡한 패턴 인식 문제를 쉽게 해결할 수 있다. 이는 명료한 이론적 근거에 기반하여 입력으로부터 어떠한 학습 방법을 이용하는가에 대한 직관적인 해석을 제공해주고, 간단하고 명료한 알고리즘을 통해 학습을 성공적으로 수행할 때, 영향을 미치는 요소들에 대해 규명할 수 있다. 실제 응용에서 복잡한 구조를 가지는 패턴의 분류를 위해 SVM 기법은 입력 공간인 높은 차수의 비 선형 특징 벡터 공간을 선형적으로 투영하여 해석할 수 있도록 해주고 각 특징 벡터 사이의 최적의 경계(최적분리면)를 제시한다. 여기에서 사용하는 커널 함수로는 Polynomial Machine, Radial Basis Function Machine 그리고 Two-layer Network Machine 등의 세 가지가 있다[12-14].

##### 4.2 학습과정

SVM을 이용해서 얼굴을 검증하기 위해서는, 데이터 학습 과정이 필요하다. 여기에는 얼굴 집단과 얼굴 아닌 집단 및 각 집단을 구성하는 데이터가 필요하다. 얼굴 집단 데이터는 입력 영상에서 얼굴 크기는 항상 가변적



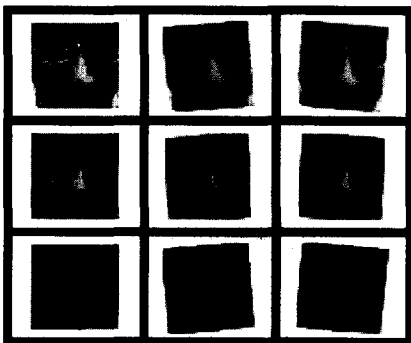
그림 11 PCA 변환 및 LDA 변환을 이용한 얼굴 검출

- (a) 원 영상의 90%크기에서 찾아진 얼굴 영상
- (b) 원 영상의 70% 크기에서 찾아진 얼굴 영상
- (c) 원 영상의 50%크기에서 찾아진 얼굴 영상
- (d) 최종 선택된 얼굴 영역

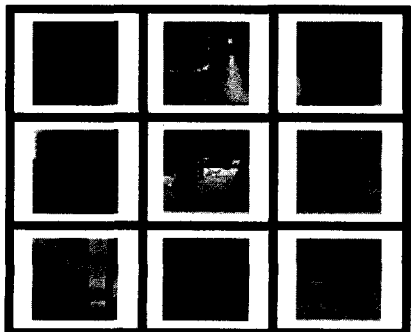


이라는 것과 영상에서 항상 정면 얼굴만 존재하는 것이 아니라 위, 아래로 돌아간 얼굴 및 옆으로 회전된 얼굴도 존재한다는 것을 고려해야 한다[15]. 따라서 얼굴 집단 학습 데이터는 여러 크기 얼굴을 정규화 한 것과 회전된 얼굴을 포함시켜야 한다.

본 연구에서 얼굴 집단 학습 데이터는 PCA 변환과 LDA 변환을 통해 얼굴로 검출된 부분을 사용하였고, 얼굴 아닌 집단을 구성하는 학습 데이터는 PCA 변환과 LDA 변환을 통해 얼굴이 아닌 부분을 얼굴이라고 잘못 찾은 부분을 사용하였다. 얼굴 집단의 학습 데이터와 얼굴 아닌 집단의 학습 데이터는 각각 300개씩을 사용했을 때 실험을 통해 가장 좋은 결과를 가짐을 알 수 있었기 때문에 본 연구에서는 각 집단을 300개의 데이터로 구성하였다. 그림 12(a)는 얼굴 집단을 구성하는 학습 데이터의 예이고 그림 12(b)는 얼굴 아닌 집단을 구성하는 학습 데이터의 예이다. 얼굴 학습 데이터는 그림 12(a)에서 보는 것과 같이 정면 얼굴 및 좌우로 5° 회전한 얼굴을 사용함으로써 여러 가지 상황의 얼굴을 검증할 수 있도록 하였다. 학습 과정 시 얼굴 영역 집단에 대해서는 타겟 값으로 +1을 주고, 얼굴 아닌 영역 집단에 대해서는 타겟 값으로 -1을 주었다.



(a) 얼굴 영역 학습 데이터



(b) 얼굴 아닌 영역 학습 데이터

그림 12 SVM 학습 데이터

#### 4.2.1 얼굴 영역 검증

입력 영상의 크기를 90%, 70%, 50% 줄인 다음 각각의 영상을 PCA 변환 및 LDA 변환한 후 형판과의 거리 차이가 가장 작은 부분을 선택하여 얼굴을 찾는 방법은 그림 23과 같이 줄어든 영상에서는 제대로 얼굴을 찾지만 결과적인 얼굴은 잘못 찾을 수 있다. 따라서, 본 연구에서는 얼굴 검출 시스템의 성능을 향상시키기 위해 SVM을 이용한 얼굴 검증 과정을 추가하였다.

입력 영상의 크기를 줄이고 이를 PCA 변환과 LDA 변환한 후에 각각의 영상에서 찾은 얼굴 영역을 얼굴 검증 과정의 후보 영역으로 선택한다. 그리고 나서 선택된 후보 영역을 40x40 크기로 정규화하고 이들을 모자이크 영상으로 만든 후에 SVM의 입력 영상으로 넣어준다. SVM 학습 과정에서 얼굴 영역의 타겟 값을 +1로 지정하고 얼굴 아닌 영역의 타겟 값을 -1로 지정하기 때문에 세 개의 영상을 SVM의 입력 영상으로 넣은 후 결과로 나온 세 개의 영상의 타겟 값을 비교한다. 이 때, 실험을 통해서 구해진 타겟 값이 임계치 이상의 값이면 얼굴 영역으로 선택하고 그렇지 않으면 얼굴이 아닌 영역으로 선택한다. 만약에 3개 입력 영상의 타겟 값이 모두 임계치를 넘지 않으면, 영상과 형판과의 차이 값을 비교한다. 형판과의 차이 값이 보통 얼굴일 때 나오던 값보다 모두 큰 경우에는 입력 영상에 얼굴이 없는 것으로 간주하고 얼굴 영역을 표시하지 않는다. 그림 13은 그림 11과 같이 PCA 변환과 LDA 변환만을 이용해서는 제대로 얼굴을 검출하지 못하지만, 여기에 SVM 과정을 추가할 때는 제대로 얼굴 검출하는 것을 보여주는 예이다.

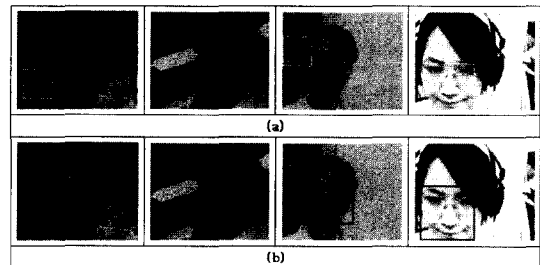


그림 13 SVM을 이용한 얼굴 검증

- (a) PCA 및 LDA 변환으로 얼굴 검출한 잘못된 얼굴 영역
- (b) SVM을 이용한 얼굴 검증해 최종 얼굴로 검증된 얼굴 영역

## 5. 실험 결과

### 5.1 실험 환경

본 논문에서 제안한 얼굴 검출 방법의 테스트 환경은 아래와 같다.

표 4 실험 환경

| 시스템 환경   |                       |                |
|----------|-----------------------|----------------|
| 운영체제     | 윈도우 2000 Professional |                |
| CPU      | Pentium IV 1.8 GHz    |                |
| 입력 장치    | LG-LPC UM10/15        | 초당 30 frame 입력 |
| 입력 영상 크기 | QCIF(176×144)         |                |
| 프로그램 언어  | MS Visual C++ 6.0     |                |

5.2 실험 결과

표 5는 PCA 변환을 이용해서 얼굴을 검출할 때와 PCA 변환과 LDA 변환을 이용해서 얼굴을 검출할 때를 비교한 표이다. 이 표를 살펴보면 PCA 변환만을 이용한 얼굴 검출 율은 약 64%이고 PCA 변환과 LDA 변환을 이용한 얼굴 검출 율은 약 81%인 것을 알 수 있다. 즉, PCA 변환과 LDA 변환을 같이 이용하여 얼굴을 검출한 것이 PCA 변환만을 이용하여 얼굴을 찾는 것보다 효과적임을 알 수 있다.

표 6은 PCA 변환과 LDA 변환을 이용하여 얼굴을 검출할 때와 거기에 SVM 검증 과정을 추가하여 얼굴 검출할 때를 비교한 표이다. 이 표에서 보면, PCA 변환과 LDA 변환을 이용한 얼굴 검출 율은 약 81%이지만, 여기에 SVM 검증 과정을 추가했을 때 얼굴 검출 율은 약 93%로 향상되었다. 따라서 얼굴 검출 시 PCA 변환과 LDA 변환만을 사용한 것보다, SVM 검증 과정을 추가한 것이 더욱 나은 얼굴 검출 율을 가짐을 알 수

표 5 PCA 변환과 PCA+LDA 변환을 이용한 얼굴 검출 율의 비교

|     | PCA | PCA+LDA |
|-----|-----|---------|
| 100 | 100 | 100     |
| 63  | 63  | 98      |
| 66  | 66  | 73      |
| 31  | 31  | 63      |
| 46  | 46  | 72      |
| 100 | 100 | 93      |
| 64  | 64  | 81      |

표 6 PCA변환과 LDA 변환을 이용한 얼굴 검출 율과 SVM을 통한 검증 과정을 추가했을 때 얼굴 검출 율 비교

|     | PCA+LDA | PCA+LDA+SVM |
|-----|---------|-------------|
| 100 | 100     | 100         |
| 98  | 98      | 100         |
| 73  | 73      | 100         |
| 63  | 63      | 76          |
| 72  | 72      | 89          |
| 91  | 91      | 100         |
| 81  | 81      | 93          |

있었다. 또한, 본 논문에서 제안한 방법이 실시간 입력 영상에 적합한 것인가의 여부를 알기 위해 본 시스템의 수행 속도를 측정했을 때, 이는 초당 약 11프레임 정도임을 알 수 있었다.

따라서, PCA변환과 LDA 변환과 SVM을 이용해서 얼굴을 검출하는 것이 가장 좋은 얼굴 검출 율을 보여 주고 또한, 실시간 얼굴 검출에도 적합하다는 것을 알 수 있었다.

6. 결론 및 향후 연구과제

본 논문에서는 기존의 얼굴 관련 연구에서 흔히 쓰이는 방법인 PCA, LDA, SVM을 이용하여 실시간 입력 영상의 얼굴 검출 및 검증에 적합하면서도 다양한 외부 환경에 민감하지 않은 새로운 얼굴 검출 알고리즘을 제안하였다.

빛이나 조명의 영향에 의한 오류를 방지하기 위해서 전처리 과정을 포함하였는데 입력 영상 밝기 성분의 평균값에 따라 다른 전처리 방법을 사용함으로써 더욱 좋은 효과를 볼 수 있었다.

얼굴 검출 시에는 형판 정합(Template Matching)방법을 기반으로 하면서도 PCA 변환 및 LDA 변환 방법을 사용함으로써 기존의 형판 정합 방법이 가지는 단점을 개선하도록 하였다. 하지만, 형판 정합의 단점이 얼굴의 크기, 모양, 회전 등에 따라 민감하고, 얼굴마다 개인 차이가 크기 때문에 얼굴을 놓치는 것인데, 이러한 단점을 개선하기 위해서 본 논문에서는 후보 영역에 대해서 SVM을 적용시킴으로써 얼굴 검출 율을 크게 향상시킬 수 있었다. 즉, 이 방법은 얼굴 영역 검출 시 흔히 사용하는 형판 정합 방법이나 외형 기반 방법보다 더 나은 결과를 가짐을 알 수 있다.

본 논문에서 제안하는 얼굴 검출 시스템은 정면 얼굴 혹은 위, 아래로 돌아간 얼굴에는 좋은 결과를 보이지만 얼굴이 옆으로 회전되어 한쪽 눈만이 보일 때 혹은 얼굴에서 한쪽 눈이 가려진 경우, 얼굴이 영상에서 너무 작을 때와 같은 경우에는 그리 좋지 않은 결과를 보인다[그림 14].

그리고 눈 혹은 눈썹 영역을 찾아서 얼굴의 정확한 영역을 다시 찾을 때는 속도의 느려짐이 발생함을 알 수 있었다. 따라서 향후 연구 과제로는 정면 얼굴에서도



그림 14 잘못 검출된 얼굴의 예

좋은 성능을 보여주지만 그 외의 경우에 대해서 좋은 성능을 보여 주도록 해야 할 것이고, 얼굴의 정확한 영역을 찾는 방법에 대해서도 성능의 개선이 요구된다.

### 참 고 문 헌

- [1] Ming-Hsuan Yang; Kriegman, D.J.; Ahuja, N., "Detecting faces in images a survey," IEEE Transactions on, Pattern Analysis and Machine Intelligence, Volume: 24 Issue: 1, pp.34~58, Jan. 2002.
- [2] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing," pp.148~156, Addison-Wesley, 1992.
- [3] <http://bi.snu.ac.kr/Courses/g-slt99/TermPrj>
- [4] <http://www.cis.hut.fi/~aapo/papers/NCS99web/node5.html>
- [5] [http://cherup.yonsei.ac.kr/leftmenu/news/biometricstudy/biometric/study5\\_3.htm](http://cherup.yonsei.ac.kr/leftmenu/news/biometricstudy/biometric/study5_3.htm)
- [6] [http://www.isip.msstate.edu/publications/reports/isip\\_internal/1998/linear\\_discrim\\_analysis/lda\\_theory.pdf/technical report](http://www.isip.msstate.edu/publications/reports/isip_internal/1998/linear_discrim_analysis/lda_theory.pdf/technical report)
- [7] Ming-Hsuan Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces : Face Recognition Using Kernel Methods," 2002. Proceeding of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp.208~213, 2002.
- [8] W. Zhao, R. Chellappa, "Discriminant Analysis of Principal Components for Face Recognition," Automatic Face and Gesture Recognition, 1998.
- [9] J. Ross Beveridge and Kai She, "Fall 2001 Update to CSU PCA Versus PCA+LDA Comparison," <http://www.cs.colostate.edu/evalfacerec/papers.html>, December, 2001 /technical report.
- [10] Jian Yang, Jing-yu Yang, "Why can LDA be performed in PCA transformed space?," Pattern Recognition 36, pp.563~566, 2003.
- [11] 고재필, "고유얼굴 기반의 얼굴 형판을 이용한 얼굴 영역 추출," 연세대학교 컴퓨터 과학과 석사학위 논문, 1998.
- [12] Nello Cristianini, John Shawe-Taylor, An Introduction to Support Vector Machines, pp.93~124, Cambridge University Press, 2000.
- [13] Simon Haykin, Neural Networks, pp.318~340, Prentice Hall, 1999.
- [14] <http://home.postech.ac.kr/~lovegod/cs499/list/6.html>
- [15] Haizhou Ai; Luhong Liang; Guangyou Xu, "Face Detection Based on Template matching and Support Vector Machines," Proceedings of International Conference on, Image Processing, Volume: 1, pp.1006~1009, 2001.



턴인식

홍 은 혜

1997년~2001년 2월 숙명여대 컴퓨터과학과 졸업(이학사). 2001년 3월~2003년 2월 연세대학교 대학원 컴퓨터과학과 졸업(공학석사). 2003년 2월~현재 LG전자 Mobile Multimedia 연구소 연구원. 관심분야는 얼굴검출, 인식, 컴퓨터 비전, 패턴



공지능

고 병 철

1998년 경기대학교 전자계산학과 졸업(이학사). 1998년~2000년 연세대학교 대학원 컴퓨터과학과 졸업(공학석사) 2000~현재는 연세대학교 대학원 컴퓨터과학과 박사과정 재학중. 관심분야는 영상검색, 비디오 인덱싱, 패턴 인식, 인



변 혜 란

1980년 연세대 수학과 학사 졸. 1987년 Univ. of Illinois at Chicago, Computer Science M.S. 1993년 Purdue University, Computer Science, Ph.D. 1994년 3월~1995년 2월 한림대학교 조교수. 1995년 3월~현재 연세대학교 컴퓨터과학과 교수. 관심분야 멀티미디어, 컴퓨터비전, 영상처리, 인공지능, 패턴인식