

고유얼굴 기반의 얼굴형판을 이용한 얼굴영역 추출 (Face Region Detection using Face Template based on Eigenfaces)

고재필[†] 변혜란^{**}
(Jaepil Ko) (Hyeran Byun)

요약 얼굴영역을 추출하기 위한 방법은 크게 얼굴의 지형적 특징추출에 기반한 방법과 얼굴형판 정합에 기반한 방법으로 분류할 수 있다. 일반적으로 복잡한 배경의 영상에서는 형판정합 방법이 우수하나, 형판의 대표성을 부여하기가 어렵다는 점이 문제시되어 왔다. 본 논문에서는 얼굴영역을 추출하기 위하여 복잡한 얼굴패턴을 몇 개의 주성분 값으로 표현할 수 있는 Hotelling 변환 과정을 이용하여 얼굴형판을 생성하고 이를 적용하여 얼굴의 크기, 영상의 명암, 얼굴의 위치에 무관하게 얼굴영역을 추출한다. 또한 휴리스틱한 임계치를 이용하여 두 사람 이상의 얼굴영역을 추출하고 기울어진 얼굴영역을 추출하기 위한 방법도 제시한다. 실험을 통하여 다양한 입력영상에 대한 추출 결과와 고유얼굴에 기반한 방법의 특징을 살펴 보았다.

Abstract There are two general approaches for face detection. One is based on the geometric feature extraction of face image and the other on a face-template matching. In general, template matching method shows better performance in a complex background. However, template matching method heavily relies on a template, which can't represent various face images. In this paper, face templates are generated through the Hotelling transform, which can compact a complex face pattern into a few principal components and the face detection system was designed to be invariant to scale, position, image intensity and to be able to detect two or more face regions by using some heuristic knowledge obtained from the test. In addition, we attempted to deal with the rotated faces in the image plane. The test showed the detection results of various input images and characteristics of the method based on eigenfaces.

1. 서론

얼굴영역 추출은 얼굴인식, 표정인식 등의 얼굴관련 연구의 핵심 기술로 많은 연구들이 수행되고 있다. 연구 초기에는 얼굴인식에 필요한 특징추출을 위한 전처리 과정으로 연구가 진행되었으나, 최근에는 얼굴영역만을 추출하기 위한 논문도 다수 발표되고 있으며 얼굴요소를 찾느냐 또는 얼굴영역 전체를 모델링 하느냐에 따라 크게 다음 두 가지로 나눌 수 있다.

① 얼굴의 지형적 특징에 기반한 방법

얼굴의 주요 구성요소인 눈, 코, 입 등을 추출하여 이들의 상관관계를 사전 지식으로 이용하여 얼굴영역을

추출하는 방법으로, 다른 물체에 의한 얼굴의 부분적인 가려짐, 얼굴의 기울어짐 등에 무관하게 얼굴영역을 추출할 수 있다. 그러나, 얼굴요소들 간의 상관관계를 규칙으로 정립하여 알고리즘화 하는 것이 어려우며, 특히 복잡한 배경에서 얼굴요소를 찾는 것은 매우 어렵기 때문에 비교적 배경이 균일한 영상에 대해서만 적용할 수 있다.

② 얼굴형판 정합에 기반한 방법

얼굴에 대한 표준적인 형판을 만들고 입력영상에 대하여 탐색윈도우를 적용하여, 각 탐색윈도우 영상을 형판과 비교하여 얼굴영역을 찾는 방법이다. 이 방법은 얼굴요소를 찾는 과정이 필요하지 않으며, 지형적 특징 추출 방법에 비해 상대적으로 복잡한 배경에 적용할 수 있다. 이러한 관점은 얼굴인식에서 형판정합과 지형적 특징 추출방법을 비교한 Brunelli에서 보여준다[1]. 그러나, 얼굴이 일부 가려지거나 얼굴의 기울어짐 등에 효과적으로 대처할 수 없다.

[†] 비회원 : 연세대학교 컴퓨터과학과
nonezero@aipiri.yonsei.ac.kr

^{**} 종신회원 : 연세대학교 컴퓨터과학과 교수
hrbyun@aipiri.yonsei.ac.kr

논문접수 : 2000년 2월 1일

심사완료 : 2000년 7월 28일

비교적 복잡한 배경에서 얼굴영역을 추출하기 위해서는 형판정합 방법을 사용하는 것이 효과적이거나 얼굴은 개인간의 차이가 크고, 개인에 있어서도 표정, 나이 등에 따른 변화가 다양하기 때문에 이들을 대표할 수 있는 형판생성이 어렵다. 얼굴형판 생성을 위한 간단한 방법은 얼굴영상들을 수집하여 이들의 평균영상을 얻는 것이나, 얼굴패턴의 복잡성에 비추어 볼 때 하나의 형판으로 대표성을 부여하기에는 부족하다.

본 논문에서는 일종의 형판정합 방법으로 하나의 형판이 갖는 취약한 대표성을 보완하기 위하여 다수의 형판을 사용하는 것과 동일한 효과를 볼 수 있는 방법을 적용한다. 이 방법은 고차원인 얼굴영상을 저차원인 얼굴공간으로의 변환 및 복원과정에 기반하여 비교하고자 하는 얼굴패턴에 가장 적합한 형판을 탐색단계에서 동적으로 생성한다. 휴리스틱한 임계치를 이용하여 두 사람 이상의 얼굴영역을 추출한다.

위에서 언급한 방법으로 구현한 얼굴영역 추출 시스템의 구성은 [그림1]과 같다. 시스템은 크게 얼굴샘플들의 고유얼굴을 이용하여 얼굴형판을 생성하는 단계와 입력영상에서 얼굴영역을 탐색하는 단계로 나뉜다. 고유얼굴은 다음 장에서 설명한다. 얼굴형판생성 단계에서는 탐색단계에서 고정크기의 윈도우영역을 입력받아 정규화 한 후 고유얼굴공간에 투영, 복원을 통해 윈도우영역에 근접한 얼굴형판을 생성한다. 탐색단계에서는 입력영상에 대해 계층적인 피라미드 영상을 만든 후 각 피라미드 영상에 대해 고정크기의 윈도우를 스캔하여 이에 해당하는 윈도우영역을 얼굴형판생성 단계에 전달하고 윈도우영역과 얼굴형판의 비교 및 휴리스틱 지식을 통해 얼굴여부를 결정한다.

디지털카메라, CCD카메라, 영화클립으로부터 다양한 영상을 수집하여 총 117장의 영상 및 카메라 앞에서 연출된 얼굴동영상을 실험에 사용하였다.

본 논문의 구성은 다음과 같다. 2장에서는 얼굴영역 추출관련 연구 소개 및 Hotelling변환에 대하여 설명하고, 3장에서는 구현한 얼굴영역 추출 시스템에 대하여 설명한다. 실험결과는 4장에서 보이고, 마지막 5장에서는 결론 및 향후연구 계획을 기술한다.

2. 배경

2.1 기존연구

얼굴영역 추출은 생물학적 인식방법 중의 하나인 얼굴인식을 위한 전처리 단계의 연구로 시작하였다. 초기 얼굴인식은 인식문제에 집중하여 인식의 대상이 되는 얼굴영역은 이미 추출되어 있거나 균일한 배경을 가정

하였다. 그러나, 자동화된 얼굴인식을 위해서는 얼굴영역 추출은 필수적이며 최근 보안관련 응용이 생김에 따라 독립적인 연구주제로 자리잡고 있다.

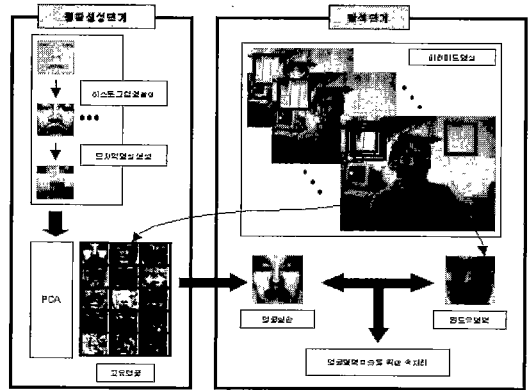


그림 1 시스템 구성

얼굴영역 추출은 얼굴인식과 마찬가지로 다음과 같은 얼굴자체의 특징 및 외적요인으로 인해 보편적인 성능을 기대하기는 어렵다[2].

- ① 눈, 코, 입과 같은 얼굴요소들의 생김새에 개인차가 존재한다.
- ② 얼굴은 비강체(Non-rigidity)운동을 통한 표정변화가 다양하다.
- ③ 나이별, 인종별 피부색 및 피부상태가 다양하다.
- ④ 헤어스타일, 안경, 화장과 같은 외적인 변화요인이 존재한다.
- ⑤ 얼굴영상은 얼굴이 3차원 물체이기 때문에 빛의 영향에 민감하다.
- ⑥ 얼굴영상은 대부분 복잡한 배경을 포함한다.

위와 같은 특성으로 인해 얼굴영역추출에 적용되는 기법들은 적절한 응용이나 처리대상을 한정짓는 것이 바람직하다. 대부분의 기존 연구는 복잡한 배경에서 수직으로 정면인 얼굴영상을 대상으로 하며, 얼굴이 눈, 코, 입과 같은 얼굴요소로 구성되어 있다는 점에 기인하여 얼굴요소들 간의 상관관계로부터 얼굴여부를 판단하거나, 또는 얼굴영상이 가진 특징을 이용하였다. 다음은 형판정합 방법 위주의 연구사례다.

(1) 얼굴요소간의 상관관계를 이용한 연구사례

Yang and Huang이 제안한 모자의 형판에 기반한 방법은 비교적 복잡한 배경에서 얼굴 및 배경을 구분하였다. 그러나, 모자의 형판으로부터 얻은 경험적 규칙은 얼굴샘플자체에 대해서도 최적의 정의라 할 수 없다[3].

신경망을 이용한 방법은 얼굴샘플내에 존재하는 얼굴오소들간의 복잡한 상관관계로부터 최적의 규칙을 찾아내어 비교적 높은 성능을 보여준다[4,5]. 두 개의 패턴을 분류하는 신경망의 학습에서 두 분류의 패턴을 끝고루 학습시키는 것은 신경망의 성능을 크게 좌우한다. 그러나, 얼굴인 경우 적절한 비얼굴영상을 확보하는 것이 쉽지 않다. 이 문제를 해결하기 위하여 Rowley는 부트스트랩 방법을 적용하였다. 이것은 초기 임의의 비얼굴영상을 확보하여 신경망을 학습한 후 얼굴이 포함되어 있지 않은 영상의 부분영상들을 신경망에 적용하여 얼굴로 잘못 인식되는 것들을 다음 번 학습시 비얼굴영상들에 첨가하는 방법이다. K.K Sung 역시 이와 유사한 방법으로 비얼굴영상을 수집하였다[2]. 이 논문에서는 영상의 픽셀수와 같은 차원의 벡터공간안에서 얼굴영상 및 배경영상이 차지하는 부분을 각각 6개의 다차원 가우시안 분포로 모델링한 후 각 분포와 입력영상과의 거리로 이루어진 12차원의 벡터를 신경망에 학습하였다. 앞의 두 경우는 비교적 높은 성능을 보여주고 있지만, 비얼굴영상 선정에 신중을 기해야 한다. 눈, 코, 입에 대해 각각의 신경망을 적용한 계층적인 방법도 소개되었다[6]. 얼굴영상만을 사용하여 얼굴을 모델링하는 방법이 Turk에 의해 소개되었는데, 본 논문에서도 이 방법을 적용하였다[7]. 2.2절에서 자세히 설명한다.

(2) 얼굴영상의 특징을 이용한 연구사례

얼굴영상에 두드러지게 나타나는 특징을 찾는 것이 중요하다. 예를 들어 눈이 있는 영역에서는 눈이 가장 어둡고 미간이 밝다는 특징은 비교적 다양한 조건하에서 변하지 않는 얼굴영상의 특징이 될 수 있다. 얼굴이 타원형이라는데 기반하여 허프변환을 응용한 Sirokey의 방법은 같은 맥락으로 파악할 수 있다[8]. 그러나, Sirokey는 비교적 배경이 단순하고 얼굴윤곽이 강한 영상들을 대상으로 하였다. 얼굴의 피부색에 대한 모델링을 통하여 얼굴영역을 추출하기도 한다[9]. 피부색 모델을 이용한 방법은 얼굴의 회전등에 무관하고 처리속도 측면에서 높은 성능을 보이나, 컬러일관성 문제로 조명 등에 매우 민감하고 피부색과 유사한 배경을 가진 영상에서는 얼굴의 경계영역을 결정하기 어렵다[10]. 이를 보완하기 위해서 스테레오 영상을 사용하기도 하며 동영상인 경우 차영상정보를 사용한다[11]. 컬러정보 사용은 스테레오정보나 시간정보가 공존할 때 유용하다.

2.2 고유얼굴을 이용한 얼굴영역 추출

고유얼굴이란 얼굴영상의 고유벡터로 얼굴영상을 얼굴공간으로 변환하기 위한 기저벡터를 말한다. Hotelling변환은 고차원의 복잡한 얼굴패턴을 이 보다

저차원의 공간(얼굴공간)에서 효율적으로 표현할 수 있게 해준다. 이 절에서는 Hotelling변환, 고유얼굴 생성 그리고 얼굴 및 배경을 구분짓는 고유얼굴의 특징을 설명한다.

2.2.1 Hotelling변환

통계적 패턴 인식에서 고려할 문제중의 하나는 특징을 추출하는 것이다. 특징추출이란 데이터공간에서 특징공간으로의 변환과정을 말한다. 보통 변환은 동일한 차원에서 이루어지나, 데이터가 가지고 있는 원래의 특징을 가지고 있으면서 효율적으로 차원을 줄일 수 있도록 설계될 수 있도록 하는데, Hotelling변환은 평균자승오차(mean-square error)측면에서 최적의 선형변환이다. Hotelling변환은 Karhunen-Loeve변환, PCA(principal component analysis)라고 불리운다[14].

차원이 n인 램뎀벡터 X를 평균이 0이라고 가정하면 즉 $E[X]=0$ 일 때, 다음과 같이 X에 대한 nxn 차원의 공분산 행렬 R은 벡터 X에 대한 외적의 기대값으로 정의된다.

$$R = E[XX^T] \tag{1}$$

공분산 행렬 R의 고유값들을 $\lambda_1, \lambda_2, \dots, \lambda_n$, 대응되는 고유벡터들을 q_1, q_2, \dots, q_n 라고 하면 다음과 같이 쓸 수 있다.

$$Rq_j = \lambda_j q_j \quad j = 1, 2, \dots, n \tag{2}$$

여기서 고유값들이 다음과 같은 순서로 정렬되어 있다고 하자.

$$\lambda_1 > \lambda_2 > \dots > \lambda_j > \dots > \lambda_n \tag{3}$$

그리고 대응되는 고유벡터를 다음과 같이 표현하자.

$$Q = [q_1, q_2, \dots, q_n] \tag{4}$$

그러면, Hotelling변환은 X의 어떤 실제벡터를 x라하고 변환행렬로 Q를 사용하여 특징벡터 y를 얻는 다음 식으로 정의된다.

$$y = [y_1, y_2, \dots, y_n]^T = [x^T q_1, x^T q_2, \dots, x^T q_n]^T = Q^T x \tag{5}$$

$$y_j = q_j^T x = x^T q_j, \quad j = 1, 2, \dots, n \tag{6}$$

x를 단위벡터 q_j 로 표현되는 주요방향으로 투영하여 얻은 결과를 y_j 라고 할 때 이를 주성분이라고 한다. 데이터 벡터 x는 특징벡터 y로부터 다음과 같이 복원 가능하다.

$$x = \sum_{j=1}^n y_j q_j = Qy \tag{7}$$

여기서 값이 큰 l개의 s만을 사용하여 복원한 x' 와 x의 오차 e는 다음 수식과 같고, q 벡터들은 서로 수직이

기 때문에 에러값은 l 을 증가시킴에 따라 급격하게 떨어진다.

$$e = x - x' = \sum_{j=1}^m \lambda_j \quad (8)$$

Hotelling 변환은 이러한 특성으로 영상처리분야에서 회전, 압축등에 사용되어져 왔고 최근얼굴인식에 도입되어 좋은 성능을 보여주고 있다[12].

2.2.2 고유얼굴

고유얼굴은 얼굴영상으로부터 구한 고유벡터 q 를 영상화했을 때 얼굴과 유사하게 나온다는데 기인한다. 얼굴영상들로부터 고유얼굴을 계산하는 것을 고려해 보자. 얼굴영상 X 의 차원이 $N \times N$ 차원인 경우, 공분산 행렬의 차원은 $N^2 \times N^2$ 이 된다. 이때 고유벡터 및 고유값은 각각 N^2 개가 존재하기 때문에 얼굴영상이 큰 경우 계산 및 저장공간의 필요량은 급격히 늘어난다. 얼굴영상의 샘플 수가 M 개일 때, $M \times M$ 차원의 행렬계산으로 변환행렬을 근사 할 수 있다. X 의 내적으로부터 고유값 λ 및 고유벡터 P 를 구한 후 다음과 같이 X 의 외적에 대한 고유벡터 행렬 Q 를 구한다.

$$Q = XP\lambda \quad (9)$$

외적에 대한 고유벡터 행렬 Q 는 M 개의 고유벡터를 가지며 λ 는 내적에 대한 것과 동일하다. [그림2]는 얼굴영상 샘플(학습영상) 및 이로부터 구한 고유얼굴을 보여준다. 고유얼굴은 얼굴의 모습과 유사하며, 왼쪽상단부터 차례대로 얼굴공간을 이루는 주요축에 해당하는 고유얼굴을 보여주고 있다.

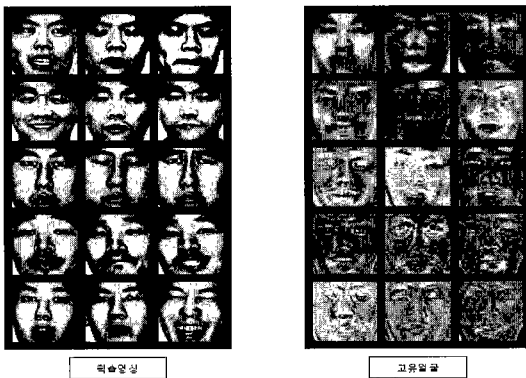


그림 2 학습영상 및 고유얼굴

2.2.3 얼굴 및 배경구분

Hotelling 변환으로부터 고유얼굴 생성에 사용한 샘플 얼굴영상은 얼굴공간에 투영한 후 다시 완벽하게 복원

할 수 있다는 것을 알 수 있다. 그러면 고유얼굴 중 일부만을 사용하여 복원할 경우 또는 고유얼굴 생성에 사용되지 않은 얼굴영상을 복원하였을 경우는 어떨까? [그림3]은 이 질문에 대한 답을 보여준다. [그림3-a]는 얼굴영상을 복원한 그림으로 왼쪽이 입력영상이고 오른쪽이 복원영상이다. 입력영상은 모두 [그림2]에서 사용되지 않은 영상으로 복원영상과 일치하지 않는다. 그러나 복원영상은 [그림2]의 학습영상 중 입력영상과 유사한 얼굴로 복원되는 것을 볼 수 있다. 얼굴이 오른쪽으로 기울어 있는 가운데 얼굴영상의 경우 복원영상은 [그림2]의 학습영상 중 정 가운데 영상과 유사함을 확인할 수 있다. 이는 고유얼굴을 통한 복원과정이 자기연상기억과 유사하기 때문에 학습의 의미로 해석할 수도 있다. 고유얼굴의 이러한 특성으로 인해 복원된 영상을 얼굴형판으로 사용할 경우 입력 얼굴영상의 변이를 반영할 수 있게 된다. 즉 얼굴의 개인차 및 표정차이에 다소 무관하게 형판의 대표성을 유지할 수 있다. [그림3-b]는 배경영상을 복원한 그림으로 배경영상을 전혀 복원하지 못하고 있으며 얼굴로도 자연스럽게 못하다.

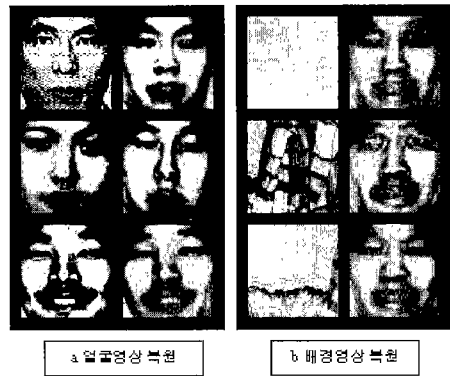


그림 3 얼굴 및 배경영상 복원

위에서 본 바와 같이 복원영상은 얼굴형판으로서 효과적이다. 수식(10)과 같이 새로운 입력영상 x 에 대하여 수식(7)로부터 복원한 x' 를 형판으로 사용하여, 이들의 차이가 임계값 이하이면 x 를 얼굴로 간주할 수 있다.

$$\|x - x'\| < \epsilon \quad (10)$$

다음 장에서는 고유얼굴의 이러한 특성을 사용하여 복잡한 배경을 가진 영상에서 얼굴영역을 추출하는 시스템을 설명한다.

3. 얼굴영역 추출 시스템

얼굴영역 추출 시스템은 2장에서 설명한 고유얼굴에 기반한 얼굴형판을 통해 수행된다. 복잡한 배경에서 얼굴의 크기, 표정에 무관한 수직 정면 얼굴영상을 대상으로 하여 다수의 얼굴을 추출하며, 좌우로 기울어진 얼굴에 대한 추출 방법도 제시한다. 시스템은 크게 고유얼굴을 생성하는 과정과 얼굴영역을 탐색하는 과정으로 나뉜다. [그림4]

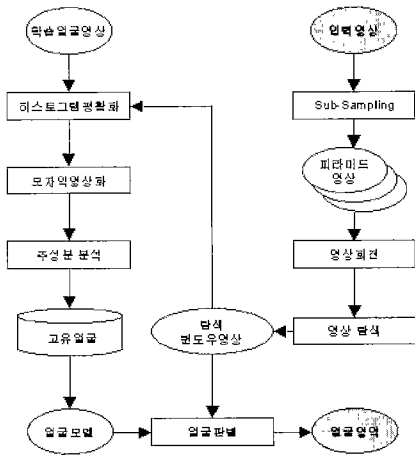


그림 4 시스템 전체 흐름도

3.1 고유얼굴 생성

이번 절에서는 고유얼굴 생성시 고려해야 할 사항을 언급한다.

3.1.1 얼굴영역 정의

얼굴영역 추출은 얼굴인식과는 달리 얼굴의 개인차를 구분하는 것이 아니라 얼굴에 존재하는 공통적인 특징을 이용하기 때문에 다양한 조건하에서도 변이가 작은 영역을 얼굴영역으로 정의하는 것이 중요하다. 이를 고려하여 본 시스템에서는 눈, 코, 입을 포함하는 최소 사각영역으로 얼굴영역을 정의하여 개인간의 차이가 큰 눈썹과 배경과 겹치게 되는 뺨의 외곽부위를 포함하지 않도록 하였다[그림5]. 이는 직관적으로 뿐 아니라 신경망을 통한 실험에 의해서도 그 타당성이 증명되었다[5]. Rowley는 얼굴영역의 각 부분에 잡음을 첨가하여 에러율을 조사하였는데, 에러율을 나타낸 그래프는 눈, 코, 입 순으로 높은 수치를 보여주었다.

3.1.2 얼굴영상 처리

명암도 차이에 따른 변화를 해결하기 위해 얼굴영역에 한정하여 히스토그램 평활화를 적용한다[그림5]. 눈



그림 5 얼굴영역



그림 6 히스토그램 평활화 영상

과 입술 영역이 더욱 어둡게 되고 볼 및 미간의 밝은 영역은 상대적으로 더욱 밝게되어 얼굴요소가 두드러져 보인다[그림6]. 또한 세 사람의 얼굴이 다소 일관성 있는 명암분포를 가지게 된 것을 볼 수 있다.

히스토그램 평활화 후 거칠어진 영상 및 개인간 또는 영상크기에 따른 얼굴요소의 위치변화에서 오는 영향을 줄이기 위해 모자이크 영상을 만든다. 4x4모자이크로 얼굴영역의 존재를 파악하고, 8x8모자이크로 얼굴요소를 구분하여 얼굴영역을 추출한 논문에 근거하여 8x8모자이크를 적용하였다.[3]. 얼굴요소의 특징이 상대적으로 빈약한 한국인의 경우 4x4모자이크로는 얼굴유무의 판단이 어려우나, 8x8모자이크는 시각적으로도 얼굴요소가 구분된다[그림7].

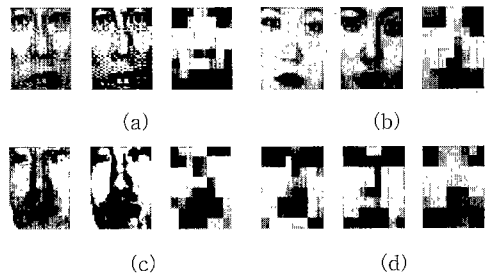


그림 7 영상처리된 얼굴영상

모자이크 영상을 통해 다음과 같은 잇점을 얻을 수 있다.

- i) 얼굴영상의 해상도 정규화
- ii) 얼굴영상의 해상도 축소를 통한 Hotelling 변환 복

잡도 감소

[그림8]은 다양한 얼굴영상에 대한 히스토그램 평활화 및 모자의 처리결과를 보여준다. (a)는 잡음이 포함되어 있으며, (b)는 왼쪽으로, (c)는 오른쪽으로 기울어져 있다. 이때 (a)(b)(c)에 무관하게 모자의 영상은 눈, 코, 입 부분에서 어둡고 볼 부분에서 밝은 명도값을 가지는 공통점을 유지한다. (d)는 정면얼굴들에 대한 모자의 예를 보여준다.



(a) 입력영상 (b) 4x4모자의 (c) 8x8모자의
그림 10 모자의 영상

3.2 얼굴탐색

입력영상에 대하여 고정크기의 탐색윈도우를 돌리고, 각 윈도우영상에 대해 3.1.2에서 설명한 방법을 적용한다. 이렇게 얻은 영상들에 대하여 2.2에서 설명한 방법에 따라 얼굴후보를 얻는다. 이 과정에서 얼굴영역의 크기 및 기울어진 얼굴에 무관하게 얼굴을 추출하기 위해 각각 피라미드 영상과 영상회전을 도입하였다. 마지막으로 경험적 규칙에 따라 얼굴후보로부터 최종 얼굴영역을 결정한다.

3.2.1 피라미드영상

얼굴영역의 크기에 무관하게 얼굴영역을 추출하기 위해 피라미드 영상을 생성하고 고정크기의 탐색윈도우를 적용하여 탐색 알고리즘의 작성 및 새로운 알고리즘 추가를 용이하게 하였다[13].

피라미드 영상은 입력영상을 단계적으로 축소한 영상들의 집합을 말한다. 본 시스템에서는 임의의 해상도를 가지는 영상을 빠르게 생성하기 위하여 인접화소 보간법을 사용하였다.

3.2.2 영상회전

고유얼굴을 이용한 방법은 수직으로 정면인 얼굴만을 추출할 수 있기 때문에, 본 시스템에서는 영상자체를 회전하여 기울어진 얼굴이 수직으로 정면이 될 수 있는 가능성을 제공하였다. 사각형의 입력영상을 임의의 사변형으로 변환함으로써 영상을 회전한다. 변환은 이계선형 변환(Bilinear)을 사용한다.(그림9)

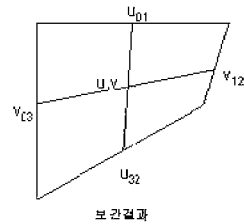
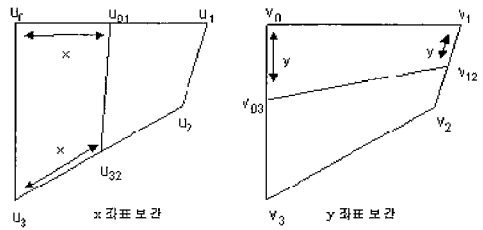


그림 9 Bilinear 변환

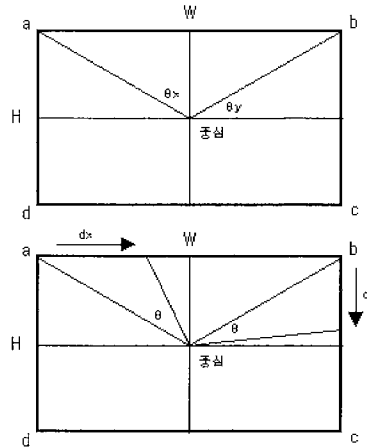


그림 10 영상회전

회전은 각 모서리 좌표를 적절히 이동하여 가능하다. 예를 들어 우 회전인 경우 각 모서리를 차례대로 다음 모서리쪽으로 이동한다(a는b로,b는c로,c는d로,d는a로이동). 일반적인 각도에 따라 이동량은 계산하기 위해 다음 수식을 정의한다([그림10] 참조).

$$dx = \frac{W}{2\theta_x} \theta, \quad dy = \frac{H}{2\theta_y} \theta$$

$$\theta_x = \tan^{-1}\left(\frac{W}{H}\right), \quad \theta_y = \tan^{-1}\left(\frac{H}{W}\right) \quad (11)$$

여기서 dx, dy는 각각 x, y의 이동화소수, W, H는 입력영상의 가로,세로 화소수, theta는 회전각도를 나타낸다. 위 방법을 통한 영상회전은 기울어진 얼굴에 대한 얼굴영

역추출의 일환으로 사용할 수 있다.

3.2.3 얼굴판별

고유얼굴을 이용한 형판을 사용하여 탐색윈도우 영역의 얼굴가능성을 구한다. 이 가능성을 0-1사이의 값으로 돌려주도록 평균자승오류에 기반 한 다음 수식을 정의하였다. 이를 고유얼굴필터라고 정의하여 필터값이 임계치 이상이면 얼굴후보로 등록한다.

$$filter = \exp\left(-\frac{f\|x-x'\|}{d}\right) > \epsilon \quad (12)$$

여기서 f 는 일반화요소, x, x' 는 각각 탐색윈도우 영역에 대한 입력패턴 및 복원패턴, d 는 x 의 차원, ϵ 은 얼굴후보를 위한 임계치를 나타낸다. 얼굴판단은 3단계를 거친다. 단계1에서는 고유필터와 휴리스틱한 지식 즉, (1)얼굴영역 근처에서 필터값이 임계치(필터값) 이상인 경우가 많으며, (2)실제 얼굴영역인 경우 이웃하는 피라미드 영상에서 임계치(필터값) 이상인 경우가 많다, 라는 점을 바탕으로 1차 후보영역 등록 및 그룹핑 정보를 수집한다. [그림11-(b)]. 단계2.에서는 단계1에서 수집한 정보를 정리하여 2차 후보영역을 등록한다[그림11-(c)]. 마지막 단계3에서는 2차 후보를 정렬하여 값이 큰 후보를 남긴다. 이렇게 하면 2명 이상의 얼굴영역을 검출할 수 있다.[그림11-(d)]

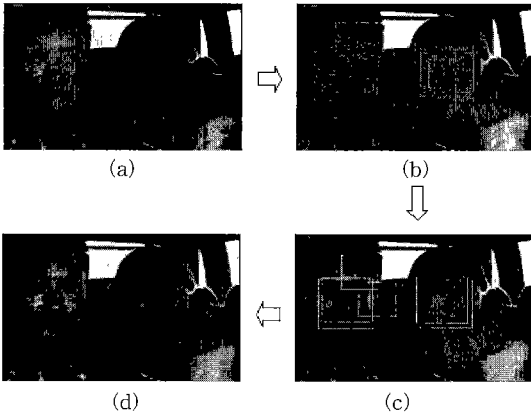


그림 16 얼굴영역 확정 과정

4. 실험결과

서로 다른 매체를 통해 획득한 179장의 영상과 동영상 사용하여 시스템의 성능과 고유얼굴에 기반한 방법의 특징을 살펴본다. 시스템은 비주얼 C++6.0으로 작성하여 PentiumIII-400 PC에서 실험하였다.

단계 1 : 1차 후보영역 등록 및 그룹핑

- (1). 1차 후보에 대한 적합도,중복횟수,위치를 저장할 배열을 초기화 한다
- (2). 피라미드 영상을 생성한다
- (3). 고정크기의 탐색윈도우를 이동한다
- (4). 탐색윈도우 영역에 대하여 모자이크처리, 히스토그램 평활화를 수행하고, 고유얼굴을 이용하여 윈도우 영상을 복원한다. 윈도우 영상 및 복원영상을 고유필터에 입력하여 필터값을 계산한다.
- (5). 필터값이 임계치 이상이면 (6)을 수행하고 아니면 (7)로 넘어간다
- (6). 피라미드 영상의 단계를 고려하여 원 영상에서 탐색 윈도우의 중심 좌표를 계산하고 이 중심 좌표를 천자로 하여 중복횟수 배열의 값을 1증가시킨다. 그리고 필터값은 적합도 배열에, 중심 좌표는 위치 배열에 저장한다(그룹핑)
- (7). 탐색 윈도우가 피라미드 영상을 모두 스캐닝 하였으면 (8)로 그렇지 않으면 (3)으로 돌아간다
- (8). 지정된 레벨만큼 피라미드 영상을 생성하였으면 단계2로 넘어가고 그렇지 않으면 (2)로 돌아간다

단계 2 : 2차 후보영역 등록

- (1). 중복횟수 배열을 내림차순으로 정렬한다.
- (2). 중복횟수를 하나씩 읽어 온다
- (3). 중복횟수가 임계치 이상이면 (4)로 넘어가고 그렇지 않으면 단계3으로 넘어간다
- (4). 찾고자 하는 사람만큼 (5)를 반복한다
- (5). 위치 배열에 있는 중심좌표에 대하여 적합도 배열에 있는 적합도에 따라 가중치를 두어 새로운 중심 좌표를 계산하고 적합도 배열에 있는 적합도의 평균을 구한다. 그리고 이 두 값을 2차 후보를 위한 배열에 저장한다

단계 3 : 얼굴영역 확정

- (1). 적합도에 따라 2차 후보영역 배열을 내림차순으로 정렬한다
- (2). 후보영역의 적합도 및 위치를 읽어온다
- (3). 현재 읽어온 점자 이후의 배열에 들어있는 후보영역에 대해 (4)를 반복한다
- (4). (2)에서 읽은 후보영역의 위치와 겹치는 다른 후보의 적합도를 0으로 설정한다
- (5). 적합도가 0이 아닌 다음 후보로 이동하여 배열의 끝이 아니면 (2)로 돌아가고 그렇지 않으면 (6)으로 넘어간다
- (6). 2차 후보영역 배열에서 적합도가 0이 아닌 후보를 모두 출력한다.(최종얼굴영역)

알고리즘 1 얼굴영역 탐색 알고리즘

4.1 실험영상 및 파라미터

다음은 실험에 사용한 영상의 특징을 정리하였다(표1).

표 1 실험에 사용한 영상 분류

	획득매체	장수	특징
A 영상	디지털카메라	52	균일한 배경이지만 명암이 다양하다 얼굴표정이 다양하다 카메라의 해상도가 높다
B 영상	CCD카메라	10	배경이 비교적 복잡하다 카메라의 해상도가 낮다
C 영상	비디오영화	117	비교적 정면 얼굴만 선정하였다 비교적 조명상태가 다양하다 배경이 다양하다
동영상	CCD카메라	200	카메라 앞에서 정면을 바라보며 자유롭게 얼굴을 움직이는 200 프레임짜리 동영상으로 얼굴의 크기, 위치, 방향이 가변적이다

얼굴영역 추출에 영향을 주는 파라미터 및 실험에서 사용한 값은 아래와 같다<표2>

표 2 실험에 사용한 파라미터 값

	값	비고
총 고유얼굴 수	22장	고유얼굴 생성을 위해 사용한 얼굴영상 수와 동일
피라미드 영상 수	[8/12장]	0.2배 또는 0.12배로 영상 축소
탐색 윈도우 이동 크기	[1,2,3,4]	피라미드 영상 크기에 의해 자동 결정
복원에 사용한 고유얼굴 수	[5,7장]	디지털, CCD(5장), 영화영상(7장)
1차 후보를 위한 임계치	[0.84 / 0.90]	사용 고유얼굴 수에 의해 실험적으로 결정
2차 후보를 위한 임계치	[3, 4개 중복]	일정범위에 속하는 1차후보의 중복 수

4.2 얼굴영역 추출 실험 1

얼굴영역 추출시간은 320x240 정지영상에 대하여 탐색윈도우의 이동크기, 찾을 얼굴의 크기 범위에 따라 가변적이긴 하지만, 평균 2-3초 가량 소요되었다. 얼굴영역 추출 결과는< 3>에서 제시하였다. 실험결과 배경이 단순한 A영상의 경우, 얼굴이 손에 가리거나 옆으로 심

하게 회전된 얼굴을 제외하고는 모두 정확하게 얼굴영역을 추출하였다. 비교적 높은 성공률을 보이는 이유는 배경이 단순한 이유도 있지만, 해상도가 높은 영상으로 피라미드 영상 생성 시 손실되는 정보가 적었기 때문이다. C영상의 경우, 비교적 정면 얼굴영상을 선정하였음에도 조명에 의한 얼굴내의 명암도 변화로 인해 성공률이 저조하였다. [그림15-(a)]는 각 얼굴영역에 대한 추출결과를 보여준다. 추출된 얼굴영역은 대체로 정확하여 3.1.1에서 정의한 얼굴영역과 일치한다. [그림15-(b)]는 기울어진 얼굴영역을 추출한 결과를 보여준다.

표 3 얼굴영역 추출 결과

	학습수	총영상수	탐색성공수	성공률
A 영상	7	52	49	94%
B 영상	3	10	9	90%
C 영상	12	117	84	72%



그림 15-(a) 정면얼굴영역 추출 예

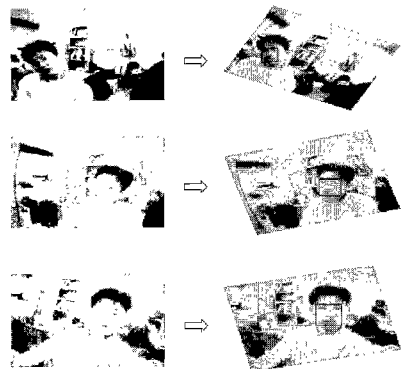


그림 15-(b) 기울어진 얼굴영역 추출 예

4.3 얼굴영역 추출 실험 2

얼굴표정에 의한 고유필터의 특성을 살펴보기 위해서 다른 표정간의 필터값(적합도)을 조사하였다. 표정간의 변화는 가장 두드러지는 표정인 웃는얼굴과 기타얼굴을 대상으로 하였다. 실험결과 표정에 따른 필터값의 변화가 미미함을 확인할 수 있다[표4,그래프1]. 이는 고유얼굴에 기반한 방법이 얼굴표정 변화에 다소 무관함을 나타낸다. 이것은 고유값이 큰 상위 몇 개의 고유얼굴을 사용하여 복원된 얼굴(형판)이 학습에 사용한 샘플들에 공통적으로 나타나는 주요 특징만을 효과적으로 담기 때문이라고 해석할 수 있다.

표 4 표정간 적합도

	기타얼굴	웃는얼굴	차이값	
얼굴영상	88.4	88.5	0.1	0.1
	91.9	91.7	-0.2	0.2
	88.4	88.8	0.4	0.4
	89.1	89.4	0.3	0.3
	89.7	90.8	1.1	1.1
	86.6	86.3	-0.3	0.3
	87.1	87	-0.1	0.1
	84.9	85.3	0.4	0.4
평균	88.26	88.48	0.21	0.36
표준편차	1.55	1.71	0.34	0.20

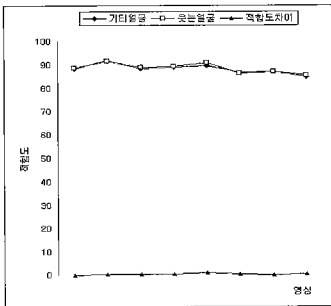


그림 12 표정간 적합도

4.4 얼굴영역 추출 실험 3

복잡한 배경에서 얼굴의 크기, 위치, 약간의 회전등을 포함한 영상에 대한 추출성능을 보기 위해 동일한 얼굴에 대한 얼굴추적 실험을 하였다. 이를 위해 1명에 대한 총 200장면이 있는 동영상에 대하여 12장을 고유얼굴을 생성하는데 사용하여 <표5>와 같은 결과를 얻었다. 얼굴이 좌,우로 회전되어 잘못 추적한 경우를 제외하면 거의 정확히 추적하였다. 적합도의 표준편차가 작다는 것은 동일인에 대한 적합도가 거의 일정하다는 것을 말하며 이는 곧 고유필터의 크기, 위치, 회전등에 대한 강

표 5 동영상에서의 얼굴추출

전체영상수	200	최저적합도	0.857
오검출	8	최고적합도	0.907
인식정확도	96%	적합도평균	0.883
		표준편차	0.0089

인함을 보여준다.

4.5 얼굴영역 추출 실험 4

다음은 학습된 얼굴에 대하여 배경이 달라졌을 경우로, 영화클립에서 배우 당 1장 즉, 총 4장의 얼굴을 학습하여 다른 장면에서의 추출 결과를 얻었다<표6>. 소규모의 인원을 대상으로 할 경우 최소한의 학습으로 높은 성공률을 얻을 수 있다는 것을 말한다.

표 6 배경변화에 따른 성공률

	성공수	실패수	합계	성공률
배우1	11	2	13	85%
배우2	5	0	5	100%
배우3	14	2	16	88%
배우4	17	2	19	89%
계	47	6	53	89%
성공률	89%			

	평균적합도	표준편차
입력영상	87.3	0.99
감마영상	87.92	1.16

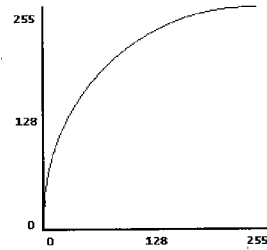


그림 13 감마곡선

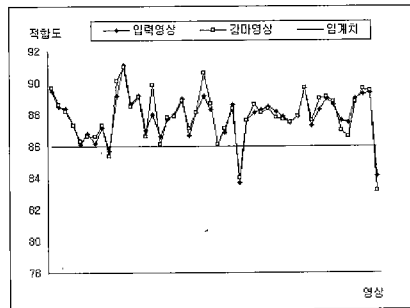


그림 14 원영상 및 감마영상에 대한 적합도 비교

얼굴영역을 추출하는데 실패한 영상은 크게 얼굴이 좌,우로 회전된 경우와 조명에 의한 얼굴내의 명도변화가 있는 영상들이었다. 명도변화에 따른 영향을 줄이기 위해 [그림13]의 감마곡선으로 영상의 명암을 변경하여 어두운 곳은 다소 밝게 밝은 곳은 다소 어둡게 영상을 변경하였으나 만족할 만한 성과를 얻지는 못했다.(평균 적합도는 감마영상이 조금 높으나, 41%만이 입력영상보다 높은 적합도를 보였다.[표6][그림14]).

5. 결론

본 논문에서는 Hotelling 변환을 이용하여 고유얼굴필터를 정의하고 이를 이용하여 등적정보 및 칼라정보 없는 비교적 복잡한 배경을 가진 정지영상에서 얼굴영역을 추출하는 시스템을 구현하고 실험을 통하여 고유얼굴필터의 특징을 살펴보았다.

시스템은 얼굴의 다양한 크기에 무관하게 하기 위해서 피라미드 영상을 만들어 이를 해결하였고 얼굴의 표정변화, 작은 기울어짐, 잡음, 영상의 명도변화에 다소 무관하게 동작하기 위해 히스토그램 평활화, 모자의 영상처리를 하였으며, 2인 이상의 얼굴영역을 추출할 수 있도록 휴리스틱한 지식을 바탕으로 한 알고리즘을 적용하였다. 또한 기울어진 얼굴영역을 찾기 위해서 영상회전을 도입하였다.

실험에 사용한 얼굴영상은 3가지 종류의 서로 다른 매체를 통해 획득하여, 실험 데이터의 다양성을 높였다. 실험을 통하여 고유얼굴필터가 소규모 집단에서, 비얼굴 영상에 대한 학습부담이 없는 효과적인 얼굴판별필터로 사용될 수 있음을 보였다. 그러나, 기울어진 영상 및 조명에 의한 얼굴변화에 민감하여 향후 이를 위한 연구가 필요하다.

참고 문헌

- [1] R.Brunelli and T.Poggio, "Face Recognition: Features versus Templates," *IEEE Trans. PAMI*, Vol.15, pp. 1042-1052, 1993
- [2] K.K.Sung and T.Poggio, "Example-Based Learning for View-Based Human Face Detection," *IEEE Trans. PAMI*, Vol 20, pp.39-51, 1998
- [3] G.Yang and T.S.Huang, "Human Face Detection in a Complex Background," *Pattern Recognition*, Vol.27, No.1, pp.53-63, 1994
- [4] Shang Hung Lin, et al, Face, "Recognition/ Detection by Probabilistic Decision Based Neural Network," *IEEE Trans. on Neural Network*, Vol. 8, No.1, pp.114-132, Jan. 1997

- [5] H.A.Rowley, S.Baluja, and T.Kanade, "Neural Network-Based Face Detection," *IEEE Trans. PAMI*, Vol 20, pp.23-38, 1998
- [6] P. Juell and R. Marsh, "A Hierarchical Neural Network for Human Face Detection," *Pattern Recognition*, Vol. 29, No. 5, pp. 781-786, 1996
- [7] Matthew A.Turk and Alex P. Pentland, "Face Recognition Using Eigenfaces," *Proc. of CVPR*, pp.586-591, 1991
- [8] S.A Sirohey, "Human Face Segmentation and Identification,," *CAT-TR-695, CS-TR-3176, CMU*, November 1993
- [9] Haiyuan Wu, Qian Chen, "Detecting Human Face in Color Images," *Proc. of IEEE*, pp.2232-2237, 1996
- [10] S.Buluswar, B.Draper, "Non-parameteric Classification of Pixels Under Varying Outdoor Illumination," *ARPA Image Understanding Workshop*, Monterey, CA, Nov.1994, pp.1619-1626
- [11] Sang-Hoon Kim, et al, "Object Oriented Face Detection Using Rang and Color Information," *Proc. of 3rd Conference on Automatic Face And Gesture Recognition*, pp.76-81, April, 1998
- [12] Pentland, B. Moghaddam, and T.Starner, "View based and modular eigenspaces for face recognition," *Proc. of IEEE Conf. On Computer Vision and Pattern Recognition*, June 1993
- [13] Richard J.Qian and T.S. Huang, "Object Detection Using Hierarchical MRF and MAP Estimation," *Proc. of CVPR*, pp.186-192, 1997
- [14] R.C.Gonzalea and R.E. Woods, *Digital Image Processing*, Addison-Wesley Publication Company, pp 148-146, 1994



고재철

1996년 연세대 전산학과 학사. 1998년 연세대 컴퓨터학과 석사. 1999년 ~ 현재 연세대 컴퓨터학과 박사과정. 관심분야는 패턴인식, 디지털방송, 영상처리

변 해 란

정보과학회논문지:소프트웨어 및 응용 제 27 권 제 1 호 참조