

# 컬러 영상에서 구성요소 융합을 이용한 얼굴 검출

이주현 이윤미 손시영 이경미<sup>o</sup>

덕성여자대학교 컴퓨터과 학부

{hyuniinara, blanchia, siyung, kmlee<sup>o</sup>}@duksung.ac.kr

## Component-fusion for face detection in color images

Ju-hyun Lee Youn-mi Lee Si-young Son Kyoung-Mi Lee<sup>o</sup>

Dept. of Computer Science, Duksung Women's University

### 요 약

본 논문에서는 컬러 영상에서 얼굴 구성요소 융합을 이용하여 얼굴 영역을 검출하는 방법을 제시한다. 먼저 광범위한 조명 환경과 인종을 포괄하는 피부색의 범위를 이용해 피부 영역을 검출하고, 영역 그룹화로 후보 얼굴 영역을 찾는다. 색 정보를 이용해 얼굴 구성요소(눈, 입)를 검출한 후, 검출된 구성요소와 구성요소 간의 관계를 융합하여 주어진 영상에서 얼굴 영역을 검출한다. 본 논문이 제안하는 구성요소 융합 방법은 구성요소 간의 관계에 대한 불확실성을 고려하고 있어, 구성요소간의 최적의 조합으로 얼굴의 크기와 포즈, 조명의 변화가 어느 정도 허용된 얼굴 검출이 가능하다.

### 1. 서 론

최근 수십 년 동안 디지털 영상으로부터 사람의 얼굴을 검출하는 문제가 지속적으로 연구되어 왔다. 얼굴 검출은 얼굴 인식 시스템의 기본 단계로, 얼굴 인식 시스템의 응용 분야는 범인 인식 시스템, 보안 시스템으로부터 신용 카드, 디지털 영화 편집 시스템, 얼굴 검색에 의한 영상의 내용 기반 검색 시스템 등을 들 수 있다. 이러한 광범위한 얼굴 인식의 적용 분야에서 볼 수 있듯이 날로 높아가는 얼굴 검출의 중요성을 실감할 수 있다. 얼굴영역 검출에 관한 다양한 접근 방법에는 신경망, 기계학습, 템플릿매칭, 휴먼한 등이 있으며 최근에는 얼굴 구성요소에 기반을 둔 검출 방법이 소개되었다.

본 논문에서는 얼굴의 기관(눈, 코, 입 등)을 찾은 후, 구성요소의 관계를 융합해 얼굴을 찾는 검출 방법을 이용하였다. 이런 구성요소에 기반한 방법은 얼굴 전체를 대상으로 한 검출 방법보다 포즈, 조명 환경, 부분적 손상에 강건하여 더 나은 검출 성능을 보이며, 검출된 얼굴 구성요소의 정보는 얼굴 인식에 적합하게 얼굴을 정규화하는데 이용될 수 있다. 또한 인체와 같이 여러 다양한 구성요소로 구성된 물체의 검출 영역에서 탁월하게 이용될 수 있다 [1].

Heisele의 논문에서는 2단계의 SVM(Support Vector Machine)을 이용하여 구성요소에 기반한 얼굴 검출 방법을 제시하였다 [1]. 첫 단계에서 얼굴 구성요소를 검출한 후, 구성요소의 조합을 이용해 얼굴을 검출한다. 이 방법은 구성요소 조합의 최대값만을 고려함으로써 영상에서 하나의 얼굴만 검출된다는 단점이 있다. Xie의 논문[3]에서는 얼굴 구성요소들을 학습시킨 AdaBoosting을 이용하여 구성요소를 검출하고 구성요소들이 얼굴모델에 기반하여 융합된다. 그러나 시간적 효율때문에 구성요소간의 관계에 대해서는 고려하지 않음으로써 동영상처럼 비슷한 포즈를 취한 얼굴을 담은 연속영상에서만 추적이 가능하다.

본 논문에서 제안하는 방법은 피부색 영역을 결정한 후, 검출된 피부색 영역을 연결 요소로 그룹화하여 불필요한

배경이나 얼굴 이외의 부분을 미리 제거하여 후보 영역을 결정한다. 그리고 색 정보를 이용해 얼굴 구성요소를 검출하고 구성요소와 구성요소간의 관계를 이용한 얼굴 모델을 이용하여 구성요소간의 관계가 함께 융합되어 얼굴이 검출된다. 이 방법은 색 정보를 이용한 얼굴 구성요소 검출로 다수의 얼굴이 포함된 영상에서도 얼굴이 검출되며 구성요소간의 최적의 조합으로 얼굴의 크기와 포즈, 조명의 변화가 어느 정도 허용된 얼굴 검출이 가능하다.

### 2. 구성요소에 기반한 얼굴 모델

우선, 얼굴의 구성요소에 대한 불확실성과 구성요소간의 관계에 대한 불확실성을 가진 확률적 얼굴 모델을 정의한다. 이러한 불확실성을 지닌 얼굴 모델은 보다 다양한 얼굴을 표현할 수 있는 장점이 있다 [3]. 얼굴은 얼굴 구성요소를 포함하며 구성요소간의 관계가 만족되어야 한다. 따라서 얼굴 모델은 다음과 같이 정의된다.

$$F=(C_i, R_j), \quad i=1, \dots, N, \quad j=1, \dots, M \quad (1)$$

여기서  $C_i$ 는 얼굴의 각 구성요소,  $N$ 은 구성요소의 개수,  $R_j$ 는 구성요소간의 관계, 그리고  $M$ 은 관계의 개수를 나타낸다.

구성요소에 기반한 얼굴검출 방법들은 어떻게 얼굴의 구성요소들의 집합을 찾아내느냐를 풀어야만 한다. 구성요소들의 조합을 학습시키거나, 대용량의 데이터에서 구성요소들이 이루는 특징을 찾아내서 사용할 수도 있다. 본 논문에서는 얼굴 구성요소로 입술과 두 개의 눈을 이용하였는데, 구성요소에 대한 정보는 영상데이터에서 통계적인 방법으로 취득되었다. 이 때, 이용된 값은 다음과 같다.

- 각 눈의 평균  $C_b, C_r$
- 입술의 평균  $C_b, C_r$
- 각 눈과 입술사이의 거리  
두 눈사이의 거리

- 세 요소가 이루는 삼각형의 세 내각

위의 값에 대한 평균( $\mu$ )과 분산( $\sigma$ )을 계산하였다.

### 3. 구성요소 융합을 이용한 얼굴 검출

#### 3.1 피부색 검출 및 후보 얼굴 영역 검출

컬러 영상(그림 1)이 주어지면 미디언 필터를 이용하여 RGB 색공간에서 평활화를 시킨다. 다음은, 적용된 피부 컬러 범위를 이용하여 피부색 영역을 결정할 수 있다. 논문 [4]는 YCbCr 색공간에서의 피부색 영역이 HSV 색공간에서의 피부색 영역보다 조밀하다고 보고했다. 따라서 우리는 YCbCr 색공간에서 YCbCr의 범위를 이용해서 피부색 영역을 결정하였다. 그림 2는 그림 1에서 피부영역을 검출한 결과 영상이다.

피부색이 검출된 영상에 대해, 후보 얼굴 영역을 결정하기 위해서 피부색인 픽셀의 4-인접 픽셀 중 피부색인 픽셀을 연결 요소로 정한다. 그룹화를 마치면, 주어진 영상에 피부 연결 영역들의 리스트가 만들어지고, 얼굴 후보 영역을 결정하기 위하여 조건에 따라 얼굴일 확률이 적은 불필요한 영역을 제거한다 [7]. 그림 3은 검출된 피부색 영역에서 찾아진 후보 얼굴 영역을 보여주고 있다.



그림 1 컬러영상 그림 2 피부색 영역 그림 3 후보 얼굴 영역

#### 3.2 구성요소 검출

후보 얼굴 영역에 대해 Hsu *et. al*의 논문에서 사용된 눈과 입술의 색정보(YCbCr)를 이용한 EyeMap과 MouthMap을 적용하여 눈과 입술을 찾는다 [2]. 그들은 눈은 Cb가 높고 Cr이 낮으며 눈동자가 어둡고 밝은 픽셀 값을 갖는다고 보고했다. 식 (2), (3)에 의한 두 EyeMap의 실험 결과를 AND한 후, 픽셀 값이 눈의 밝기를 만족하면 눈으로 검출한다. 이 때, AND를 적용한 영상에 대해서는 반드시 팽창과 마스킹을 통해 눈 영역을 강조함으로써 구성요소 검출율을 높인다 (그림 4(a)).

$$EyeMapC = \frac{1}{3} \{ (Cb^2) + (255 - Cr)^2 + (Cb/Cr) \} \quad (2)$$

$$EyeMapL = \frac{Y^+(x, y)}{Y^-(x, y) + 1} \quad (3)$$

여기서  $Y^+$ 는 팽창,  $Y^-$ 는 침식을 나타낸다. 그리고 눈은 일반적으로 가로가 세로보다 길이가 길다는 점을 이용해 가로/세로의 비율이 눈의 조건에 맞지 않으면 제외시킨다 (그림 4(b)).

입술은 Cr이 Cb보다 더 큰 값을 갖는다는 점에서 식 (4)을 적용하고 (그림 4(c)), 픽셀 값이 입술의 밝기를 만족하면 입술로 검출한다 (그림 4(d)).

$$MouthMap = Cr^2 \cdot (Cr^2 - \eta \cdot Cr/Cb)^2 \quad (4)$$

여기서  $\eta$ 는 평균  $Cr/Cb$ 에 대한  $\sigma^2$ 의 비율을 나타낸다.

그림 4는 후보 얼굴 영역에서 검출된 구성요소들을 보여주고 있다. 그러나 색 정보만을 이용하여 구성요소를 검출하는 것은 눈, 입이 아닌 부분도 구성요소로 검출되어 불충분하므로 구성요소간의 적절한 융합을 통해 얼굴을 검출하는 것이 필요하다.



그림 4 얼굴 구성요소 검출: (a) 식 (2)과 (3)를 적용한 후 AND한 결과, (b) 눈 후보 검출, (c) 식 (4)을 적용한 결과, (d) 입술 후보 검출

#### 3.3 얼굴 검출

이 단계에서는 3.2장에서 검출된 얼굴 구성요소를 구성요소의 색 정보와 얼굴의 기하학적 조건을 이용하여 최적의 얼굴 조합을 찾아 얼굴을 검출한다. 구성요소 융합 시에 검출된 모든 요소에 대한 얼굴 조합을 고려하는 것 보다는 동일 후보영역 내의 눈 후보와 입술 후보에 대해서만 계산하는 것이 얼굴 검출에 보다 타당하며 시간-효율적이다.

2장에서 정의한 구성요소와 그들 간의 관계에 대한 확률적 얼굴 모델(식 (1))을 이용하는 것은 구성요소간의 관계를 융합함에 있어 어느 한 조건이 만족되지 않아 얼굴임에도 얼굴로 검출되지 못하는 것을 방지하기 위함이다. 이렇게 하면 모든 구성요소 및 관계를 고려할 수 있어 다수의 얼굴이나 기울어진 얼굴을 포함한 영상에서도 최적의 얼굴 조합을 찾아낼 수 있다.

확률적 얼굴 모델에 대한 얼굴 가능성(face score)을 다음과 같이 계산한다.

$$FaceScore = \prod_{i=1}^N \alpha_i L(C_i) + \prod_{j=1}^M \beta_j L(R_j) \quad (5)$$

여기서  $L()$ 는 확률적 모델과의 유사도(likelihood)를,  $\alpha_i$ 와  $\beta_j$ 는 모델내에서의 중요도를 나타낸다.

$$L(A) = \frac{1}{|A - \mu_A| + 1}, \quad \alpha_i = \frac{1}{\sigma_i}, \quad \beta_j = \frac{1}{\sigma_j}$$

여기서  $A$ 는 개별적인 구성요소( $C_i$ )나 관계( $R_j$ )를,  $\mu_A$ 는  $A$ 에 대한 평균을,  $\sigma_i$ 는  $C_i$ 에 대한 분산,  $\sigma_j$ 는  $R_j$ 에 대한 분산을 나타낸다. 그림 5는 식 (5)을 이용하여 검출된 결과를 보여준다.



그림 5 검출된 얼굴 영상



그림 6 PC카메라에 의해 획득된 영상에 대한 실험 결과: 제안된 방법은 기울어진 얼굴이나 촬영거리에 의한 다양한 크기의 얼굴, 여러 얼굴이 포함된 경우에도 성공적으로 얼굴을 검출하였다.



그림 7 얼굴 검출 결과: 각 영상은 여러 얼굴들을 포함하고 있다.

4. 실험 결과 및 결론

제안된 알고리즘은 운영체제 Windows 2000XP 상에서 JAVA를 이용하여 구현되었으며, 실험은 Pentium-IV 1.8GHz인 CPU와 512 MB 메모리 사양의 Windows 2000 XP 환경에서 수행하였다. 실험에 사용된 영상은 AnyCam Player2.0에서 획득한 영상, Miss Korea 데이터베이스 [5], champion 데이터베이스 [6], 인터넷에서 얻은 영상 등을 사용하였다. 영상은 다양한 얼굴 크기, 자세, 사람 수, 조명 환경 등을 포함한다.

표 1 제안된 방법을 이용한 검출 결과  
FP (False Positive): 잘못 검출된 영상

데이터베이스	Miss Korea	Champion
영상 개수	232	268
영상 크기(화소)	350×370	150×200
FP 개수	19	36
검출율 (%)	91.81%	86.57%
수행시간	3.64±1.05	1.22±0.5

실험 결과는 얼굴 크기, 포즈, 각도의 변화에도 강건하게 얼굴이 검출되었다 (그림 6). 여러 얼굴이 포함된 영상에서의 검출은 보다 어려운 문제인데 그림 7은 좋은 검출 결과를 보여주고 있다. 표1은 Miss Korea 데이터베이스와 Champion 데이터베이스를 이용하여 획득한 결과이다. 여기서 검출율은 잘 검출된 영상의 개수를 전체 영상의 수로 나눈 수치이다. 잘못 검출된 경우(FP)는 대부분 조명의 영향으로 눈, 입의 색이 손상되거나 머리칼과 눈이 붙어 있어 얼굴 구성요소(눈, 입)가 검출되지 않는 경우이다. 또한, 2장에서 정의한 통계적 얼굴 모델과의 차이가 커서 융합이 적절히 되지 않은 경우도 있었다. 검출 시간은 주어진 영상에서 피부색 검출부터 구성요소를 기반한 얼굴 검출까지이다. Champion 데이터베이스인 경우는 Hough변환을 사용하는 Hsu et. al [2] 의 5.78±4.98보다 최고 6.26배 빨랐다.

본 논문에서 제안한 방법은 얼굴 구성요소 융합을 이용하여 얼굴 영역을 검출하는 것으로 다양한 얼굴 크기, 각도, 포즈의 변화에도 얼굴이 검출되었다. 그러나 조명 보정과 좀 더 광범위한 얼굴 데이터를 통해 얼굴 모델을 정의하여 얼굴 검출율을 높이는 노력이 필요하다고 본다.

Acknowledgements

본 연구는 한국과학재단 목적기초연구 (R04-2003-000-10092-0) 지원으로 수행되었음.

참고문헌

- [1] B. Heisele, T. Serre, M. Pontil and T. Poggio, "Component-based face detection," *Proceedings of 2001 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001)*, vol. 1, pp. 657-662, 2001.
- [2] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, 2002.
- [3] B. Xie, D. Comaniciu, V. Ramesh, M. Simon and T. Boult, "Component fusion for face detection in the presence of heteroscedastic noise," *Proceedings of Annual Conf. of the German Society for Pattern Recognition (DAGM'03)*, pp. 434-441, 2003.
- [4] C. Garcia and G. Tziritis, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," *IEEE Tran. on Multimedia*, vol. 1, no. 3, pp. 264-276, 1999.
- [5] Miss Korea 데이터베이스, <http://misskorea.hankooki.com/>.
- [6] The Champion database, [http://www.libfind.unl.edu/alumni/events/breakfast\\_for\\_champions.htm](http://www.libfind.unl.edu/alumni/events/breakfast_for_champions.htm).
- [7] 이주현, 이경미, "얼굴 영역 검출을 위한 이질적 템플릿 매칭의 융합", *제 16회 영상처리 및 이해에 관한 워크샵 발표 논문집*, pp. 391-396, 2004