

논문 2004-41C1-2-13

# 표정분석을 위한 얼굴 구성 요소 검출

## (Detection of Face-element for Facial Analysis)

이 철 희\*, 문 성 룡\*\*

(Cheul-Hee Lee and Sung-Ryong Moon)

## 요 약

미디어의 발달에 따라 다양한 정보가 미디어에 실리게 되는데 이 중에서 표정은 흥미있는 정보중에 하나이다. 표정에는 인간 내면의 의중이 포함되어 있기 때문이다. 내면의 상태는 표정 뿐만 아니라 제스처로 나타나기도 하지만 중요도는 표정이 높다고 할 수 있다. 이 표정은 임의로 조작할 수도 있지만 대체로 내면의 의중을 포함한다. 또한 표정은 사람마다 독특한 개성을 가지고 있지만 대체로 일정 구분가능한 공통점 또한 가지고 있다. 본 논문에서는 USB CCD카메라 환경의 동영상에서 표정을 분석하기 위해서 얼굴의 구성 요소를 검출 하고자 한다. 얼굴 구성요소에는 사람마다의 공통적인 표정 변화에 따른 특징점이 분포하기 때문이다. 구성 요소 검출을 위해서 먼저 동영상에서 한 프레임을 캡처하여 얼굴의 위치를 파악하고 얼굴 영역을 분리한 다음 특징 점을 검출하게 된다.

## Abstract

According to development of media, various information is recorded in media, expression is one during interesting information. Because expression includes of relationship of human inside. Intention of inside is expressed by gesture, but expression has more information. And, expression can manufacture voluntarily, include plan of inside on the man. Also, expression has unique character in a person, have alliance that do division possibility. In this paper, to analyze expression of USB camera animation, wish to detect facial building block. Because characteristic point by person's expression change exists on face component. For component detection, in animation one frame with Capture, grasp facial position, and separate face area, and detect characteristic points of face component.

keywords : face component search, facial detection, face detection

## I. 서 론

멀티미디어의 발달에 따라 다양한 정보가 미디어에 실리게 되는데 인간의 표정 정보는 흥미 있는 정보중 하나이다. 미디어의 종류도 다양해져서 다양한 장치로 부 터 정보를 가져올 수 있다. 표정에는 인간의 내면의 감정이 나타나고, 얼굴에 집중되어 있기 때문에 인지하는데

집중도를 가진다. 다른 감정의 표출 수단으로 제스처나 음성이 있지만 이것들은 각 각 다른 난이도를 가지고 있다. 표정 인식은 다양한 방법으로 시도되는데 두가지로 구분할 수 있다. 근육의 움직임을 직접 센서로 측정하여 얻는 방법과 영상처리를 이용하는 방법이다.

전자는 초창기에 시도되었고, 후자는 연구의 진행 에 따라 시도되었다. Ekman은 FACS( Facial Action Coding Sytem)을 만들어 얼굴의 움직임을 46개의 AU(Action Unit)로 표정을 만들어 내었다<sup>[1]</sup>. Bernd 와 Thomas는 두 단계를 거쳐서 얼굴의 구성요소를 추출했는데, 먼저 얼굴의 필수 구성요소인 눈, 코, 입에 기초한

\* 학생회원, 원광대학교 전자공학과  
(Dept. of Electronic Eng., Wonkwang Univ.)

\*\* 정회원, 원광대학교 전자공학과  
(Dept. of Electronic Eng., Wonkwang Univ.)  
접수일자: 2003년3월19일, 수정완료일: 2004년2월16일

형판으로 구성요소들을 기하학적 분류기를 이용하여 찾고, 각 구성 요소를 선형 분리기를 이용하여 찾았으며 구성 요소의 회전과 깊이 변화를 학습을 통하여 보정하였다.<sup>[3]</sup> 이러한 연구들은 정지영상에서 이루어 졌으며, 미디어의 발달에 따라 동영상에서의 표정인식에 대한 욕구가 증대되고 있다. 본 논문에서는 표정인식의 전단계로 화상 카메라로부터 들어오는 영상에서 얼굴의 구성요소에 분포하는 특징점을 찾고자 한다. 전체 과정은 다음과 같다. 카메라 영상에서 피부색을 이용하여 찾은 얼굴 영역의 무게 중심을 찾고, 무게 중심점을 이용하여 얼굴 구성 요소의 기하학적인 특징을 이용하여 각각 구성 요소를 찾고, 찾은 구성 요소에서 표정 특징 점을 찾아 표시한다.

### II. 표정 특징 점의 구성

표정을 인식하기 위하여 우선 MPEG4에 기술된 FAP (Facial Animation Parameter)에 대해서 살펴보면 다음과 같다. FAP는 애니메이션에서 표정을 합성하기 위해서 사용하는 점을 모아놓은 국제 표준이다. 이 점들은 MPEG4에 표시되어 있는데 MPEG4에는 FDP(Face Definition Point)와 FAP를 기술했다. FAP는 FDP에 포함되어 있다. 그림 1은 FDP와 FAP를 표시하였다. 그림

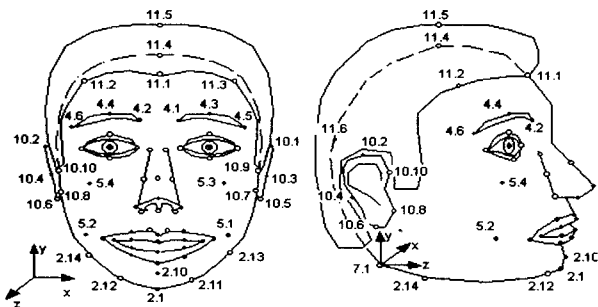


그림 1. FDP와 FAP.  
Fig. 1. FDP and FAP.

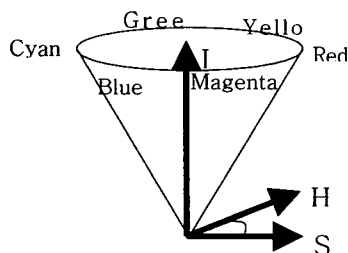


그림 2. HSI 좌표계.  
Fig. 2. HSI coordinate system.

에서 검은 점이 FAP이다. 모든 FAP가 표정에 관여하는 것은 아니고, 얼굴의 형태를 규정하는 것도 있고, 영상처리를 이용하여 찾기 어려운 부분과 표정에 직접적인 영향을 미치지 않는 점들을 제외하면 19개의 특징 점을 찾을 수 있다. 선행 연구에서 보면 표정 특징 점은 일정 얼굴 구성요소 즉, 눈, 코, 입 주변에 분포하는 것을 알 수 있다.<sup>[2][3][5]</sup> 따라서 본 논문에서는 얼굴의 구성요소를 먼저 찾을 필요가 있다.

### III. 전처리과정

먼저 USB카메라로부터 들어온 320×240 24bit RGB영상은 HSI 좌표계로 바뀌어진다. HSI 좌표계는 RGB보다 다루기 쉽고, 색이 각도로 표시되므로 원하는 색의 각도 범위만 알고 있으면 영상에서 해당 색을 찾을 수 있다. 그림 2는 HSI 좌표계 모델을 보였다.

컬러영상에서 찾으려고 하는 피부색은 범위가 일정한 영역에 존재한다. 실험결과 피부색은 H의 값이 24-27에 집중되어 있는 것을 알 수 있다. 그러나 실제 영상은 주변에 피부색과 비슷한 색도 부분적으로 있을 수 있으므로 범위를 크게 잡는 것은 오히려 검출하는데 방해가 된다. 또한 H값의 범위만으로는 피부색을 전체 영상에서 찾을 수 없다. 이유는 USB 카메라는 조명의 영향에 따라 피부색의 분포 값이 변하기 때문이다. 따라서 얼굴색이 가지는 I 값도 고려해서 얼굴 부분을 찾는다. 그림3의 (a)는 입력영상에서 피부색을 찾은 결과이다.

결과는 피부색과 비슷한 색까지 포함되어 있어 얼굴에 해당하는 피부색 정보만을 찾아내기는 힘들다. 따라서 형상학을 이용하여 범위가 작은 값들은 제거한다. 사용한 형태학 연산자는 Opening이다. 그림 3의 (b)는 형상학을 적용한 결과이다.



(a) 얼굴색 검출 (a) Detection of face color  
(b) 형상학 적용 결과 (b) Result of morphology application

그림 3. 전처리 결과.  
Fig. 3. Result of preprocessing.

#### IV. 무게중심을 이용한 얼굴영역 검출

그림 3에서 얻은 결과를 가지고 얼굴 영역의 검출하는데 있어서 얼굴색이 차지하는 영역이 가장 크므로 형상학 적용 결과 남은 영상의 픽셀의 위치 정보에 따른 무게 중심도 얼굴색 영역 내부에 존재하게 된다. 무게 중심을 구하는 식은 식(1), (2)와 같고, 그림 4에서 여러 영상에서 무게 중심 점을 찾은 결과를 보였다.

$$x_m = \frac{\sum_{i=0}^{i < Height} \sum_{j=0}^{j < Width} X_{ij}}{N_1} \quad (1)$$

$$y_m = \frac{\sum_{i=0}^{i < Height} \sum_{j=0}^{j < Width} Y_{ij}}{N_1} \quad (2)$$

$X_{ij}$ ,  $Y_{ij}$ 는 픽셀  $ij$ 의 좌표  
 $N_1$  은 형상학 적용결과 1인 픽셀의 수



그림 4. 무게 중심 점 검출 결과.  
 Fig. 4. Weight central point detection result.

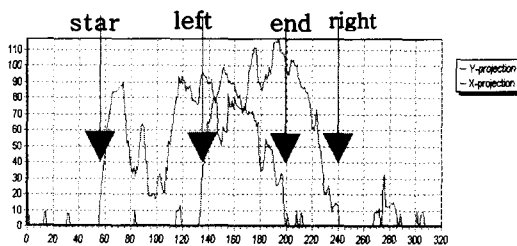


그림 5. 프로젝트에 의한 얼굴영역의 검출.  
 Fig. 5. Detection of face area by projection.

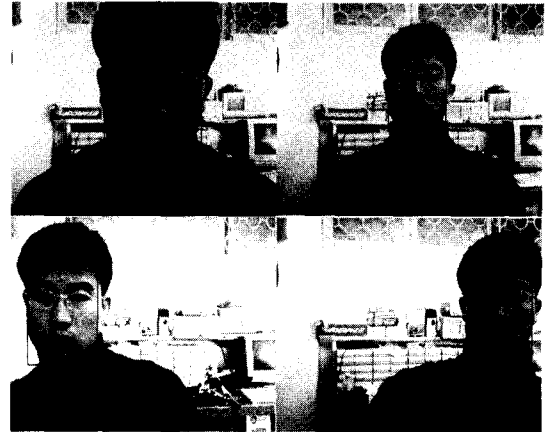


그림 6. 영상에서 찾은 얼굴 영역.  
 Fig. 6. Face areas that find in images.

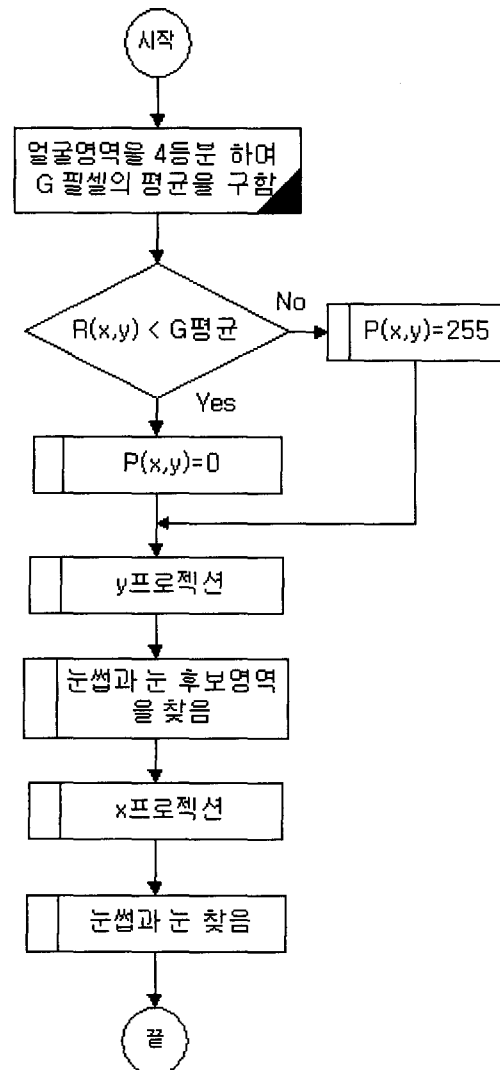


그림 7. 눈썹과 눈 검출 순서도.  
 Fig. 7. Flowchart of eyebrows and eye detection.



그림 8. 눈썹과 눈을 찾은 결과.  
Fig. 8. Result that find eyebrows and eye.

찾은 무게 중심을 이용하여 x축과 y축의 프로젝션을 통하여 얼굴 영역을 찾는다. 무게중심에서 위쪽 끝이 얼굴 영역의 끝이 되고 아래쪽 끝이 얼굴 영역이 된다. 마찬가지로 좌우의 끝을 찾을 수 있다.

무게 중심 법으로 얼굴영역을 찾는데 걸리는 시간 약 0.5초정도 소요된다.<sup>[5]</sup>

### V. 얼굴 구성요소 검출

다음으로 찾은 얼굴 영역에서 얼굴 구성요소(눈, 눈썹, 입)를 찾는다. 얼굴 구성요소는 찾은 얼굴 영역에서 검은 부분에 해당함을 알 수 있다. 따라서 얼굴 영역에서 픽셀 값이 낮은 픽셀을 찾음으로 얼굴 구성요소를 찾을 수 있다. 먼저 눈썹을 찾고, 눈을 찾는 과정은 다음의 순서도에 따른다.

$P(x,y)$ 는 얼굴 영역의 픽셀 값이고,  $R(x,y)$ 는 얼굴 영역의  $P(x,y)$ 의 RGB에서 Red 값이다. RGB 중 G 값을 이용하여 영역의 밝기정보를 얻은 이유는 G가 다른 값에 비하여 밝기에 더 많은 영향을 미치기 때문이다.

입 영역은 눈을 찾은 결과에서 얼굴 구성요소간의 형태학적 관계를 이용하여 후보영역을 찾는다. 일반적으로 입 영역은 눈에서 일정거리에 일정한 영역을 차지한다.

따라서 이 영역을 후보영역으로 정하고 이 영역에 대

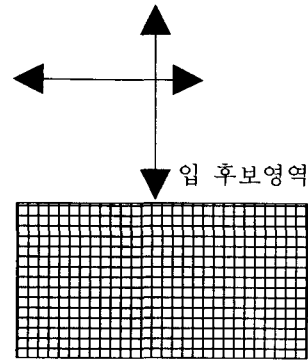


그림 9. 얼굴 구성요소의 형태학적 관계.  
Fig. 9. Relation of face component.



그림 10. 구성요소 추출 결과.  
Fig. 10. Result of characteristic point detection.

하여 입술이 가지는 색상 분포를 이용하여 입 후보영역에서 입을 찾을 수 있다.

찾은 구성요소에서 주변에 분포하는 특징 점을 찾으면 다음과 같다.

### VI. 결론

본 논문에서는 USB카메라 환경에서 얻은 영상에서 표정을 검출할 목적으로 먼저 영상에서 얼굴의 구성요소를 추출하였다. USB 카메라 환경은 조명에 민감하고, 색상 변화도 잦으므로 일반적인 RGB로는 얼굴영역을 찾

아내기가 쉽지 않다. 따라서 HSI영역에서 얼굴 영역을 찾고, 찾은 얼굴 영역에서 얼굴 구성 요소 중 표정에 관계있는 특징 점을 포함하는 눈썹, 눈, 입을 프로젝션을 통하여 찾았다. 정확한 얼굴 영역을 찾아내기 위해서 형상학을 이용하여 크기가 작은 영역은 제거하였으며, 무게 중심을 이용하여 얼굴영역을 추적하였다.

결과로 얻은 얼굴 영상은 1초에 약 2장 정도의 처리속도를 보였다. 향후 연구과제로는 얼굴의 뒤틀림이나 회전에 의한 왜곡을 보정하여 다양한 각도에서도 얼굴의 구성요소를 찾을 수 있도록 얼굴 영상을 보정하는 연구가 필요하다.

추출 방법”, 멀티미디어 학회 추계 학술 발표회, p 256~269, 2000.

## 참 고 문 헌

- [1] P. Ekman, W. V. Friesen, "Facial Action Coding System, Consulting Psychologists Press Inc, 577 College Avenue, 1978.
- [2] A. Blum and P. Langley. "Selection of relevant features and examples in machine learning", *Artificial Intelligence*, 10:245~271, 1997
- [3] B. Heisele, T. Serre, M. Pontil, T. Poggio. "Component-based face detection" In Proc. IEEE Conference in Computer Vision and Pattern Recognition, volume 1, pages 657~662, Hawaii, 2001.
- [4] P. Viola, M. Jones, "Robust real-time face detection", *Proc ICCV*, 20(11):1254~1259, 2001
- [5] H.A. Rowley, S. Baluja, T.Kanade, "neural network-based face detection", *IEEE Trans. Pattern Anal. Machine Intell.*, vol20,p23-28, 1998
- [6] B. Heisele, A. Verri, T. Poggio, "Learning and Vision Machines", *Proceedings of The IEEE*, vol. 90, No. 7, July 2002
- [7] C. Papageorgiou, T. Poggio, "A trainable system for object detection", In *International Journal of Computer vision*, vol. 38, 1, pages 15~33, 2000
- [8] 문대성 "피부색 영역 확장에 의한 얼굴 영역

— 저 자 소 개 —



이 철 희(학생회원)  
 2000년 원광대학교 석사 학위  
 취득  
 2000년 현재 원광대학교  
 전자공학과 박사과정  
 <주관심분야:Nonlinear system을  
 이용한 영상처리, 인 식분야>



문 성 룡(정회원)  
 2004년 현재 원광대학교  
 전자공학과 부교수  
 <주관심분야:신경망 시스템, 인식  
 분야, 영상처리 >