

# 액티브 카메라와 피부색상에 의한 다중 얼굴 검출 및 추적

김 광 희, 이 배 호  
전남대학교 컴퓨터공학과

## Multiple face detection and tracking using active camera and skin color

Kwang-Hee Kim, Bae-Ho Lee  
Dept. of Computer Engineering, Chonnam National University  
E-mail : ghkim@cisom.chonnam.ac.kr

### 요 약

본 논문에서는 실내에서 액티브 카메라를 사용하여 다중 인물의 얼굴의 위치를 검출하고, 추적할 수 있으며 조명과 배경 등의 영향에 강인한 추적 알고리즘을 제시하고자 한다. 알고리즘은 얼굴영역 검출, 추적의 2 단계로 구성되며, 빠르고 효율적인 얼굴영역 검출은 추적 알고리즘의 성능향상으로 이어지므로, 이를 위해 독특한 색상영역 분포를 갖는 피부 색상 특징을 이용하였다. 표본영상에서 추출된 피부색상 픽셀들을 바탕으로 YCbCr 색상계를 사용하여 얼굴 색상모델을 구축한 후, Gaussian 함수를 사용하여 입력 영상의 픽셀과 얼굴 색상모델과의 유사도를 결정하였다. 최종 얼굴 영역은 추출된 영역에 대한 얼굴의 타원특징, 해부학적 특징을 이용하여 결정된다. 추적은 추출된 얼굴영역과 temporal Gaussian 필터를 적용한 움직임 추정을 통한 움직임 검출의 조합으로 이루어진다. 또한, 예측버퍼의 사용으로 탐색영역의 축소로 인한 계산량 감소와 처리 속도의 증가시켰으며, pan/tilt가 가능한 카메라를 사용하여 상호 피드백이 가능하도록 하였다. 제시된 알고리즘은 PC 상에서 시뮬레이션되었으며, 좋은 결과를 얻을 수 있었다.

### 1. 서 론

실시간 얼굴 검출 및 추적은 영상 보안 및 감시 시스템에 있어서 핵심적인 요소이다. 그러나, 지금까지

얼굴 인식 및 검출, 추적에 대한 많은 연구가 진행되어 왔으나, 영상 내에 분포하는 얼굴의 다양한 크기 및 형태, 위치, 기울기와 같은 일반적인 특징으로 인해 효율적인 알고리즘을 구현하기에는 많은 문제점이 존재한다. 또한, 얼굴 영상 자체뿐만 아니라 복잡한 배경, 영상의 질, 조명의 영향과 같은 외적인 요소들 역시 얼굴 영역 검출 및 인식 등을 어렵게 만든다.

최근에는 얼굴 피부색상과 형태를 이용한 얼굴 영역의 분리 및 추적에 관한 연구가 활발히 진행중이다[1-7]. 특히, 피부색상의 경우 매우 훌륭한 전처리 필터의 역할을 수행하므로, 얼굴색상모델 구축에 관한 다양한 방법이 제시되었다[1-7].

본 논문에서는 실내에서 액티브 카메라를 사용하여 다중 인물의 얼굴의 위치를 검출하고, 추적할 수 있으며 조명과 배경 등의 영향에 강인한 추적 알고리즘을 제시하고자 한다. 제안한 방법 역시 피부색상과 얼굴 형태에 바탕을 두었으며, 프레임 간의 움직임을 추출하고 탐색 영역 예측을 수행하여 프레임당 계산량 감소와 처리속도 향상에 주안점을 두었다.

알고리즘은 얼굴영역 검출, 추적의 2단계로 구성되며, 얼굴 영역 검출을 위해 표본영상에서 추출된 피부색상 픽셀들을 바탕으로 YCbCr 색상계를 사용하여 얼굴 색상모델을 구축한 후, Gaussian 함수를 사용하여 입력 영상의 픽셀과 얼굴 색상모델과의 유사도를 결정하였다.

추적은 추출된 얼굴영역과 temporal Gaussian 필터를 적용한 움직임 추정을 통한 움직임 검출의 조합으로 이루어진다. 또한, 예측버퍼의 사용으로 탐색영역의 축

소로 인한 계산량 감소와 처리 속도의 증가시켰으며, pan/tilt가 가능한 카메라를 사용하여 상호 피드백이 가능하도록 하였다. 실험결과 다중 인물이나 배경에 크게 관계없이 효율적인 얼굴영역 검출 및 추적이 가능함을 알 수 있었다.

## II. 얼굴 영역 검출 알고리즘

### 1. 피부 색상추출

얼굴은 좌우 대칭, 얼굴 특징 요소(눈, 코, 입 등), 피부색상과 같은 몇 가지 주요한 특징을 가지며, 특히 피부색상은 다른 사물과 구별되는 특정한 색상 영역을 가지고 있다. 물론, 얼굴 비슷한 색상을 포함하고 있는 사물들도 많이 존재하지만 대부분의 경우 얼굴 색상과는 많은 차이를 보이고 있다. 따라서, 이를 이용할 경우 피부색상은 복잡한 배경에서 얼굴 영역 추출을 위한 훌륭한 선행 처리필터가 된다. 또한, 얼굴 색상은 개인, 인종마다 다르지만 색차(chrominance) 영역에서는 매우 비슷한 영역을 가지며, 단지 밝기의 차이(intensity) 만이 있을 뿐이다[8].

본 논문에서는 YCrCb 색상계와 90개의 표본 얼굴 영상을 바탕으로 얼굴 색상모델을 구축하였다. YCrCb 색상 모델에서 Y는 명암 정보를 나타내므로, 조명의 영향을 감소시키기 위해 Y를 제외한 CrCb 평면이 색상정보를 나타낸다. 식(1)에 의해 YCbCr 색상계로의 변환이 수행된다.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

그림 1(a)는 CbCr 평면상의 얼굴 색상모델을 나타내며, 그림 1(b)는 색상모델에 기반한 피부색상 추출 결과를 나타내고 있다.

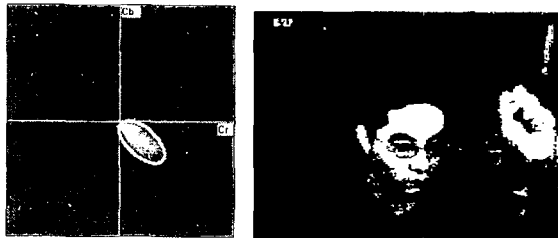


그림 1. (a) 얼굴 색상모델 (b) 피부색상 추출

그림 1(a)의 얼굴 색상 모델은 식(2)의 Gaussian 색상모델에 의해 CbCr 평면에서의 영역이 결정된다.

$$p(x, y) = \frac{e^{-\frac{1}{2(1-\rho_{xy}^2)} \left[ \frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho_{xy}(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right]}}{2\pi\sigma_x\sigma_y\sqrt{(1-\rho_{xy}^2)}} \quad (2)$$

식(2)에 사용된 각 파라미터들은 식(3)에 의해 구할 수 있다.

$$\begin{aligned} \mu_x &= \frac{1}{N} \sum_{n=1}^N X_n, \quad \mu_y = \frac{1}{N} \sum_{n=1}^N Y_n \\ \sigma_x &= \sqrt{\frac{1}{N} \sum_{n=1}^N (X_n - \mu_x)^2}, \quad \sigma_y = \sqrt{\frac{1}{N} \sum_{n=1}^N (Y_n - \mu_y)^2} \\ \rho_{xy} &= \frac{\frac{1}{N} \sum_{n=1}^N (X_n - \mu_x)(Y_n - \mu_y)}{\sigma_x\sigma_y} \end{aligned} \quad (3)$$

본 논문에서는 식(2)의 결과를 일반화하기 위해 식(4)와 같이 표본색상 히스토그램과 가중치를 곱한 후 이를 피부 색상을 판별하기 위한 기준으로 사용하였다.

$$Th = C(x, y) [2\pi\sigma_x\sigma_y\sqrt{(1-\rho_{xy}^2)}] p(x, y) \quad (4)$$

( 단,  $C(x, y)$  는 표본색상 히스토그램 )

### 2. 후보 영역 추출

#### 2.1 Filtering

피부색상 추출은 픽셀 단위로 이루어지므로 추출된 얼굴 픽셀들은 완전한 영역을 이루지 못하며, 입력 영상에 따라 다양한 형태를 나타내게 된다. 따라서, 이를 적당한 영역으로 묶고 잡음을 제거하기 위해 5\*5 median filtering을 수행하였다. 또한, 일반적인 영상에서 얼굴 영역은 최소한의 크기를 가지므로 20\*20 이하의 크기를 갖는 영역은 모두 제거하였다.

#### 2.2 형태 제약조건

얼굴은 여러 특징 중 타원형의 형태를 나타내므로 필터링을 거친 영역들에 대해 각 영역의 타원적인 특성을 계산함으로써, 좀 더 얼굴에 가까운 후보 영역들을 추출할 수 있다. 그러나, 일반적인 영상에서 얼굴의 크기 및 회전 등은 매우 다양하게 나타나므로, 템플릿을 이용한 방법에는 한계가 있으며, 영상에 나타나는 얼굴의 형태에 상관없이 영역의 타원특성을 구할 수 있는

방법이 필요하게 된다. 이에 본 논문에서 영역에 대한 타원 파라미터를 구할 수 있는 best-fit ellipse 알고리즘을 적용하여 타원특성을 구하고, 미리 정의된 제약조건에 의해 얼굴 후보 영역을 결정하였다.

각 영역에 대한 best-fit ellipse의 파라미터 - 기울기  $\theta$ , submajor  $a$ , subminor  $b$  - 는 식(5)에 의해 계산된다.

$$\theta = \frac{1}{2} \tan^{-1} \left[ \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right] \quad (5)$$

$$a = \left( \frac{4}{\pi} \right)^{1/4} \left[ \frac{(I'_{\max})^3}{I'_{\min}} \right]^{1/8}, \quad b = \left( \frac{4}{\pi} \right)^{1/4} \left[ \frac{(I'_{\min})^3}{I'_{\max}} \right]^{1/8}$$

( 단,  $I$  는 관성모멘트 )

해부학 구조상 일반적인 얼굴은 1:4~1:6의 비율을 가지므로 식(5)에 의해 계산된 파라미터 값 중, 최대  $a : b$  의 비율이 1:2,  $|\theta| > 45^\circ$ 인 영역을 얼굴후보 영역으로 추출하였다.

### 3. 얼굴영역 검출

추출된 후보 영역들 중 일부 영역들은 배경이거나, 신체의 다른 영역(팔, 다리)일 수 있다. 따라서, 추출된 각 후보 영역에 대해 얼굴 영역 여부를 판별해야 할 필요성이 있으며, 본 논문에서는 이를 위해 신경망을 적용하였다. 신경망 학습을 위해 모넨텀 역전파 알고리즘을 사용하였으며, 후보 영역을 20\*20 픽셀 크기로 일반화하고 히스토그램 equalization을 수행한 후 마스크를 통과한 영상을 입력 패턴으로 사용하였다. 신경망 학습을 위해 150개의 얼굴 패턴 영상과 200개의 비 얼굴 패턴 영상을 사용하였다.

## III. 얼굴 추적 알고리즘

### 1. 움직임 검출

동영상에서 움직임은 반드시 존재하며, 이는 물체 추적을 위한 강력한 실마리를 제공한다. 움직임은 프레임간의 차이에 의해 발생되나, 단순히 차영상에 의한 움직임 추출은 조명 변화나 배경 등에 의해 많은 에러를 포함하게 된다. 따라서, 조명이나 배경의 영향에 둔감한 움직임 검출 방법이 필요하게 된다.

본 논문에서는 움직임 검출을 위해 움직임 추정 방법을 사용하였다[9]. 노이즈를 제거하기 위해 Gaussian smoothing 필터를 적용한 각 프레임  $I(x, y, t)$ 에 2차 미분 temporal-Gaussian 필터를 사용하였으며, 식(6)에 잘 나타나있다.

$$Z(x, y, t) = \frac{\partial^2 G(t)}{\partial t^2} \otimes I(x, y, t) \quad (6)$$

이때, temporal-Gaussian 필터는 식(7)과 같이 정의된다.

$$\frac{\partial^2 G(t)}{\partial t^2} = \frac{-2s^3}{\sqrt{\pi}} (1 - 2s^2 t^2) \exp(-s^2 t^2) \quad (7)$$

( 단,  $s = \sqrt{3/(2n^2)}$ ,  $n = 1, 2, \dots$  )

식(6)에 의해 움직임은  $Z(x, y, t)$  상의 zero-crossing 지점을 검출함으로써 구할 수 있다.

### 2. 얼굴 추적

얼굴 추적 프로세스는 이전 단계에서 추출된 움직임 영역과 색상영역의 곱에 의해 수행된다. 그림 2는 전체 추적 프로세스를 나타내고 있다.

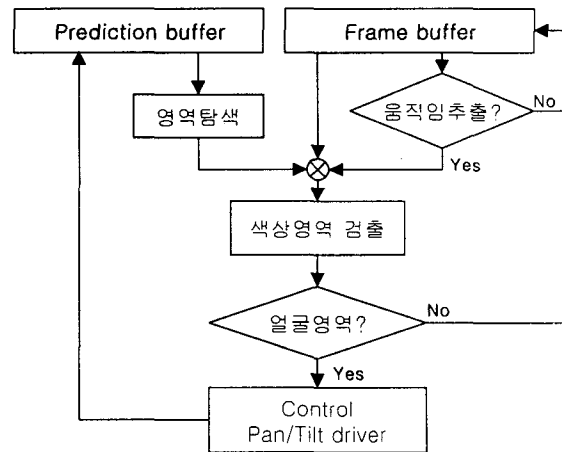


그림 2. 얼굴추적 프로세스

먼저, 프레임 버퍼로부터 움직임 추출 프로세스가 프레임을 넘겨받아 움직임 추정에 의해 움직임을 추출한다. 움직임이 추출되면, 현재 프레임과 움직임 추출 영역과의 교집합 부분에 대해 색상영역 검출을 수행함으로써 계산량의 감소와 처리속도가 향상된다.

현재 프레임의 얼굴영역 위치가 결정되면, Pan/Tilt driver를 제어한 후 그 결과를 예측 버퍼에 저장하여 다음 프레임 처리시 탐색 영역을 예측할 수 있도록 한다. 예측 버퍼에 데이터가 존재하면 움직임 추출영역, 현재 처리할 프레임 그리고 예측 영역을 모두 곱하여 색상 영역검출을 수행한다.

#### IV. 실험결과

본 논문에서는 제안한 알고리즘을 바탕으로 PC상에서 얼굴 추적 시스템을 구축하였다. 영상 취득을 위해 Matrox MeteorII-1394와 SONY VL500 카메라를 사용하였으며, 입력 영상의 해상도는 320\*240 이었으며, 초당 15-20 프레임을 캡처하여, AMD 850Mhz 환경에서 VC++6.0으로 테스트를 수행하였다.

그림 3은 얼굴 추적 결과로서 180 프레임으로 구성된 동영상 시퀀스 중 얼굴 영역 검출에 성공한 일부 프레임들을 나타내고 있다.

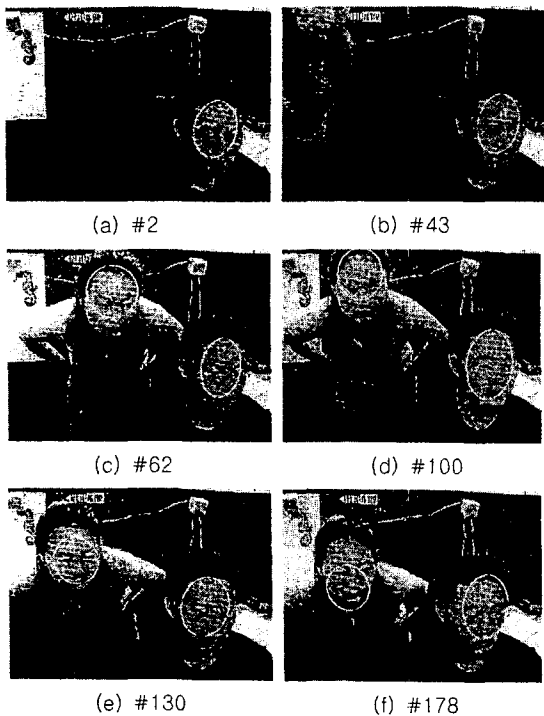


그림 3. 얼굴 추적 결과( 180 프레임 )

위 그림 3에서와 본 논문에서 제시한 알고리즘에 의해 구현된 추적 시스템에서는 인물의 수에 상관없이 효율적인 얼굴추적이 가능함을 알 수 있다.

#### V. 결론

본 논문에서는 얼굴의 피부색상과 타원 형태의 특징, temporal-Gaussian 필터를 바탕으로 특별한 제약이 없는 실내영상에서 얼굴 추적을 수행할 수 있는 알고리즘을 제시하였다. 배경에서 효과적으로 얼굴 영역을 분리하고 조명의 영향을 감소시키기 위해 YCrCb 색

상계를 사용하여 얼굴 색상 모델을 구축하였다. 또한, 얼굴의 타원적인 특징 형태를 이용하기 위해 best-fit ellipse를 사용하여 얼굴의 크기 및 회전 등에 적합한 얼굴 영역추출 방법을 제안하였다.

효율적인 실시간 추적이 가능하도록 예측버퍼에 의한 탐색영역 예측과 움직임 추출영역을 이용하여 계산량을 줄이고, 프레임당 얼굴영역 검출 속도를 향상시켰다.

그러나, 컬러 영상을 사용하는 만큼 입력 영상의 색상 강도에 따라 전체 시스템이 많은 영향을 받게된다. 특히, 추적 장소의 변경이나, 사용 카메라의 변경은 많은 색상 왜곡을 야기하므로 이에 대한 색상 보정 알고리즘의 개발이 요구된다. 또한, Pan/Tilt driver의 기계적인 특성상 이동속도에 따른 에러가 발생하는데, 기계적인 특성에 맞는 움직임 보상 및 예측 알고리즘을 적용할 경우 보다 향상된 결과를 얻을 수 있을 것으로 기대된다.

#### 참고 문헌

- [1] J. Yang, A. Waibel, "A real-time face tracker", IEEE Work. on Appl. Comp. Vis., Sarasota, pp.142-147, 1996
- [2] K. Sobottka, I. Pitas, "Segmentation and tracking of faces in color images", Proc. Automatic Face and Gesture Recognition, pp.236-241, 1996
- [3] P. Fieguth, D. Terzopoulos, "Color-based tracking of heads and other mobile objects at video frame rates", IEEE Conf. on Comp. Vis. and Pat. Rec. Puerto Rico, pp.21-27, 1997
- [4] GR. Bradski, "Computer vision face tracking as a component of a perceptual user interface", IEEE Work. on Appl. Comp. Vis., Princeton, pp.214-219, 1998
- [5] K. Schwerdt, J.L. Crowley, "Robust face tracking using color", Automatic Face and Gesture Recognition, pp.90-95, 2000
- [6] L. Jordao, M. Perrone, J.P. Costeira and J. Santos-Victor, "Active face and feature tracking", Proc. Image Analysis and Processing, pp.27-29, 1999
- [7] D. Comaniciu, V. Ramesh, "Robust detection and tracking of human faces with an active camera", Proc. Visual Surveillance, pp.11-18, 2000
- [8] H. Wang, S. Chang, "A Highly Efficient System for Automatic Face Region Detection in MPEG Video", IEEE Tr. TCSVT, Vol. 7, 1997
- [9] S. Gong, S. J. McKenna and A. Psarrou, *Dynamic vision from images to face recognition*, Imperial College Press, 2000