

화상회의 카메라 제어를 위한 안면 검출 알고리즘

온승엽¹, 박계현², 박규식³, 이준희¹

1. 홍익대학교 전자 전기 컴퓨터 공학부,

2. 명지대학교 전자 정보 통신 공학부

3. 상명대학교 컴퓨터 정보 통신 학부

1. 전화 : (0415) 860-2478 / 팩스 : (0415) 865-0460

Face Detection Algorithm for Video Conference Camera Control

Syng-Yup Ohn¹, Jae-Hyun Park², Kyu-Sik Park³, Jun-Hee Lee¹

1. School of Electronics, Electrical, and Computer Engineering, Hongik University

2. School of Electronics, Information, and Communication Engineering, Myungji University

3. School of Computer, Information, and Communication, Sangmyung University

1. E-mail : syohn@wow.hongik.ac.kr

Abstract

In this paper, we propose a new algorithm to detect human faces for controlling a camera used in video conference. We model the distribution of skin color and set up the standard skin color in YIQ color space. An input video frame image is segmented into skin and non-skin segments by comparing the standard skin color and each pixels in the input video frame. Then, shape filter is applied to select face segments from skin segments. Our algorithm detects human faces in real time to control a camera to capture a human face with a proper size and position.

I. 서론

본 연구는 한국과학재단 특정 기초 연구 (과제번호 1999-2-302-018-3) 지원으로 수행되었음.

안면(얼굴)은 사람 대 사람의 통신에서 큰 역할을 담당한다. 또한, 최근 사람 대 컴퓨터의 상호 작용(HCI)의 중요성이 부각되어 안면이 새로운 인터페이스의 하나로써 등장하고 있다. 영상 프레임 내에서 안면의 위치와 크기 정보를 알아내는 안면 검출은 안면인식, 표정인식, 영상코딩 등의 다양한 응용분야에서 필수적이다. 안면검출에 관한 과거 연구 동향을 살펴보면 그레이 스케일 화상을 이용하는 알고리즘[1, 2, 3], 컬러 영상을 이용하는 알고리즘[4, 5, 6] 등이 있다. 이미 발표된 대부분의 알고리즘들은 안면의 크기나 방향 등을 제한적으로 가정하고 있거나[1, 2, 3, 4, 6], 패턴 매칭 등에 의존하여 계산시간이 길어지는 단점이 있다[3, 4, 6, 7, 8]. 그러나, 화상회의 시스템의 경우 회의 환경이라는 속성 때문에 안면검출 단계에서 안면의 속성에 대하여 제한적인 가정을 할 수 없고 영상이 실시간으로 전송되어야 하므로 짧은 계산시간이 요구된다.

본 논문에서는 영상 프레임 안에서 사람의 안면을 검출하는 알고리즘을 제안한다. 여기에서 검출된 안면의 위치 및 크기 정보는 화상회의 시스템 카메라의 pan, tilt, zoom을 제어하여 안면을 화상 프레임 안에 적당한 크기와 위치에 포착할 수 있도록 하는 카메라

제어 시스템의 입력으로 활용된다. 본 논문에서 소개되는 알고리즘은 피부색을 모델링하여 얻어진 기준 피부색과 영상 안의 각 화소 값을 비교하여 영상프레임을 피부 영역과 비피부 영역으로 분리하고, 피부 영역으로부터 안면의 크기와 위치를 검출한다. 이 알고리즘은 검출 대상이 되는 안면의 속성에 대하여 제한적인 가정을 하지 않고 보통 수준의 PC 환경에서 짧은 수행시간을 가진다. 본 논문의 구성은 다음과 같다. 2장은 피부색의 모델링과 피부 영역의 분리에 대하여 설명하고, 3장은 분리된 피부 영역들에 대하여 잡음을 제거하고 영상 필터를 적용하여 안면 영역을 선택하는 과정을 보이고, 4 장은 결론을 기술한다.

II. 피부영역의 분리

YIQ 컬러 표준 형식은 휘도(luminance) 성분, Y와 색조(color) 성분, IQ로 나누어지므로 휘도와 색조 성분을 분리하여 영상을 처리할 수 있다. 특히, 화상회의 시스템을 위한 안면 검출의 경우, 조명이 이상적으로 설치되지 않은 사무실 상황에서 영상이 얻어지므로 역광, 글레어(glare) 등이 발생할 수 있고, 이것에 따른 영상의 넓은 휘도 변동 범위를 효과적으로 수용할 수 있어야 하므로 YIQ 형식이 RGB 형식보다 적합하다.

YIQ 컬러 공간에서 피부색의 분포를 3차원 정규 분포로써 모델링하고, 피부색의 분포의 중앙값과 카메라 영상의 각 화소 값을 비교함으로써 해당 화소가 피부 영역 또는, 비피부영역의 화소인지를 판별할 수 있다. 화소의 판별은 피부색 분포의 중앙값과 카메라 영상의 각 화소 값간의 Mahalanobis 거리를 계산하여 정하여진 값으로 임계함으로써 이루어진다.

일반적인 3 차원 정규 분포는

$$p(x) = \frac{\|C^{-1}\|^{1/2}}{(2\pi)^{3/2}} \exp\left[-\frac{1}{2} (x-m)^T C^{-1} (x-m)\right]$$

으로 나타내어진다. 피부색의 분포를 정규 분포로 모델링하기 위한 중앙값 m 과 공분산 행렬 C 는 샘플 화상들의 안면 영역의 화소들을 수동적으로 채취하여 다음의 식을 이용하여 구할 수 있다.

$$m = \frac{1}{n} \sum_{i=1}^n x_i, \quad C = \frac{1}{n} \sum_{i=1}^n (x_i - m)(x_i - m)^T$$

여기에서 $x_i = [Y \ I \ Q]^T$ 는 YIQ 형식으로 나타내어진 샘플 영상의 안면 부분의 화소 값이다. 피부색분포의 중앙값은

$$m = \begin{bmatrix} 122.169 \\ 24.209 \\ -3.719 \end{bmatrix}$$

이고, 공분산 행렬은

$$C = \begin{bmatrix} 437.664 & 35.712 & -15.672 \\ 35.712 & 16.757 & -2.228 \\ -15.672 & -2.228 & 4.791 \end{bmatrix}$$

이다.

피부색 분포의 중앙값, μ 와 카메라 영상의 각 화소 값, x 간의 Mahalanobis 거리[9]는

$$D = (x - m)^T C^{-1} (x - m)$$

으로 나타내어진다.

만일, Mahalanobis 거리가 정해진 임계값 이하이면 피부영역의 화소로 분류된다. 그림 1의 입력 영상에 대한 시뮬레이션 결과가 그림 2에 보여진다.



그림 1 입력 영상

Fig. 1 Input Video Frame

III. 잡음 제거 및 영상 필터링

피부영역과 비피부영역에 나타나는 잡음을 제거하기 위하여 피부영역의 mask image에 Morphology opening과 closing 연산이 적용된다[10]. opening 과 closing은 각각 Morphology dilation 과 erosion 연산을 이용하여 다음과 같이 정의된다.

$$\cdot \text{Opening} : X \circ B = (X \ominus B) \oplus B$$

$$\cdot \text{Closing} : X \cdot B = (X \oplus B) \ominus B$$

opening과 closing operation은 화상의 작은 파편 등을 제거하고 영역의 복잡한 경계부분을 단순화시키는 역할을 한다.

mask image의 잡음 제거를 위하여 opening 과 closing이 다음과 같이 두 번 반복된다.

$$X \circ B_1 \cdot B_2 \circ B_1 \cdot B_2$$

여기에서 X 는 mask image 이며, 적용된 structuring element B_1, B_2 는 다음과 같다.



그림 2 D < 35로 임계 된 영상
Fig. 2 Video Frame thresholded by D < 35

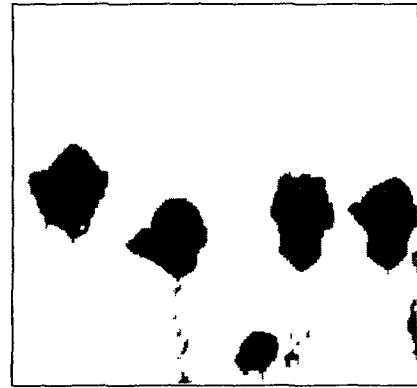


그림 3 잡음 제거후의 화상
Fig. 3 Skin segment image after removing noise

$$B_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad B_2 = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

그림 3 은 잡음 제거후의 mask image를 원래의 화상 프레임에 적용한 화상이다.

피부영역 영역 구분을 위해 mask image에 Blob Coloring 알고리즘[11]을 적용하고 각 피부영역의 eccentricity, 크기, 위치 등의 feature를 검출해낸다. 사람의 얼굴은 일반적으로 세로가 긴 타원의 기하학적인 특성을 가지므로 샘플 조사를 통하여 기준 안면 형상을 모델링할 수 있다. 이 기준 안면 형상과 각각의 피부영역의 feature를 비교하여 피부 영역 중 안면이 아닌 영역을 제거한다. 그림 4 에는 검출된 안면 영역을 사각형으로 표시했다.

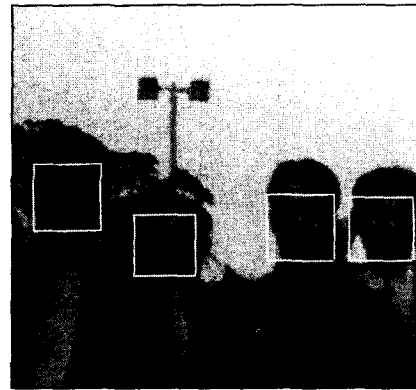


그림 4 안면이 검출된 화상
Fig. 4 Image with faces detected

IV. 결론

본 논문에서는 화상회의 카메라의 제어를 위한 안면 검출 알고리즘이 제시되었다. 여기서 제시된 알고리즘은 사진에 검출 대상 안면의 크기, 위치, 방향 등에 대하여 제한받지 않으며 짧은 계산 시간을 가진다. 따라서, 여기서 제시된 알고리즘은 화상회의 시스템뿐만 아니라 다양한 실시간 응용 분야에 적용될 수 있을 것으로 기대한다.

참고문헌

- [1] G. Yang and T. Huang. "Human face detection in a scene", In Proc. IEEE Conf. Comput. Vision Pattern Recog., 453-458. IEEE, 1993.
- [2] X. Song, C. Lee, G. Xu, and S. Tsuiji. "Extracting facial features with partial feature template", In Proc. First Asian Conf. Comput. Vision, 751-754, Osaka, 1993.
- [3] A. Tsukamoto, C. Lee, and S. Tsuiji. "Detection and tracking of human face with synthesixed template", In Proc. First Asian Conf. Comput. Vision, 183-196.
- [4] Q. Chen, H. Wu, and M. Yachida. "Face

detection by fuzzy pattern matching", In Proc. Fifth Int'l Conf. Comput. Vision, 591-596, Boston, 1995. IEEE.

[5] K. Sobottka and I. Pitas. "Extraction of facial region and features using color and shape information", In Proc. 10th Int'l Conf. Pattern Recog., C421-425, Vienna, 1996.

[6] H. Wu, Q. Chen, and M. Yachida. "Face detection from color images by fuzzy pattern matching", IEICE Trans. JDI, 80(7):1774-1785, 1997.

[7] A. Tsukamoto, C. Lee, and S. Tsuji. "Detection and tracking of human face with synthesized template", In Proc. First Asian Conf. Comput. Vision, 183-186, Osaka, 1993.

[8] H. Rowley, S. Baluja, and K. T. "Human face detection in visual scenes", Technical Report CMU-CS-95-158R, School of Computer Science, Carnegie Mellon University, 1995.

[9] J. T. Tou, R. C. Gonzalez. *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, 1974.

[10] D. H. Ballard, C. M. Brown. *Computer Vision*, Prentice Hall, Englewood Cliffs, NJ, 1982.

[11] Jean Serra. *Image Analysis and Mathematical Morphology*. Vol 1, Academic Press, London, 1982.