

색상 정보를 이용한 얼굴 영역 추출

정회원 장 선 아*, 유지 상*

Face Detection Using Color Information

Sun-ah Chang*, Jisang Yoo* *Regular Members*

요 약

본 논문에서는 컬러 정지 영상에서 색상 정보를 이용한 얼굴객체 검출 알고리즘을 제안하였다. 제안한 알고리즘에서는 먼저 입력 컬러 영상을 YCbCr로 색상 변환을 한 후, 컬러 정보를 갖고 있는 Cb 와 Cr 에 대하여 살색 범위 내의 값의 부분을 추출하게 된다. 이 결과에 잡음 성분을 제거하기 위하여 형태학적 필터(Morphological filter)를 적용하였고, 수평, 수직 방향으로 스캔하여 1차 후보 영역을 찾아낸다. 1차 후보 영역이 얼굴 영역인지 아닌지를 판별하고, 얼굴 영역이 아니라면 2차 후보 영역을 찾게 된다.

제안한 알고리즘은 YCbCr에서 색상 성분인 Cb와 Cr 만을 사용하기 때문에 조명에 큰 영향 없이 얼굴 영역을 추출할 수 있었고, 비교적 간단한 계산 방법으로 고개를 좌우로 돌린 영상, 고개를 위 아래로 움직인 영상, 얼굴의 표정, 얼굴의 크기, 영상의 화질에 관계없이 얼굴 영역을 찾아낼 수 있었다.

ABSTRACT

In this paper, This paper presents a new algorithm which is used for detecting and extracting human masks from a color still image. The regions where each pixel has a value of skin-color were extracted from the Cb and Cr images, after the tone of the color image is converted to YCbCr form. A morphological filter is used to eliminate noise in the resulting image. By scanning it in horizontal and vertical ways under threshold value, first candidate section is chosen. If it is not a face, secondary candidate section is taken and is divided into two candidate sections.

The proposed algorithm is not affected by the variation of illuminations, because it uses only Cb and Cr components in YCbCr color format. Moreover, the face recognition was possible regardless of the degree of shifting face, changed shape, various sizes of the face, and the quality of image.

1. 서론

블록 최근 수 년 동안 디지털 영상에서 사람의 얼굴을 찾아내고 인식하는 분야가 지속적으로 연구되어 왔다. 특히 얼굴은 많은 정보를 포함하고 있어, 얼굴 영역 추출은 얼굴 인식, 표정 인식 등의 얼굴 관련 연구의 핵심 기술이다. 추출된 얼굴 영역은 범죄자 검색 시스템, 개인용 컴퓨터의 접근 제

어, 공공 기관에서의 보안 시스템 등의 분야에 이용되어 오고 있다.

얼굴 영역을 검출하는데 사용될 수 있는 정보는 크게 모양 정보, 색상 정보, 움직임 정보 등으로 나눌 수 있다. 모양 정보는 모노크롬 영상에서 주로 에지나 명안의 편차 등을 추출하고 얼굴의 사전 지식을 적용하여 얼굴 영역을 추출하는데 사용되거나, 색상 정보를 이용하여 살색 영역을 찾아내고, 얼굴 모양이 타원이라는 정보를 이용하여 살색 영역 중

* 광운대학교 전자공학과 디지털미디어 연구실(jsyoo@daisy.kwangwoon.ac.kr, sachang@notre.co.kr)

논문번호 : 00043-0207, 접수일자 : 2000년 2월 7일

※ 이 논문은 1998년 한국학술진흥재단의 학술연구비에 의하여 지원되었음.

타윈 모양을 가진 영역이 얼굴 영역이라고 판단하는데 사용된다. 색상 정보는 인간의 얼굴이 가지는 고유의 색상을 이용하는 것으로 인종이나 개인마다 약간의 편차를 가지고 있으나 색상 정보와 다른 정보를 같이 이용하면 성능이 더 좋아질 수 있고, 현재의 대부분의 영상이 컬러 영상이므로 이용가치가 높다. 움직임 정보는 동영상에서 물체의 움직임으로 인한 정보를 사용하는 것이다.

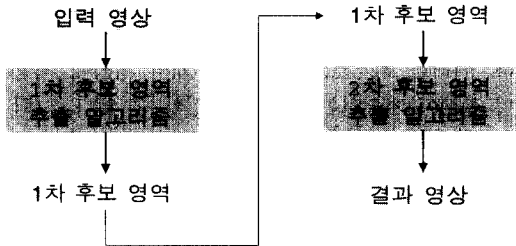


그림 1. 제안된 알고리즘 구성도

본 논문에서는 컬러 정지 영상에서 얼굴 객체를 검출하기 위한 알고리즘을 제안하였다. 제안한 알고리즘의 구성은 그림 1에 보인 바와 같이 크게 1차 후보영역 추출과정과 2차 후보영역 추출과정으로 구성되어 있다. 컬러 영상에 대하여 얼굴을 검출하기 위한 기본적인 방법으로 색상 정보를 이용하였다. 우선 입력 영상은 사진이나 비디오 시퀀스 중 한 프레임 또는 텔레비전 화면을 캡처하여 획득하였고, 획득된 영상을 YCbCr로 색상 변환을 한 후, Cb와 Cr의 색상 성분이 살색을 나타내는 일정한 범위 내의 값을 갖는 영역을 추출해 낸다. 이 결과 영상에서 잡음 성분들을 제거하기 위하여 형태학적 필터를 사용하였고, 수평 방향과 수직 방향으로 스캔하여 일정한 임계값을 가지고 형태학적 필터를 가지고 제거하지 못한 나머지 잡음 성분들을 제거하여 1차 후보 영역을 찾아낸다.

1차 후보 영역은 얼굴 영역일 수도 있고, 잡음으로 인하여 얼굴보다 큰 영역을 찾아낼 수 있다. 1차 후보 영역이 얼굴 영역인지 여부는 타윈 넓이를 이용한 임계값을 이용하여 판별하고, 얼굴 영역이 아니라면 2차 후보 영역을 찾아내게 된다. 2차 후보 영역은 두 개의 후보 영역으로 나누어 각각의 후보 영역이 얼굴 영역인지 아닌지를 타윈 넓이를 이용한 임계값과 가로 세로 비를 이용하여 구한 임계값을 이용하여 얼굴 영역을 추출하게 된다.

본 논문의 구성은 다음과 같다. II장에서는 얼굴

영역 추출과 관련된 연구를 소개하고, 색상 공간에 대한 기본적인 설명과 본 논문에서 사용한 핵심 기술인 살색 정보를 추출하는 과정에 대해 설명하였다. III장에서는 살색 정보를 이용한 알고리즘에 대해 자세히 설명하고, IV장에서는 제한된 알고리즘의 성능 평가를 위한 실험 결과를 소개하였다. 마지막 V장에서 결론을 언급하였다.

II. 얼굴 영역 추출을 위한 색상 정보

2.1 색상 공간

색상 공간은 컬러들 간의 관계를 표현하는 방법으로 컬러 영상들은 여러 가지 색상 공간 모델들 중에 하나로 표현된다. 대표적으로 많이 쓰이는 공간이 RGB, HSI, YCbCr 공간이다^[1]. RGB 색상 공간은 서로 가산될 수 있는 빨강(red), 초록(green), 파랑(blue)의 주요 3 가지색으로 구성되어 있다. 이들 컬러의 분광 요소들이 부가적으로 복합되어 결과적인 컬러를 만들어 낸다. 이것은 Hardware-oriented 모델이고, 컬러 모니터 디스플레이에 주로 사용되고 있다.

RGB 모형은 각 축의 모서리가 빨강, 초록, 그리고 파랑인 3차원 입방체로 표현된다. 검은색은 원점이며 흰색은 입방체의 반대쪽 끝점이 된다. 명암도는 검은색에서 흰색으로 이어지는 선을 따라 표현된다. 컬러는 24비트를 가지는 그래픽 시스템에서 각 색상 별로 8비트를 가지게 된다. 즉, R,G,B 각각이 8비트씩 할당을 받아 컬러를 나타낸다. 예를 들어 빨간색은 (255, 0, 0)의 값을 가지게 된다.

HSI 색상 공간은 색상, 채도, 그리고 명도라는 세 가지 특성들이 컬러를 설명하는데 사용된다. H는 Hue를 나타내는 것으로 순수한 색상을 묘사하는 정보이고, S는 Saturation이며 상호적인 순수성이거나 hue에 혼합된 흰색 빛의 합계로 정의된다. HSI 색상 공간을 사용할 때, 어떤 컬러를 만들어 내기 위해서 몇 퍼센트의 파란색이나 녹색이 필요한지 알 필요가 없다. 진한 빨간색을 분홍색으로 바꾸기 위해서는 단순히 채도를 조절하며 어두운 것을 밝게 하려면 명도를 변경한다.

본 논문에서 사용한 YCbCr은 컬러 정보로부터 명도를 분리하는 또 하나의 컬러 공간이다. Y는 휘도(luminance) 성분분, CbCr은 색상(chrominance) 성분분을 나타낸다. 명도는 Y로 기호화되고, 푸른 정도와 붉은 정도는 Cb와 Cr로 기호화된다. YCbCr 값은 RGB 값으로부터 다음 식을 이용하여 변환된다.

$$Y = 0.29900R + 0.58700G + 0.11400B$$

$$C_b = -0.16874R - 0.33126G + 0.50000B \quad (1)$$

$$C_r = 0.50000R - 0.41869G - 0.08131B$$

본 논문에서는 YCbCr 색상 공간을 사용하여 색상 분할을 하였으며 그 이유는 사람의 살색 정보를 생성하기 위해 chrominance 성분의 효율적인 사용이 YCbCr 색상 공간에서 이루어질 수 있기 때문이다. 즉, 명도에 영향을 받지 않고, 색상 성분만을 뽑아낼 수 있다는 장점이 있기 때문이다. 실험에 의해 YCbCr과 YIQ, YUV 중에 살색의 범위가 넓고 살색을 잘 표현하는 색상 공간이 YCbCr이었다. 그래서 YCbCr을 사용하게 되었다.

2.2 관련 연구

RGB 색상 모델을 이용한 방법으로 김남호 등이 제안한 컬러 비디오 시퀀스 상에서 색상 정보와 움직임 정보를 복합적으로 사용하였다. 색상 정보는 정규화된 색상 모델을 사용하였으나 CCD 카메라 입력이라는 제한된 조건 하에서 제안한 방법이다^[1]. Li는 비디오 클립에서 얼굴 영역을 찾아내어 내용 검색을 용이하게 하기 위한 전처리 방법으로 색상 정보를 이용하여 사람 얼굴을 찾아내는 방법을 제안하였다. 이 방법은 사람의 얼굴의 크기가 어느 정도 커야하고, RGB 모델을 바로 이용하기 때문에 조명에 의해 영향을 많이 받는 단점이 있다^[2]. Sun은 RGB 색상 모델에서 살색 정보를 가우시안 모델로 정규화 한 후에 살색 영역만을 추출하고, 얼굴의 지형적 특징을 이용해 얼굴 영역을 판단하는 방법을 사용하였으나, 얼굴의 지형적 특징을 퍼지 이론을 이용하여 구하므로 시간이 오래 걸리는 단점을 가지고 있다^[3].

HSI 색상 모델을 이용한 방법으로는 Sobotka가 hue 와 saturation 값을 이용해 발견된 영역에 타원 모양이 잘 맞는 영역을 얼굴 영역이라고 결정하는 방법을 제안하였다. 그러나 영상 내에 사람이 한 사람이고, head-shoulder 영상에만 적용하여 실험하였다^[4,5].

YCbCr 색상 모델을 이용한 방법으로는 Chai가 제안한 방법이 있다. Chai는 정지 영상에서 영상 내에 사람이 한 사람이 있고, head-shoulder의 영상일 경우에 얼굴의 위치를 찾아내는 방법을 제안하였다. 제한된 조건 하에서는 얼굴의 위치를 잘 찾아내지만, 정지 영상은 제한된 조건일 경우가 많지 않으므로, 이러한 조건이외의 경우에도 얼굴 영역을 찾아

내는 알고리즘이 필요하다^[6,7].

2.3 색상 분할

YCbCr 색상 공간에서 색상 정보는 Cb와 Cr이 가지고 있다. Cb와 Cr 성분에서, 일정한 색을 나타내는 것은 일정한 범위의 Cb와 Cr의 조합으로 나타난다. 이러한 성질을 이용해서 살색이 가지는 정보를 이용하기 위해 살색이 차지하는 범위를 찾아내어 그 범위의 값만 추출해 낸다. 살색이 차지하는 범위를 찾아내기 위해 영상에서 살색 부분만을 뽑아내어 살색이 차지하는 범위를 히스토그램으로 구한다. 이러한 히스토그램을 여러 개의 영상에 적용하여 살색이 차지하는 범위를 결정할 수 있다. 그림 2 (a)의 그림은 영상에서 살색영역만을 뽑아낸 그림이고, 그림 2 (b)는 이 컬러 영상에서 Cb 성분을, 그림 2 (c)는 Cr 성분을 뽑아 히스토그램을 그린 그림이다.

III. 얼굴 영역 추출

본 논문에서 제안한 얼굴 영역 추출 알고리즘은 살색 성분을 이용하고, 간단한 후처리 기법을 통해 얼굴의 크기, 얼굴의 자세, 얼굴의 표정, 영상의 조명 변화에 비교적 무관하게 얼굴 영역을 추출하도록 제안되었다. 또한 영상 내에 사람의 얼굴이 두 개일 경우에도 적용될 수 있는 알고리즘을 제안하였다. 제안된 알고리즘은 살색을 이용하기 때문에 동양인에 국한을 두었고, 얼굴의 크기는 최소 32 × 32 이상이어야 한다고 가정하였다. 그림 3 은 제안된 알고리즘의 전체 구성도이다.

3.1 1차 후보 영역 추출

1차 후보 영역을 추출하는 과정은 살색 정보를 이용하여 영상 내에 살색의 값을 가지는 화소들을 찾아내고, 얼굴 영역 이외의 부분에서 발생하게 되는 작은 잡음들을 제거하는 과정으로 모두 3 단계로 구성되어 있다.

3.1.1 색상 분할

이 과정에서는 색상 정보를 이용하여 빠르게 살색 영역을 추출해 낸다. 이 과정은 입력 영상에서 살색과 살색이 아닌 영역을 구분 짓는 것이 목적이다. YCbCr 색상 공간은 색상의 명암도가 Y라는 Luminance 값에 의해 나타난다. 따라서 색상 정보를 가지고 있는 Cb와 Cr의 Chrominance 값은 명암 정보가 들어 있지 않으므로 조명에 영향을 받지

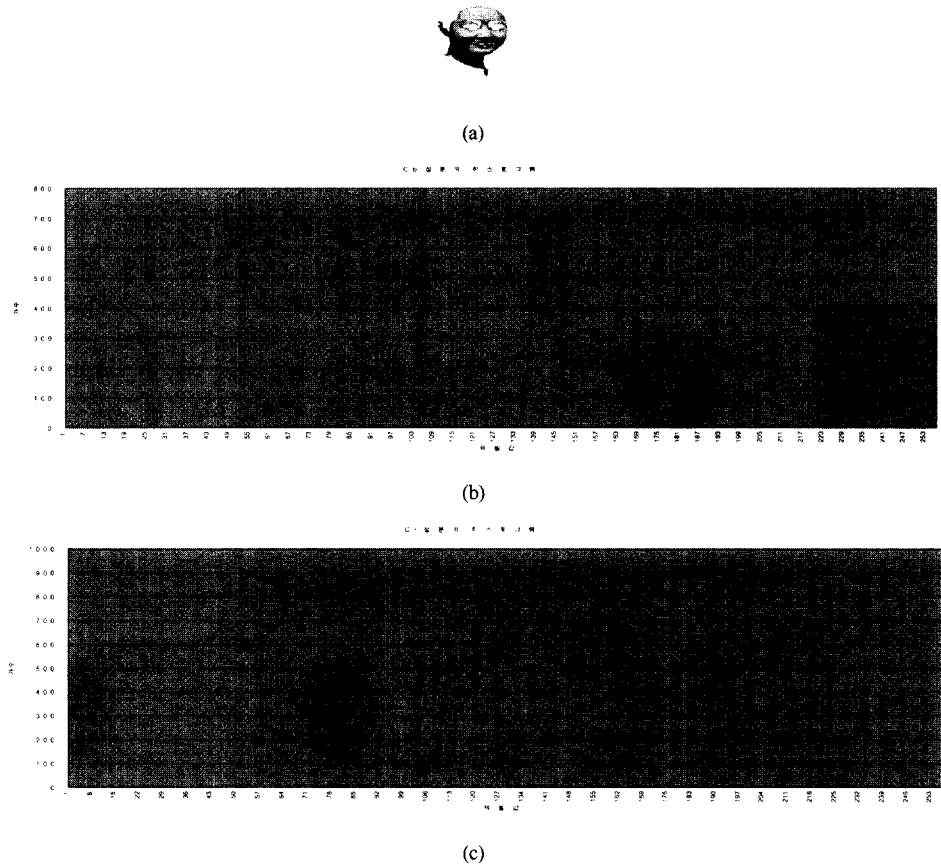


그림 2. (a)컬러 영상에서 추출된 살색영역 (b)Cb 성분의 히스토그램 (c)Cr 성분의 히스토그램

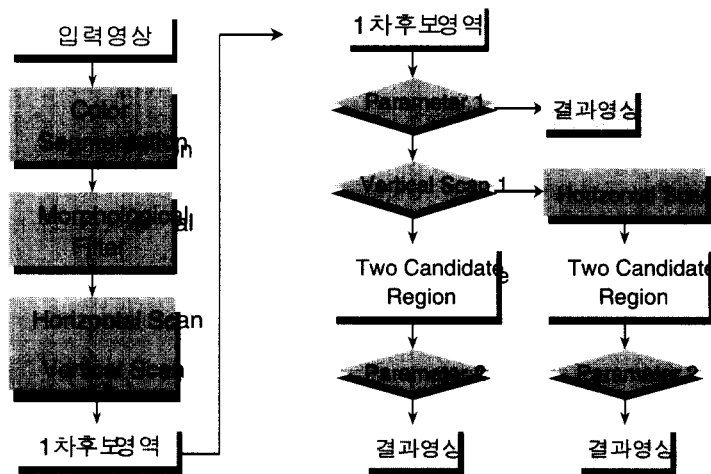


그림 3. 얼굴 영역 추출 알고리즘의 전체 구성도

않고, 살색 범위를 효과적으로 찾아낼 수 있다^{6,7)}.

색상 분할 시에는 색상 정보를 가지고 있는 Chrominance 값만을 사용하며, 다음 식을 이용하여 Cb와 Cr 성분에서 살색인 부분은 1로, 살색이 아닌 부분은 0으로 바꾼다.

$$f_c(x, y) = \begin{cases} 1 & \text{if } (133 < Cb < 173) \cap (77 < Cr < 127) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

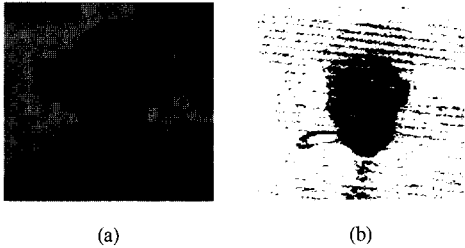


그림 4. (a) 입력 영상 (b) 색상 분할 과정을 거친 이진 영상

그림 4에 이 과정의 결과로써 얻은 이진 영상을 보였다. 이 과정은 매우 중요한 부분으로 사람의 살색 모델을 기초로 하여 살색이 아닌 부분의 많은 화소들을 제거할 수 있다. 나머지 과정들은 이 과정에서 남아있는 원하지 않은 화소들을 제거하는 과정이다.

3.1.2 형태학적 필터

형태학은 경계 골격, 볼록과 같은 영역 형태를 표현하거나 서술하는데 있어서 유용한 영상요소들을 추출하기 위한 도구로 수학적 형태론의 개념으로 사용된다. 이 기법은 영상 전처리 작업이나 초기 객체 분류 또는 이러한 처리 뒤에 이어서 물체의 내재된 구조를 명확히 하는데 이용된다. 이것은 물체의 외곽선을 물체의 가장 현저한 한 화소 두께의 외곽선이나 골격선을 간단하게 함으로써 이루어진다. 다시 말하면, 최외각선의 한 화소를 빼던가, 더하던가 하는 작업을 말한다. 그러므로 한 화소의 잠음은 제거될 수 있고, 두 화소로 확장될 수도 있다.

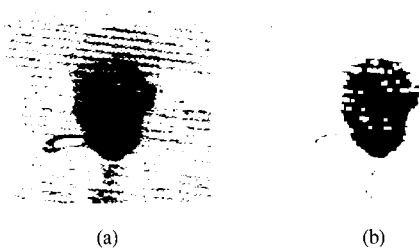


그림 5. (a) Closing 연산 과정의 입력 이진 영상 (b) Closing 연산 과정을 거친 영상

Closing 연산은 dilation 연산 후 erosion 연산 과정을 거친다. 즉, 원 영상에서 dilation을 하여 검은색 물체의 크기를 축소하여 작은 잠음들을 제거하고, erosion 연산을 하여 영상을 원래의 크기로 돌아오게 하는 방법이다. 그림 5 은 Closing 연산 과정을 영상에 적용하여 보여 준 그림이다.

3.1.3 수평, 수직 방향 스캔

Closing 연산을 거친 이진 영상, 여러 개를 분석하여 특징을 찾아내어 일정한 임계값을 정하였다. 하나의 열을 수평 방향으로 스캔한 후에 검은색 화소의 개수가 임계값 이상 일 경우에는 그 열을 그대로 보존하고 임계값 이하 일 경우에는 그 열에 있는 모든 화소를 모두 흰 색이 되도록 변환한다. 수직 방향 스캔도 수평 방향 스캔과 마찬가지로 한 행에 검은색 값을 가지는 화소들의 개수가 임계값 이상 일 경우에는 그 열을 모두 흰 색의 값을 가지는 화소로 변환한다.



그림 6. (a) 수평, 수직 방향 스캔의 입력 영상 (b) 수평, 수직 방향 스캔 후의 영상

그림 6 (a) 그림은 closing 과정을 거친 영상을 보여 주고, 그림 6 (b)는 그림 6 (a)의 영상을 가지고 수평, 수직 방향으로 스캔한 경우에 나오는 결과 영상이다.

3.2 타원 넓이를 이용한 임계값

1차 후보 영역은 얼굴 영역일 수도 있고 아닐 수도 있다. 이것을 판별하는 임계값이 타원 넓이를 이용한 임계값이다. 이 임계값은 1차 후보 영역이 얼굴 영역이라면 얼굴이 타원이라는 특징을 이용해 사각형 안에 타원이 차지하는 비율을 구한 것이다⁸⁾.

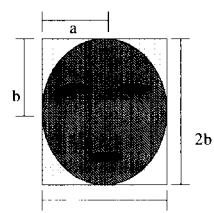


그림 7. 사각형 넓이와 타원 넓이

그림 7 은 가로, 세로 길이가 각각 $2a$ 와 $2b$ 인 사각형 안에 장축이 b 이고 단축이 a 인 타원의 관계를 보인 것이다. 사각형과 타원의 넓이는 각각 식 (3)와 (3)과 같이 표현된다.

또한 눈과 입 부분은 살색이 아니므로 그림에서 보는 것처럼 타원의 넓이에서 식 (5)에서 구한 눈과 입 부분의 넓이를 제외시켰다. 따라서 식 (6)에서 구한 것처럼 사각형 안에서 얼굴 영역이 차지하는 비율은 대략 76.5%가 된다.

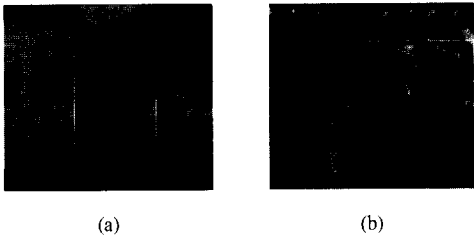


그림 8. (a) 1차 후보 영역이 얼굴 영역인 영상 (b) 2차 후보 영역을 찾아야하는 영상

그림 8 은 임계값을 넣어 계산한 결과 영상들로 그림 8 (a)의 영상은 1차 후보 영역이 얼굴 영역이 되는 경우이고, 그림 8 (b)의 영상은 사각형 안에 검은색 화소의 비율이 임계값에 미치지 못하여 2차 후보 영역을 찾는 과정을 거쳐야 하는 경우이다.

3.3 2차 후보 영역 추출

1차 후보 영역을 수직방향으로 스캔하여 한 열씩 그 열의 각각의 화소의 값을 다 더한 것이 0이 되는 열이 존재하는지 여부에 따라 다음의 두 가지 경우로 나누었다.

① 존재하지 않는 경우 : 수평 방향으로 스캔하여 2개의 후보 영역으로 나눈다. 각 후보 영역이 얼굴인지는 2개의 매개변수를 가지고 판단한다. 두 개의 매개변수는 가로대 세로 비와 후보 영역 안에 살색이라고 판별된 화소의 수로 정의된다.

② 존재하는 경우 : 수직 방향으로 스캔하여 2개의 후보 영역으로 나눈다. 이 2개의 후보 영역 중에서 얼굴인지 아닌지를 판별하는 방법은 바로 위에서 설명한 방법과 같다.

그림 9 은 이러한 과정을 거쳐서 두 개의 후보 영역을 구한 영상이다.

탐색해낸 영역이 얼굴인지 아닌지를 판단하기 위해 두 개의 임계값을 사용한다. 첫 번째 임계값은 앞에서 설명한 타원의 넓이를 이용한 것이고, 두 번째 임계값은 사람의 얼굴이 일반적으로 세로가 가

로에 비해 1.5배 크다는 점을 이용하여 정하였다. 이 두 번째 임계값은 일반적으로 공인된 사실을 참고로 하여 가지고 있는 영상 데이터에서 실험적으로 구한 값이다. 그림 10 (a)는 두 개의 후보 영역을 나눈 그림이다. 그림 10 (a)의 왼쪽 사각형은 얼굴이고, 오른쪽 사각형은 얼굴이 아니지만, 그림 10 (b)에서는 2차 후보 영역 중 하나로 찾아내었다. 그림 10 (b)는 이러한 임계값을 적용하여 왼쪽의 얼굴 영역만을 찾아낸 결과 영상이다.

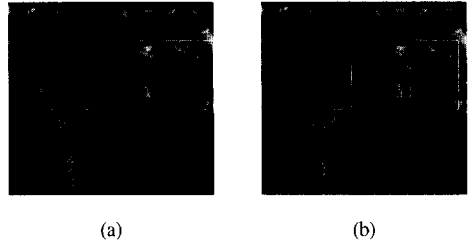


그림 9. (a) 영역 분리 전 그림 (b) 영역 분리 후의 그림

IV. 실험 결과

서로 다른 특징을 가지고 있는 정지 영상에 대하여 얼굴 영역을 추출하는 실험을 하여, 알고리즘의 성능을 검사하였다. 실험은 펜티엄 MMX 200MHz의 CPU와 메인 메모리 64MB의 윈도우즈 98 시스템 환경에서 Visual C++ 6.0을 이용하여 구현하였다.

4.1 제한한 알고리즘의 실험결과

제한한 알고리즘은 동양인일 경우에만 적용하였다. 실험에 사용된 데이터는 표1과 같은 특징을 가지고 있다. 실험 데이터는 다양한 얼굴 크기, 정면 영상, 고개를 돌린 영상, 고개를 기울인 영상, 고개를 숙이거나 든 영상, 다양한 배경, 다양한 표정을 가진 영상들이다. A 영상군은 사진을 스캔한 디지털 데이터로 일반 CCD 카메라나 텔레비전 캡춰 영상보다 영상이 선명하고, 색상이 다양한 분포를 가지고 있고, 대부분의 정면 영상이다. 그에 비해 B 영상군은 영화나 드라마를 홍보하는 사이트에서 제작한 사진들을 실험 데이터로 사용했으므로 사람의 수가 두 명 이상의 경우가 많으며 A 영상군보다는 배경이 더 선명하게 나온 사진들을 묶은 영상군이다. C 영상군은 실제로 화질이 깨끗하지 않은 텔레비전을 보면서 캡춰한 영상으로 영상 내에 잡음이 많이 섞여 있는 영상을 말한다. 이러한 영상 데이터

의 입력 형태는 256*256 크기의 RGB 색상 공간을 가지는 영상으로, 본 논문에서 제안한 알고리즘을 적용하기 위해 RGB 색상 공간을 YCbCr 색상 공간으로 바꾸었다.

실험에 영향을 미칠 수 있는 파라미터의 값은 표 2와 같다. 1차 후보 영역 추출 시 수직, 수평 방향 스캔 시에 한 행과 한 열에 0 값의 개수의 합을 구하여 최소의 값을 실험적으로 정했다. 수평 임계값이 32인 것은 얼굴의 최소 크기에 맞춘 것이고, 수직 임계값은 수평 임계값을 적용한 결과에 수직 임계값을 적용한 것이다. 따라서 아래 그림처럼 수평 임계값을 적용한 결과가 나왔을 때도 적용 가능해야 하므로 수직 임계값은 실험적인 결과에 의해 15로 정하였다. 1차 후보 영역을 추출하고, 1차 후보 영역이 얼굴 영역일 경우와 2차 후보 영역을 추출해야 하는 경우로 판별하는 파라미터로 타원의 넓이를 이용한다. 타원의 넓이는 사각형의 약 78.5%로 차지하고 눈, 입 부분은 살색이 아니므로 이 부분까지 제외하면, 사각형안에 살색 영역은 약 76.5%의 비율을 차지하게 된다. 이것을 좀 더 유연하게 적용하기 위해서 76.5%에 ±3%의 범위를 가지도록 실험적으로 정했다. 가로세로 비율은 사람의 얼굴은 세로가 가로에 비해 약 1.5배 정도이므로 실험적으로 1에서 2사이의 값을 가지면 얼굴이라고 판단하도록 하였다.

얼굴의 크기에 관계없이 얼굴 영역을 찾아낸 경우로 영상 내에 사람이 한 명이 있을 경우나, 두 명이 있을 경우에 관계없이 얼굴 영역을 찾아내며, 두 명이 있을 경우 두 명의 얼굴 크기가 틀려도 각각을 찾아내는 결과를 보였다. 또한 영상의 화질에 관계없이 얼굴 영역을 찾아낸다. 화질이 좋은 사진 영상과 화질이 좋은 많은 텔레비전 화면을 캡처한 영상의 두 가지의 경우를 실험했을 때 화질이 좋은 경우와 화질이 좋지 않은 두 경우 모두에서 얼굴 영역을 찾아내는 결과를 보였다. 고개의 움직임에 관계없이 얼굴 영역을 찾아내어 제안한 알고리즘이 고개의 위치에 따라 영향을 받지 않음을 보여준다. 사람의 얼굴 표정에 무관하게 얼굴 영역을 찾아낸다.

고유 얼굴을 이용한 형판 정합 방법은 학습 비율에 의해 영향을 받고, 탐색에 성공해야 얼굴 영역을 추출하게 된다. 또한 형판 정합 방법은 추출 시간이 오래 걸리는 단점을 가지고 있으며, 얼굴의 기울어짐, 얼굴의 크기에 영향을 많이 받으며, 이러한 경우에 얼굴 영역을 추출하려면 다른 연산 과정이 추

가로 필요하다. 학습율이 높은 경우 90%까지의 결과를 보였으나, 학습율이 낮은 경우에는 72%의 낮은 성공률을 보였다. 그러나 본 논문에서는 제안한 알고리즘의 경우 찾아내지 못하는 경우는 거의 없고, 대신 얼굴 영역 이외의 부분까지 포함하여 찾아지는 경우만 있다. 그러나 본 논문에서 제안한 알고리즘의 경우 얼굴의 기울어짐, 얼굴의 크기에 관계없이 얼굴 영역을 추출하며, 찾아내지 못하는 경우는 거의 없고, 대신 얼굴 영역 이외의 부분까지 포함하여 찾아지는 경우가 있다.

YCbCr 색상 정보를 이용한 다른 방법은 흑인, 백인, 동양인에 모두 적용되는 알고리즘으로 약

표 1. 실험에 사용된 영상

영상 분류	장 수			특 징
	총	한명	두명	
A 영상군 (사진)	42	33	9	화질이 좋다 다양한 배경 대부분이 정면 영상 단순한 표정
B 영상군 (영화와 드라마)	45	18	24	배경이 복잡하다. 다양한 포즈와 표정 사람의 수가 두 명 이상인 경우
C 영상군 (텔레비전 캡처)	54	51	3	화질이 좋지 않다 다양한 표정 다양한 배경 다양한 얼굴 크기

표 2. 실험에 영향을 미치는 파라미터

파라미터	사용된 값	비 고
살색의 임계값	133<Cb<173 77<Cr<127	히스토그램의 분포를 이용해 나온 화소의 값
수평, 수직방향 스캔 시 임계값	a=32 b= 15	한 줄에 살색이라고 판단된 화소의 개수 (실험적으로)
타원 넓이의 임계값	73.5 - 79.5	얼굴 영역이 사각형안에서 차지하는 비율
가로세로 비	1-2	일반적인 사람들 얼굴의 가로 세로 비

표 3. 얼굴 영역 추출율

영상 종류	총 (장)	제대로 찾은 영상 (장)	잘못 찾은 영상 (장)	추출율 (%)
사진	42	36	7	85.4
영화와 드라마	48	41	6	85.7
텔레비전 캡처	55	48	7	87.3
합계	145	125	20	86.2

표 4. 2인일 경우와 1인일 경우 추출율 비교

영상 분류	총(장)	제대로 찾은 영상(장)	잘못 찾은 영상(장)	추출율 (%)
영상 내에 얼굴이 한 개인 경우	106	93	13	87.7
영상 내에 얼굴이 두 개인 경우	39	32	7	82.1

82%의 결과가 나왔으나 이 방법은 제안한 알고리즘에 비해 어느 정도 균일하거나 배경이 없는 영상들에 적용한 방법이고, 영상 자체를 head-shoulder 영상으로 제한을 걸고 얼굴 영역을 추출하였다.

RGB와 움직임 정보를 이용한 방법은 동영상에 적용되는 알고리즘으로 일정한 배경에서 사람이 움직일 때 얼굴을 찾아내어 얼굴의 특징점을 추출하는데 약 94%의 성공률을 보였으나 이 방법의 단점은 배경이 바뀌지 않는다는 것이다.

다른 방법들과 얼굴 영역 추출율은 비슷하나, 얼굴의 기울어짐, 고개를 좌우로 돌림, 고개를 숙이거나 든 영상, 얼굴의 크기에 관계없이 얼굴 영역을 추출해 낼 수 있다는 장점이 있다.

V. 결론 및 향후 연구 계획

본 논문에서는 다양한 배경 및 다양한 조건을 가지는 정지 영상에서 동양인 기준의 살색 정보를 이용하여 얼굴 영역을 추출하는 알고리즘을 제안하였다. 제안된 알고리즘에서는 형판 정합의 방법과 같이 얼굴의 크기에 무관하게 하기 위해서 다른 연산의 과정을 거치지 않고, 살색을 가진 성분만을 추출하여 사용하기 때문에 얼굴의 크기, 얼굴의 표정변화, 고개가 돌아간 정도, 고개를 숙인 정도, 고개를 든 정도, 얼굴의 기울어짐에 관계없이 얼굴 영역을 찾아내도록 하였다.

실험에서 사용한 데이터는 영상의 화질, 인물의 몸 전체의 자세, 얼굴의 자세에 따른 3가지 영상군으로 나누어 실험의 다양성을 기했다. A 영상군의 경우 비교적 정면의 얼굴의 영상이 많고, 화질이 좋기 때문에 B 영상군에 비해 높은 추출율을 보였으며, B 영상군의 경우 2인일 경우의 영상이 다른 영상군에 비해 많음에도 불구하고 다른 영상군과 비슷한 결과를 보임으로써 본 논문에서 제안한 알고리즘이 2인 이상인 경우의 영상에도 적용할 수 있음을 보여준다. C 영상군의 경우에는 텔레비전의 화면을 캡처한 영상들의 집합으로 A, B 영상군에

비해 화질이 좋지 않으나, 높은 얼굴 영역 추출율을 보여 화질이 좋지 않은 영상에도 본 논문에서 제안한 알고리즘이 잘 적용됨을 보여준다. 그러나 머리카락의 색이 밝은 갈색에 가까운 경우 머리카락도 살색으로 판단하여 얼굴 영역보다 큰 영역을 추출하고, 2인 이상일 경우 얼굴이 겹쳐져 있는 경우에는 얼굴이 하나있는 영상으로 얼굴 영역을 추출해 내는 경우가 발생하였다.

본 논문에서 제안한 알고리즘은 얼굴인식, 표정인식 등의 전처리 단계로 응용되거나 현재 많은 데이터들이 디지털 화되면서 영상 데이터들을 검색하는 시스템에 효율적으로 사용될 것으로 기대된다.

향후 과제로는 3인 이상의 영상에서도 얼굴 영역을 추출할 수 있도록 알고리즘을 최적화하는 것과 다양한 살색, 즉 백인, 흑인, 동양인 모두에게 광범위하게 적용될 수 있도록 알고리즘을 구현하는 것이다.

참 고 문 헌

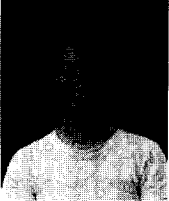
- [1] 김남호 외, "색상 움직임을 이용한 얼굴 특징점 자동 추출," *전자공학회지논문지*, 제 35권 S편 제 8호, p 1077-1089, 1998
- [2] Y. Li, B. Tao, S. Kei and W. Wolf, "Semantic Image Retrieval Through Human Subject Segmentation and Characterization," *SPIE Vol. 3022*, p.340-351, 1997
- [3] Q. B. Sun, W. M. Huang and J. K. Wuosm, "Face Detection Based on Color and Local Symmetry Information," *International Conference on Automatic Face and Gesture Recognition*, pp. 130-135, 1998
- [4] K. Sobottka and I. Pitas, "Extraction of Facial Regions and Features Using Color and Shape Information," *ICPR '96*, p. 421-425, 1996
- [5] A. Eleftheriadis and A. Jacquin, "Automatic face location, detection and tracking for model-assisted coding of video conferencing sequences at low bit-rates," *Signal Processing : Image Communication* 7(3), p. 231-248, 1995
- [6] D. Chai and K. N. Ngan, "Locating Facial Region of a Head-and-Shoulders Color Image," *International Conference on Automatic Face and Gesture Recognition*, p. 124-129, 1998
- [7] D. Chai and K. N. Ngan, "Face Segmentation

Using Skin-Color Map in Videophone Application," IEEE. Trans. Circuits and Systems for Video Tech, p. 551-564, June 1999

[8] <http://forum.swarthmore.edu/dr.math/faq/formulas/faq.ellipse.html>

장 선 아(Sun-Ah Chang)

정회원



1998년 2월 : 홍익대학교

전자공학과 졸업 (학사)

2000년 2월 : 광운대학교 대학원

전자공학과 졸업 (석사)

2000년 2월~현재 :

(주)노뜨르 근무

<주관심 분야> 영상인식, 컴퓨터 비전

유 지 상(Jisang Yoo)

정회원



1985년 2월 : 서울대학교

전자공학과 졸업(공학사)

1987년 2월 : 서울대학교 대학원

전자공학과 졸업(공학석사)

1993년 5월 : Purdue대학교

전기공학과 졸업(Ph.D.)

1993년 5월~1994년 8월 : 현대전자산업(주) 산전연구
구소 선임연구원

1994년 9월~1997년 8월 : 한림대학교 전자공학과
조교수

1997년 9월~현재 : 광운대학교 전자공학과 조교수

<주관심 분야> 웨이블릿 기반 영상처리, 영상압축,
영상인식, 비선형 신호처리