

# 동영상에서의 색상 정보와 차영상을 이용한 얼굴 영역 추출에 관한 연구

박형철, 전병환  
공주대학교 전자계산학과

## The Extraction of Face Regions in Dynamic Image Using Color Information and Difference Images

Hyung Chul Park, Byung Hwan Jun  
Dept. of Computer Science, Kongju National University

### 요 약

동영상에서의 얼굴 영역 추출은 헤드 제스처 인터페이스를 위한 기본적인 필수적인 기법이다. 얼굴 영역 추출을 위해서는 색상 정보와 차영상을 이용한 방법이 많이 사용되며, 색상 정보를 이용하는 방법에는 HSI의 H(hue)성분과 YIQ의 I(in-phase)성분이 널리 알려져 있다. 본 논문에서는 먼저 얼굴 색상에 해당하는 각 색상 성분의 구간을 탐색하고, 다음으로 각 색상 정보를 이용한 얼굴 영역 검출의 정확도를 비교 실험한다. 또한, 색상 정보와 차영상을 결합한 방법에 대해서도 얼굴 영역 검출의 정확도를 비교한다. 실험 결과, YIQ의 경우 구간 130~150, HSI의 경우 구간 0~20에서 얼굴색을 잘 표현하는 것으로 나타났다. 얼굴 영역 검출의 정확도 측면에서는, 색상 정보만을 이용한 실험의 경우 YIQ가 HSI에 비해 약 10%의 향상된 성능을 보였고, 색상 정보와 차영상을 결합한 경우에서도 YIQ가 약 5%의 향상된 성능을 보였다.

### 1. 서 론

헤드 제스처 인터페이스는 얼굴의 방향이나 눈의 깜빡임, 혹은 고개의 움직임 등의 정보를 이용하여 보다 자연스럽고 편리한 사용자 인터페이스를 구현하기 위한 연구분야이다. 이를 위해서는 사람의 얼굴 영역을 보다 정확히 추출해 내는 연구가 선행되어야 한다. 얼굴 영역을 추출하는 방식은 크게 두 가지로 나눌 수 있다. 첫째, 정지영상의 경우, 에지 추출, 형판 정합, 타원 정합 등의 방법을 이용하고[1], 동영상의 경우, 차영상이나 색상 정보를 이용한 방법 등이 있다[2,3,4,5].

본 연구에서는 얼굴 색상에 해당하는 색상값의 구간을 찾는 방법에 대해 연구하고, 색상 정보와 차영상을 함께 적용하여 얼굴 영역을 추출하는 방법에 대해 연구한다.

### 2. 얼굴 영역 추출 방법

#### 2.1 색상 정보 : HSI와 YIQ

색상 정보를 이용하는 방법으로 색상(hue), 채도(saturation), 명도(intensity)의 세 성분으로 구성되는 HSI를 고려할 수 있다[5]. RGB를 HSI로 변환하는 식은 다음 식(1)과 같다.

$$I = \frac{1}{3}(r + g + b)$$

$$S = 1 - \frac{3}{(r + g + b)} [Mm(r, g, b)] \quad (1)$$

$$H = \cos^{-1} \left\{ \frac{\frac{1}{2}[(r-g) + (r-b)]}{\sqrt{(r-g)^2 + (r-b)(g-b)}} \right\}$$

또한, YIQ는 TV 전송에 사용되는 칼라모델로 Y(luminosity) 혹은 흑백 디스플레이 장치 상에 두 개의 다른 색들을 구별하는데 이용), I(in-phase 오렌지색(orange)-청록색(cyan) 색상 정보를 포함), Q(quadrature 녹색(green)-자홍색(magenta) 색상 정보를 포함)로 나타낸다[4]. 변환식은 다음 식(2)와 같다.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

본 연구에서는 HSI의 H(hue)성분과 YIQ의 I(in-phase)성분을 이용하여 얼굴 영역을 추출한다.

#### 2.2 차영상

시간적으로 연속된 영상들을 이용하여 인접 영상간의 명암차에 의해 차영상을 추출한다[2].

2.3 색상 정보와 차영상의 결합

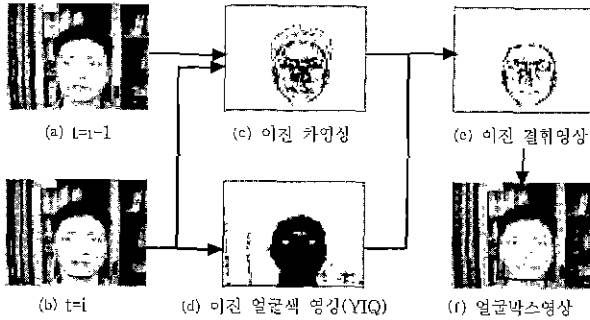


그림 1 얼굴 영역 추출 과정

그림 1은 얼굴 영역을 추출하는 과정이다. 시간적으로 연속된 원영상 (a), (b)를 이용하여 이진 차영상 (c)를 얻어내고, 원영상 (b)로부터 이진 얼굴색 영상 (d)를 추출한다. (c)와 (d)를 결합 (AND)하여 이진 결합영상 (e)를 얻을 수 있고, 마지막으로 얼굴 영역이 추출된 얼굴박스영상 (f)를 얻을 수 있다.

3. 실험 및 고찰

3.1 실험 환경

본 실험은 HSI와 YIQ의 얼굴 영역 추출의 정확도를 비교하기 위한 실험이다. 실험을 위해 한 사람으로부터 실내 환경에 따른 다양한 영상을 320×240 화소 크기로 취득하여 사용하였다. 영상 데이터는 일반 사무실에서 흔히 볼 수 있는 복잡한 배경을 갖고 컴퓨터 화면을 주시하는 얼굴 영상을 대상으로 한다. 이때, 창문을 통한 자연광의 유무나 방향에 따라 다음의 4가지 경우로 구분하여 영상 데이터를 취득하였다

- A 영상 : 정면에서 자연광이 들어오는 경우
- B 영상 : 자연광을 커튼으로 가린 경우
- C 영상 : 측면에서 자연광이 들어오는 경우
- D 영상 : 후면에서 자연광이 들어오는 경우(역광)

실험을 위해 Pentium MMX 200MHZ PC, Telegeny 캡처 보드, 그리고 CCD 카메라를 이용하였고, Visual C++ 5.0을 사용했다.

3.2 실험 방법

3.2.1 얼굴 색상의 한계 구간 설정

H(hue)는 색상값 0~360을 갖고, I(in-phase)는 색상값 -152~152를 갖는다. 본 논문에서는 이들을 동일하게 0~255로 정규화하여 사용한다. H 성분과 I 성분의 한계구간을 설정하기 위해서는 각 성분에 대해 최고 경계치와 최저 경계치를 결정해야 한다. 그러나 이 두 경계치를 동시에 찾는 것은 지나치게 많은 연산을 요구한다. 따라서, 이 논문에서는 먼저 하나의 경계치 ( $\theta_I$ )에 대해 영상을 이진화해 가면서 가장 원영상의 얼굴 부분과 일치하는 값을 탐색한다. 이때, 얼굴 부분은 원영상에서 얼굴과 목 부분만을 검게 칠하고 나머지 부분은 하얗게 칠하여 얻을

수 있다. 그런 다음, 최고치를 중심으로 한 여러 구간에 대해 얼굴색 영역 추출의 정확도를 측정하여 그 값이 최대가 되도록 최종 구간을 확정한다. 얼굴색 영역 추출의 정확도를 측정하기 위해 다음과 같은 0과 1사이의 값을 갖는  $\alpha, \beta, \gamma$ 를 사용한다

$$\alpha = \frac{|F_s \cap F_d|}{|F_s|}$$

$$\beta = \left(1 - \frac{|F_s \cap F_d^c|}{|F_s^c|}\right)$$

$$\gamma = \frac{\alpha + \beta}{2}$$

$F_s$ : 찾아진 얼굴 영역

$F_d$ : 실제의 얼굴 영역

$\alpha$ 는 원영상의 얼굴 영역내에서 얼굴이라고 찾은 영역의 비율이고,  $\beta$ 는 얼굴을 제외한 영역에서 실제로도 얼굴이 아니라고 찾은 영역의 비율이다.  $\gamma$ 는  $\alpha$ 와  $\beta$ 의 평균값으로서 최종적인 얼굴 영역 추출의 정확도를 의미한다.

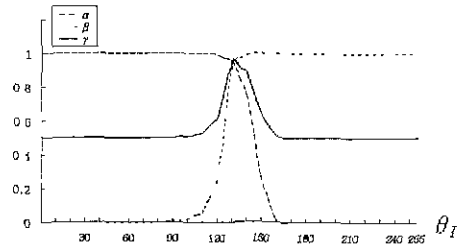


그림 2 영상 A에 대해 경계치  $\theta_I$ 에 따른  $\alpha, \beta, \gamma$ 의 변화

영상 A에 대한 I 성분의 경계치  $\theta_I$ 의 탐색 결과, [그림 2]의 값이  $\theta_I = 130$  일 때  $\gamma$ 가 최대( $\max(\gamma(A)) = 0.95$ )가 됨을 알 수 있다. 다른 영상 B, C, D에서도 일관되게  $\theta_I = 130$ 에서  $\gamma$ 가 최대가 된다( $\max(\gamma(B)) = 0.95, \max(\gamma(C)) = 0.91, \max(\gamma(D)) = 0.87$ ). H성분의 경우에도, 영상 A, B, C, D에 관계없이 일관되게  $\theta_H = 20$  일 때  $\gamma$ 가 최대가 된다.

그런 다음, 각각의  $\theta_I$ 와  $\theta_H$ 에 대해 5씩 증감하면서 그 주변의 경계 구간에 대해  $\alpha, \beta, \gamma$ 를 조사하였다. 그 결과, H 성분의 경우에는 경계구간 0~20, I 성분의 경우에는 경계구간 130~150이 얼굴 색상을 찾는 최적의 구간으로 판명되었다

3.2.2 색상 정보에 의한 얼굴 영역 추출

HSI와 YIQ의 색상 정보 추출 능력을 비교하기 위해 앞에서 설정된 경계구간을 이용하여 각 영상 A, B, C, D에 대해  $\alpha, \beta, \gamma$ 를 측정하면 각각 표 I-1, 표 I-2와 같다.

	A	B	C	D	평균
$\alpha$	0.78	0.86	0.57	0.64	0.71
$\beta$	0.83	0.81	0.80	0.76	0.81
$\gamma$	0.81	0.85	0.69	0.70	0.76

<표 I-1> HSI의 경우

	A	B	C	D	평균
$\alpha$	0.63	0.83	0.87	0.89	0.81
$\beta$	0.96	0.94	0.90	0.85	0.91
$\gamma$	0.80	0.89	0.89	0.87	0.86

<표 I-2> YIQ의 경우

실험 결과, YIQ를 이용한 방법이 HSI에 비해  $\alpha, \beta, \gamma$  모두 0.1(10%)정도의 향상된 성능을 보였다.

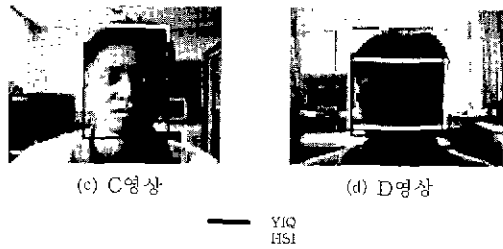


그림 3 색상 정보와 차영상을 이용해 추출된 얼굴 영역

3.2.3 색상 정보와 차영상의 결합에 의한 얼굴 영역 추출  
 [그림 1]에서 보인 바와 같이, 이진 차영상과 이진 얼굴색 영상으로부터 이진 결합영상을 얻을 수 있다. 이 영상을 각각 x축과 y축으로 투영(projection)시켜 얼굴 영역 분포를 나타내는 히스토그램(histogram)을 작성하고, 이를 실험에 의해 구한 경계치( $\phi$ )에 의해 이진화한 다음, 해당 얼굴 영역을 둘러싸는 박스를 구한다. 이렇게 구한 얼굴 박스에 대해 얼굴 영역 추출의 정확도를 평가하기 위해서, 차영상에서 추출되지 않는 목을 제외하고 남은 얼굴 부분만을 겹쳐 칠한 이진 원영상과 비교한다. 히스토그램을 살펴보면, 배경의 일부 작은 영역이 얼굴 영역에 해당하는 것으로 잘못 판단되는 경우가 발생하는데, 이를 해결하기 위해 다음과 같은 휴리스틱 정보를 사용한다.

- 영상의 가장자리에서부터 시작해서  $\phi$  이상이 되는 첫 구간이 극히 좁은(3 pixel 이내) 구간이고,
- 이 구간으로부터 인접한 다음 구간이  $\frac{2}{3}\theta$ 보다 작으면서 넓게(30 pixel 이상) 나타나는 경우, 잡음으로 간주하여 얼굴 영역이 아닌 것으로 판단

얼굴 영역 추출의 정확도를 평가하기 위해 앞에서 정의한  $\alpha, \beta, \gamma$ 와 유사하게 다음과 같은  $\alpha', \beta', \gamma'$ 을 사용한다.

$$\alpha' = \frac{|B_s \cap B_d|}{|B_s|}$$

$$\beta' = \left(1 - \frac{|B_s \cap B_d^c|}{|B_s^c|}\right)$$

$$\gamma' = \frac{\alpha' + \beta'}{2}$$

$B_s$ : 찾아진 얼굴 박스  
 $B_d$ : 실제의 얼굴 박스

위의 정의된  $\alpha', \beta', \gamma'$ 을 이용하여 다음과 같은 결과를 얻었다.

<표 II-1> HSI + 차영상의 경우

	A	B	C	D	평균
$\alpha'$	0.94	0.79	0.98	0.70	0.85
$\beta'$	0.99	0.93	0.93	0.98	0.97
$\gamma'$	0.97	0.89	0.96	0.84	0.91

<표 II-2> YIQ + 차영상의 경우

	A	B	C	D	평균
$\alpha'$	0.99	0.97	0.98	0.83	0.94
$\beta'$	1.0	1.0	0.98	0.99	0.97
$\gamma'$	1.0	0.99	0.98	0.91	0.96



(a) A영상



(b) B영상

실험 결과, YIQ의 색상 정보에 차영상을 결합시켜 추출된 영상이  $\alpha'$ 의 경우 0.1(10%)정도 향상된 성능을 보이고,  $\beta'$ 의 경우는 거의 동일하게 나타났다.  $\gamma'$ 은 0.05(5%)정도의 성능 향상을 보인다 [그림 3]은 각 영상에서 추출된 얼굴 영역의 박스를 나타낸다.

#### 4. 결론

본 논문은 헤드 제스처 인터페이스를 위한 기초 연구로 HSI와 YIQ의 색상구간을 탐색하여 경계구간을 설정하였으며, 색상 정보만을 이용한 얼굴 영역 추출 방법과 색상 정보에 차영상을 결합한 얼굴 영역 추출 방법에서 두 색상 정보에 대해 얼굴 영역 추출의 정확도를 비교 평가하였다. 실험결과, HSI의 구간은 0~20, YIQ는 130~150이 얼굴색을 표현하는 최적의 구간으로 판명되었다. 또한, 색상 정보만을 이용한 방법에서는 YIQ를 이용한 방법이 HSI를 이용한 방법보다 10%정도 얼굴 영역을 더 정확히 추출하였으며, 차영상과 색상 정보를 함께 이용한 경우에서도 YIQ 적용방법이 5%정도 높게 나타났다. 한편, 처리시간에서는 YIQ의 경우 약 0.15초, HSI의 경우 약 0.55초로 나타났다. 향후, 조명의 영향, 배경에서의 움직임 발생 등의 주변환경 변화와 연령, 성별, 인종 등의 사용자에 대한 다양한 환경에서의 연구가 필요하고, 나아가 제스처 인식을 위한 얼굴 요소(눈, 코, 입, 귀)등의 특징점 추출에 관한 연구가 뒤따라야한다

#### 참고 문헌

- [1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey", in Proc. of The IEEE, Vol 83, No 5, pp.704-740, 1995
- [2] W K Chow and J K Aggarwal, "Computer Analysis of Planar Curvilinear Moving Images", IEEE Transaction Computer, C-26, pp.179-185, 1977
- [3] T. C. Chang, T. S. Huang, and C. Novak, "Facial Feature Extraction from Color Images", Proc of 12th International Conference on Pattern Recognition, Vol 2, pp.39-43, 1994
- [4] Y. Dai and Y. Nakano, "Face-Texture Model Based on SGLD and Its Application in Face Detection a Color Scene", Pattern Recognition, Vol 29, No. 6, pp.1007-1017, 1996.
- [5] R. Crane, A Simplified Approach to Image Processing, Prentice-Hall, 1997.