

# 동영상에서 최적의 얼굴색 정보와 움직임 정보에 기반한 얼굴 영역 추출

(The Extraction of Face Regions based on Optimal Facial  
Color and Motion Information in Image Sequences)

박형철<sup>†</sup>    전병환<sup>††</sup>

(Hyung Chul Park) (Byung Hwan Jun)

**요약** 자연스러운 사용자 작업 환경인 헤드 제스처 인터페이스를 구현하기 위해서는 얼굴 영역 추출에 관한 연구가 선행되어야 하는데, 최근에는 동영상에서 얼굴 영역을 추출하기 위해 색상 정보를 이용하는 연구가 많이 진행되고 있다. 본 논문에서는 대표적으로 사용되는 HSI 컬러 모델과 YIQ 컬러 모델의 각 색상 성분인 H(hue)와 I(in-phase)를 이용하여 얼굴 영역의 추출 성능을 비교한다. 먼저 각 색상 성분에 대해 임계 구간 변화에 따른 최적의 얼굴색 구간을 설정하여 얼굴 영역 추출의 정확도를 비교한다. 다음으로 설정된 최적의 얼굴색 정보와 움직임 정보를 결합한 영상에 대해 얼굴 박스를 추출함으로써 최종적인 얼굴 영역 추출의 정확도를 비교 평가한다.

실험 결과, 최적의 얼굴색 정보를 표현하는 구간은 HSI의 H 성분의 경우  $0^{\circ} \sim 14^{\circ}$ , YIQ의 I 성분은  $-22 \sim -2$ 인 것으로 나타났다. 이 최적 구간에 의한 각 색상 성분의 얼굴 영역 추출의 정확도에서는 I 성분이 H 성분에 비해 약 10% 정도 높은 추출율을 나타냈으며, 최적의 얼굴색 정보와 움직임 정보를 결합한 영상의 경우에서도 I 성분이 약 3% 정도의 보다 나은 얼굴 영역 추출 결과를 얻었다.

**Abstract** The extraction of face regions is required for Head Gesture Interface which is a natural user interface. Recently, many researchers are interested in using color information to detect face regions in image sequences. Two most widely used color models, HSI color model and YIQ color model, were selected for this study. Actually H-component of HSI and I-component of YIQ are used in this research. Given the difference in the color component, this study was aimed to compare the performance of face region detection between the two models. First, we search the optimum range of facial color for each color component, examining the detection accuracy of facial color regions for variant threshold range about facial color. And then, we compare the accuracy of the face box for both color models by using optimal facial color and motion information.

As a result, a range of  $0^{\circ} \sim 14^{\circ}$  in the H-component and a range of  $-22 \sim -2$  in the I-component appeared to be the most optimum range for extracting face regions. When the optimal facial color range is used, I-component is better than H-component by about 10% in accuracy to extract face regions. While optimal facial color and motion information are both used, I-component is also better by about 3% in accuracy to extract face regions.

## 1. 서론

컴퓨터 기술의 눈부신 발전으로 기존의 사용자 인터페이스에 비해 보다 자연스럽게 편리한 사용자 인터페이스에 대한 요구가 날로 증가하고 있다. 이에 따라 사람의 필체 인식이나 음성 인식, 그리고 헤드 제스처 인식 등의 분야에서 이러한 사용자의 요구를 충족시키기 위해 많은 연구가 진행 중이다. 특히 헤드 제스처 인터페이스는 컴퓨터 사용자의 흥미 유발과 자연스러운 작업 환경을 제공해 주는 것이 가능하므로, 여러 응용 분

· 본 논문은 한국과학재단 지정 지역협력 연구센터인 공주대학교 자원 재활용 신소재 연구센터의 연구비 지원에 의해 연구되었음

† 학생회원 : 공주대학교 전자계산학과  
thinkmac@sky.kongju.ac.kr

†† 정회원 : 공주대학교 전자계산학과 교수  
bhjun@knu.kongju.ac.kr

논문접수 : 1999년 4월 12일

심사완료 : 1999년 12월 9일

야에 편리하게 적용될 수 있다. 헤드 제스처는 사람의 얼굴 방향이나 고개의 움직임 또는 눈의 깜빡임 등의 정보를 이용하여 컴퓨터와 사용자간의 의사 소통을 원활히 해주는 인터페이스로, 이를 구현하기 위해서는 우선적으로 사람의 얼굴 영역을 정확히 추출해내는 연구가 선행되어야만 한다[1].

에지 추출(edge detection), 형판 정합(template matching), 그리고 타원 적합(ellipse fitting) 등의 얼굴 영역 추출 방법은 주로 초기에 배경이 단순한 정지 영상에서 사용되어온 대표적인 방법들이다. 에지 추출 방법 등[2-4]은 얼굴 영역을 추출하기 위해 많이 사용된 방법으로, 명암도에 따른 윤곽선을 추출하여 사용한다. 형판 정합 방법 등[2-6]은 주로 가변 형판을 사용하여 형판과 일치하는 영역을 얼굴 영역으로 추출한다. 타원 적합 등[4,7]의 방법에서는 일반적으로 얼굴이 타원형에 가깝다는 사실에 기초하여 Hough 변환에 타원 방정식을 적용하여 얼굴 영역을 추출하였다. 그러나 이와 같은 얼굴 영역 추출 방법은 잡음에 민감하고 처리 시간이 오래 걸리며 복잡한 배경에서의 얼굴 영역 추출이 어렵다는 단점을 가지고 있다.

앞에서 언급한 방법들은 얼굴 영역을 추출하기 위해 사용한 영상이 대부분 그레이 영상이다. 이와 달리 비교적 복잡한 배경을 가진 컬러 정지 영상에서 사람이 가진 얼굴색 고유의 색상 정보를 이용하는 방법도 사용되고 있다. 이 방법은 영상 처리 분야에서 사용되는 몇 가지 컬러 모델의 색상 정보를 이용하는데, RGB, HSI, YIQ 등의 컬러 모델이 주로 이용되어 왔다. Chang 등[8]은 RGB 컬러 모델을 이용하여 얼굴색 정보를 포함하는 색상 정보의 범위를 설정하고 이 범위에 속한 화소값을 바탕으로 얼굴 영역을 추출하였다. 유태웅 등[9]은 RGB 컬러 모델에서 HSI 컬러 모델로 변환된 정지 영상에서 색채 성분(H와 S)만을 색채 히스토그램에 적용하여 각 화소를 얼굴색일 확률값 영상으로 변환하고, 이 영상에 얼굴의 대칭성 정보를 이용하기 위해 타원 마스크를 적용하여 얼굴 영역을 추출하였다. Herodotou 등[10]은 HSV 컬러 모델에서 H, S, V가 나타내는 피부색 구간의 히스토그램을 측정하고, 여기에 Median Filtering 및 Morphological 연산자를 이용한 후처리 과정을 거쳐서 분할된 영역을 재정의 한 후, 퍼지 집합을 사용하여 얼굴 영역을 추출하였다. Dai 등[11]은 색상 정보를 이용한 얼굴 영역 추출을 위해 RGB 컬러 모델에서 YIQ 컬러 모델로 변환하여 색상 정보를 나타내는 I 성분의 얼굴색 범위를 설정하고, SGLD matrix의 특징 파라미터를 이용하였다. Terrillon 등[12]은 RGB 컬러

모델을 이용해 자신들이 변형한 STV 컬러 모델에서 임계값을 설정하고 S와 T에 기반한 히스토그램을 작성하여 얼굴색 범위를 설정했으며, 이를 신경망에 적용하여 얼굴 영역을 추출하였다. 한편 동영상에서 얼굴 영역을 추출하기 위해 Stringa[13]는 시간적으로 연속된 두 영상간의 명암도 차이를 이용하는 방법을 사용하였다. 이러한 방법들은 초기의 얼굴 영역 추출 방법에 비해 처리시간이 짧고 복잡한 배경에서도 적용이 가능하다는 장점으로 인해 많이 사용되고 있다. 최근에는 앞에서 언급한 색상 정보와 움직임 정보를 결합하여 동영상에서 얼굴 영역을 추출하는 방법[14]이 소개되고 있는데, 이 방법의 가장 큰 장점은 배경에서의 얼굴색 정보를 포함하는 물체나 움직임이 있는 물체를 제거할 수 있다는 점이다.

색상 정보를 이용한 기존의 얼굴 영역 추출 방법들은 얼굴색이 나타나는 구간을 설정하여 얼굴 영역을 추출하기도 하였으나, 각 컬러 모델의 색상 성분에 대한 최적의 얼굴색 구간을 설정하는 연구가 미흡하고, 사용된 각 컬러 모델간의 얼굴색 표현 능력에 대한 성능 비교가 이루어지지 않았다. 이러한 연구의 필요성은 최적의 색상 구간 설정을 통해 각 컬러 모델의 특성을 파악하여 주위 환경에 영향을 덜 받는 컬러 모델을 얼굴 영역 추출에 이용하는 것이 중요하기 때문이다. 따라서 본 논문에서는 색상 정보를 이용한 얼굴 영역 추출시 많이 사용되는 HSI 컬러 모델과 YIQ 컬러 모델의 각 색상 성분 H(hue)와 I(in-phase)를 이용하여 각 색상 성분의 임계치 변화에 따른 얼굴색 정보를 포함하는 구간을 탐색하고, 이 구간 중 얼굴색을 최적으로 표현하는 구간을 설정하여 얼굴 영역 추출의 정확도를 비교한다. 그런 다음 각각의 얼굴색 정보에 움직임 정보를 결합한 영상에 대한 얼굴 박스를 추출하여 얼굴 영역 추출의 정확도를 비교 평가하여 색상 정보를 이용한 얼굴 영역 추출시 더 좋은 성능을 나타내는 컬러 모델을 알아보기 위한 연구를 수행한다.

그림 1은 본 논문에서의 얼굴 영역 추출 과정이다. 먼저 시간적으로 연속된 두 원영상 (a)와 (b)를 이용하여 인접 영상간의 움직임 정보에 의해 차영상 (c)를 얻어낸다. 또한 원영상 (b)로부터 H 성분 및 I 성분 각각의 얼굴색 탐색을 통해 설정된 최적의 얼굴색 구간을 적용하여 얼굴색 영상 (d)를 추출한다. 다음으로 차영상 (c)와 얼굴색 영상 (d)를 결합(AND)하여 얼굴 영역 추출을 위한 결합 영상 (e)를 얻는다. 마지막으로 영상 (e)를 투영해서 얻어진 히스토그램을 분석하여 얼굴 영역을 나타내는 영상 (f)를 얻는다.

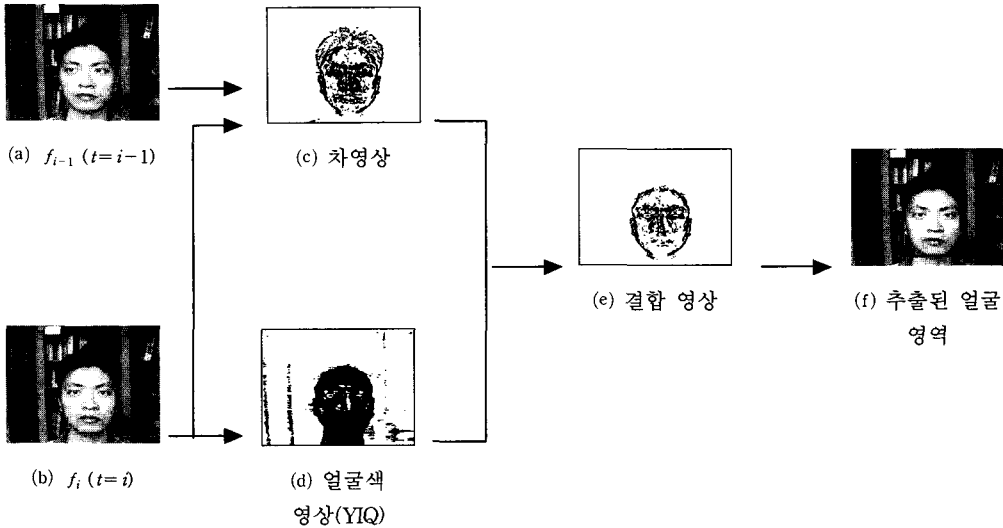


그림 1 얼굴 영역 추출 과정

2. 얼굴색 정보를 이용한 얼굴 영역 추출

2.1 컬러 모델의 변환

컴퓨터 화면상에서 처리되는 RGB 컬러 모델은 색상 정보와 명도 정보의 분리가 복잡하므로, 본 논문에서는 이 두 가지 정보를 쉽게 분리 할 수 있는 HSI 컬러 모델과 YIQ 컬러 모델로 각각 변환하여 사용한다.

색상(hue), 채도(saturation), 명도(intensity)의 세 가지 성분으로 구성된 HSI 컬러 모델은 영상 처리 분야에서 널리 사용되는 컬러 모델이다[15,16]. 이를 RGB 공간에서 변환하기 위한 식은 다음과 같다.

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

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

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

YIQ 컬러 모델은 Y(luminosity: 흑백 디스플레이 장치 상에 두 개의 다른 색들을 구별), I(in-phase: 오렌지색(orange)-청록색(cyan) 색상 정보를 포함), Q(quadrature: 녹색(green)-자홍색(magenta) 색상을 포함)로 구성되며 TV 전송에 사용되는 컬러 모델이다[15]. 마찬가지로 다음 식에 의해 RGB 공간에서

YIQ 공간으로 변환한다.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

앞에서 언급한 두 가지 컬러 모델의 성분 중, 본 논문에서는 색상 정보를 포함하는 HSI의 H(hue) 성분과 YIQ의 I(in-phase) 성분을 이용하여 각각의 얼굴 영역 추출의 정확도를 비교한다. 이 때 H 성분은 구간 0° ~ 360° 의 값으로 표현되고, I 성분은 구간 -152 ~ +152 의 값으로 표현되므로, 본 연구에서는 이들 구간을 0 ~ 255로 정규화하여 실험하도록 한다.

2.2 얼굴색 구간 탐색 및 최적 구간 설정

H 성분과 I 성분의 색상 정보를 이용하여 각각의 얼굴 영역을 추출하기 위해서는 각 성분의 얼굴색이 존재하는 구간을 탐색한 후, 최적의 얼굴색 구간을 설정해야 한다. 이를 위해 본 논문에서는 식 (1)의 측정치  $\alpha$ ,  $\beta$ ,  $\gamma$  를 사용한다.

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

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

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

$F_s$  : 찾은 얼굴색 영역  
 $F_o$  : 실제 얼굴 영역

$\alpha$ 는 원영상의 실제 얼굴 부분에서 얼굴색 정보를 이용해 얼굴색이라고 찾은 영역의 비율이고,  $\beta$ 는 얼굴 부분을 제외한 영역에서 실제로도 얼굴색이 아니라고 찾은 영역의 비율이다.  $\gamma$ 는 최적의 얼굴색 정보를 포함하는 구간을 설정하기 위한 기준으로  $\alpha$ 와  $\beta$ 의 평균값이다. 측정치  $\alpha, \beta, \gamma$ 는 모두 0 ~ 1 사이의 값을 갖는다. 여기서, 절대값  $|F|$ 는 영역  $F$ 의 면적이다.

식 (1)의 측정치를 이용하여 얼굴색 구간을 탐색하는 과정은 다음과 같다. 먼저 원영상에서 실제 얼굴색에 해당하는 부분(얼굴과 목)과 나머지 부분을 구별하여 얼굴색 부분을 검게 칠하는 이진화 작업을 수행하여 이진 영상( $F_o, F_o^c$ )을 생성한다. 그런 다음  $\alpha$ 에 대한 임계치( $\theta_H$  또는  $\theta_l$ )를 0 ~ 255까지 10씩 증가시키면서 원영상을 이진화하고 앞에서 생성된 이진 영상과의 결합에 의해  $\alpha$ 를 측정한다. 측정된  $\alpha$ 는 얼굴색 정보를 포함하는 부분에서 급격히 변하기 시작하므로,  $\alpha$ 가 급변하는 구간을 얼굴색 정보를 탐색하기 위한 구간으로 설정한다. 급변하는  $\alpha$ 의 구간에서 각각의 상한 경계치( $\theta_{Top}$ )와 하한 경계치( $\theta_{Down}$ )의 구간 폭을 변화시키면서  $\alpha, \beta, \gamma$ 를 측정하여  $\gamma$ 가 최대가 되는 구간을 얼굴 영역 추출을 위한 최적의 얼굴색 정보를 포함하는 구간으로 각각 설정한다.

### 3. 최적의 얼굴색 정보와 움직임 정보에 의한 얼굴 영역 추출

추출된 얼굴색 정보와 움직임 정보를 결합한 영상에 대하여 얼굴 영역 추출의 정확도를 비교하기 위해 얼굴 박스 영상을 이용한다. 이 때 추출되는 얼굴 박스의 정확도를 측정하기 위해 얼굴색 정보 이용시 사용한  $\alpha, \beta, \gamma$ 와 유사한  $\alpha', \beta', \gamma'$ 를 사용한다.

$$\begin{aligned} \alpha' &= \frac{|B_s \cap B_o|}{|B_o|} \\ \beta' &= \left(1 - \frac{|B_s \cap B_o^c|}{|B_o^c|}\right) \\ \gamma' &= \frac{\alpha' + \beta'}{2} \end{aligned} \tag{2}$$

$B_s$  : 찾은 얼굴 박스  
 $B_o$  : 실제 얼굴 박스

$\alpha'$ 는 원영상의 얼굴 박스내에서 얼굴 박스라고 찾은 영역의 비율이고,  $\beta'$ 는 얼굴 박스 이외의 영역에서 얼굴 박스가 아니라고 찾은 영역의 비율이다.  $\gamma'$ 는  $\alpha'$ 와  $\beta'$ 의 평균값으로서 추출된 얼굴 박스의 정확도를 판단하는 기준이다. 측정치  $\alpha', \beta', \gamma'$ 도 0 ~ 1 사이의 값을 가진다. 여기서, 절대값  $|B|$ 는 얼굴 박스  $B$ 의 면적

이다.

얼굴 박스를 추출하기 위해서는 히스토그램의 각 경계치( $\varphi_{HSI_H}, \varphi_{HSI_V}, \varphi_{YIQ_H}, \varphi_{YIQ_V}$ )를 설정해야 하는데, 본 논문에서는 각각의 수평 경계치( $\varphi_H$ )와 수직 경계치( $\varphi_V$ )의 조합을 통해 설정된 식 (2)의  $\gamma'$ 가 최대가 되는 두 경계치를 얼굴 박스 추출을 위한 최종 경계치로 설정한다. 설정된 최종 경계치를 바탕으로 다음과 같은 과정을 통해 얼굴 영역을 나타내는 얼굴 박스를 추출한다. 원영상의 실제 얼굴과 나머지 부분을 구분하여 얼굴을 검게 칠하는 이진화 작업을 수행한 후, 수평 및 수직 투영(projection)을 통해 실제 얼굴 박스( $B_o$ )를 추출한다. 다음으로 2절에서 추출된 최적의 얼굴색 정보와 움직임 정보를 결합하여 결합(AND)영상을 생성하고 이를 수평 및 수직 방향으로 투영시켜 얼굴 영역 분포를 나타내는 히스토그램을 생성한다. 이를 바탕으로 실험에 의해 측정된 경계치( $\varphi_{HSI}$  또는  $\varphi_{YIQ}$ )를 얼굴 박스 추출을 위해서 적용하고, 최종적으로 추출된 얼굴 박스에 대한 얼굴 영역 추출의 정확도를 평가하기 위해 실제 얼굴 박스와 비교한다.

투영에 의해 생성된 히스토그램을 분석해 보면, 배경의 일부 작은 영역에서 빛의 산란 효과에 의해 얼굴색 정보와 움직임 정보를 동시에 포함하는 영역이 발생한다. 이로 인해 이 영역이 얼굴 영역으로 오인식 되는 경우가 발생하므로 경계치( $\varphi_{HSI}$  또는  $\varphi_{YIQ}$ )에 만족하는 히스토그램의 구간 폭이 극히 좁으면 노이즈로 판단하는 휴리스틱 정보를 이용하여 이러한 문제점을 해결한다.

## 4. 실험

### 4.1 실험 환경

실험에 사용되는 영상 데이터는 남/여 각 4명씩, 총 8명의 데이터를 학습 및 시험 데이터로 분류하였다. 영상 데이터의 크기는 320×240 화소로, 뒷배경이 복잡하고 배경에서의 움직임은 없으며 사용자가 컴퓨터 화면을 주시하는 얼굴 영상을 대상으로 한다. 이 때 자연광의 유·무나 방향에 따라 표 1과 같은 4가지 장면으로 구분하여 영상 데이터를 취득하였다.

표 1 영상 데이터의 분류

장면	정의
A 영상	실내 조명만 있는 경우
B 영상	정면에서 자연광이 들어오는 경우
C 영상	측면에서 자연광이 들어오는 경우
D 영상	후면에서 자연광이 들어오는 경우(역광)

실험을 위해 Pentium MMX 200MHz PC 상에서 CCD 카메라와 Telegeny 캡처보드를 사용하여 영상 데이터를 취득하였고, Visual C++ 5.0으로 프로그래밍 작업을 수행하였다.

4.2 실험 결과 및 분석

실험은 색상 정보를 표현하는 H(hue) 성분과 I(in-phase) 성분에 대한 각각의 얼굴색 구간 탐색 및 최적의 얼굴색 구간 설정을 통한 얼굴 영역 추출의 정확도 비교와 최적의 얼굴색 정보에 움직임 정보를 결합한 경우에 대해서 얼굴 영역 추출의 정확도를 비교해 본다.

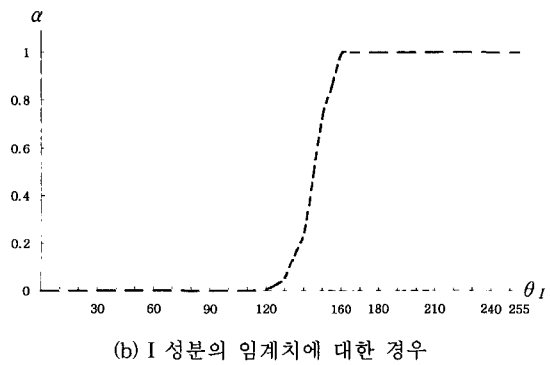
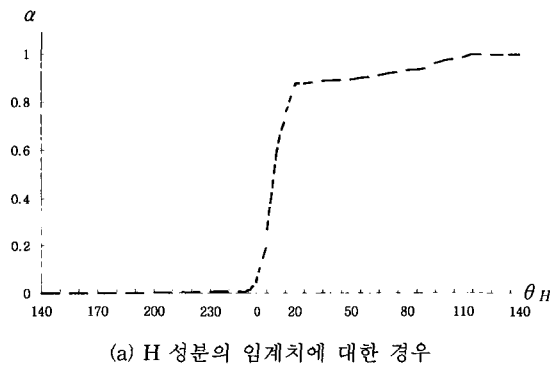


그림 2 학습 데이터에 대한 경계치 변화에 따른  $\alpha$ 의 변화도

4.2.1 색상 성분별 최적의 얼굴색 구간 설정

그림 2는 학습 데이터에 대한 식 (1)의  $\alpha$ 를 적용하여 얻어진 임계치  $\theta_H$ 와  $\theta_I$ 의 얼굴색 정보 탐색 변화도이다. 2.1절에서 언급했듯이 H 성분의 구간은  $0^\circ \sim 360^\circ$ , I 성분은  $-152 \sim +152$ 인데, 이들 구간을 본 논문에서는  $0 \sim 255$ 로 정규화하였다. 그림 2-(a)를 보면 구간이  $0 \sim 255$ 의 환형으로 되어있는데, 이는 H(hue) 성분

이  $0^\circ \sim 360^\circ$ 의 환형 구간을 가지기 때문이다. 실험에 의해  $\theta_H$ 는 구간  $0 \sim 20$ ,  $\theta_I$ 는 구간  $120 \sim 160$ 에서  $\alpha$ 값이 급격히 변하므로 이들 구간이 대부분의 얼굴색 정보를 포함하는 것으로 판단하여 최적의 얼굴색 구간을 설정하기 위한 탐색 구간으로 설정한다. 이 탐색 구간을 바탕으로 얼굴 영역 추출의 정확도( $\gamma$ )가 가장 높게 나타나는 최적 구간을 설정해야 하는데,  $\theta_H$ 와  $\theta_I$ 에 대해 각각의 상한 경계치( $\theta_{Top}$ )와 하한 경계치( $\theta_{Down}$ )에 대한 조합을 통해  $\gamma$ 값을 조사하였다. 학습 데이터에 대한 최적의 얼굴색 구간을 탐색한 결과, H 성분의 경우는 구간  $0 \sim 20$ , I 성분의 경우에는 구간  $130 \sim 150$ 에서 최적의 얼굴색을 표현하는 구간으로 판명되었다.

설정된 구간을 바탕으로 영상 데이터에 대한 H 성분과 I 성분의 각각의 얼굴 영역 추출의 정확도를 비교해 보면 표 2와 같다.

표 2 최적의 얼굴색 구간에 따른 얼굴 영역 추출율( $\gamma$ )  
(a) H 성분의 최적 구간 :  $0 \sim 20$

영상	A	B	C	D	평균
$\alpha$	0.872	0.795	0.540	0.511	0.679
$\beta$	0.829	0.815	0.796	0.823	0.816
$\gamma$	0.850	0.805	0.668	0.667	<b>0.748</b>

(b) I 성분의 최적 구간 :  $130 \sim 150$

영상	A	B	C	D	평균
$\alpha$	0.867	0.786	0.894	0.765	0.828
$\beta$	0.892	0.891	0.844	0.796	0.856
$\gamma$	0.880	0.839	0.869	0.781	<b>0.842</b>

실험 결과에 의하면, YIQ 컬러 모델의 I(in-phase) 성분을 이용한 방법이 HSI 컬러 모델의 H(hue) 성분에 비해 얼굴 영역 추출율( $\gamma$ )의 정확도에서 약 0.1(10%) 정도 높게 나타났다.

4.2.2 색상 성분에 따른 얼굴 영역 추출의 정확도 비교

최적의 얼굴색 정보와 움직임 정보를 결합하여 생성된 영상을 이용하여 얼굴 박스를 추출하는데 사용하는 두 경계치( $\theta_H$ 와  $\theta_I$ )는 식 (2)의  $\gamma'$ 에 의해 표 3과 같이 측정되었다.

H 성분의 얼굴색 정보와 움직임 정보를 결합한 경우

표 3 얼굴 박스 추출을 위한 각각의 경계치  
(a)  $\varphi_{HSI}$ 의 각 경계치에 대한  $\gamma'$ 값

$\varphi_V \backslash \varphi_H$	8	9	10
9	0.9051	0.9068	0.9060
10	0.9059	<b>0.9073</b>	0.9066
11	0.9003	0.9016	0.9009

(b)  $\varphi_{YIQ}$ 의 각 경계치에 대한  $\gamma'$ 값

$\varphi_V \backslash \varphi_H$	9	10	11
9	0.9505	0.9507	0.9506
10	0.9518	<b>0.9519</b>	0.9516
11	0.9515	0.9515	0.9512

의 경계치는  $\varphi_{HSI\_H} = 9$ ,  $\varphi_{HSI\_V} = 10$ 이고, I 성분의 얼굴색 정보와 움직임 정보를 결합한 경우는  $\varphi_{YIQ\_H} = 10$ ,  $\varphi_{YIQ\_V} = 10$ 으로 이를 각각의 최종 경계치로 설정하였다. 표 3에서 설정된 경계치를 영상 데이터에 적용하여 추출된 얼굴 영역의 정확도는 표 4와 같다. 실험 결과, I 성분의 얼굴색 정보에 움직임 정보를 결합시켜 추출된 영상이 H 성분의 얼굴색 정보에 움직임 정보를 결합시킨 영상에 비해 약 0.03(3%) 정도 얼굴 영역을 더 정확히 추출하는 것으로 나타났다.

표 4 얼굴 영역 추출의 정확도( $\gamma'$ )  
(a) H 성분 영상 & 차영상의 경우

영상	A	B	C	D	평균
$\alpha'$	0.972	0.992	0.962	0.734	0.915
$\beta'$	0.936	0.917	0.833	0.960	0.912
$\gamma'$	0.954	0.955	0.898	0.847	<b>0.914</b>

(b) I 성분 영상 & 차영상의 경우

영상	A	B	C	D	평균
$\alpha'$	0.944	0.995	0.957	0.866	0.941
$\beta'$	0.979	0.958	0.914	0.936	0.947
$\gamma'$	0.962	0.977	0.936	0.901	<b>0.944</b>

그림 3은 각 영상 데이터에서 추출된 얼굴 영역을 나타낸 것이다. H 성분을 적용한 영상의 경우 머리카락 상단 경계 부분까지를 얼굴 영역의 경계 부분으로 오인식하는 문제가 많이 발생하였다. 또한 사용자의 흰색 상의에 살색이 약하게 비치고 상체의 미세한 움직임이 있는 경우, H 성분을 적용한 영상에서 이를 얼굴 영역으로 잘못 추출하는 문제점도 발생하였다. 역광 영상 D의 경우에는 대체적으로 얼굴색 정보의 추출율이 다소 떨어지는 것으로 나타났다.

### 5. 결론 및 향후 연구 과제

헤드 제스처 인터페이스의 구현을 위한 선결 과제인 얼굴 영역을 추출하기 위해 본 논문에서는, 색상 정보를 표현하는 HSI 컬러 모델의 H(hue) 성분과 YIQ 컬러 모델의 I(in-phase) 성분의 임계치 변화에 따른 각각의 얼굴색 구간을 탐색하고, 탐색된 구간내에서 임계치 구간의 변화에 따른 최적의 얼굴색 구간을 설정하여 두 성분의 얼굴 영역 추출 정확도를 비교하였다. 또한 추출된 각각의 최적의 얼굴색 정보와 움직임 정보를 결합한 영상에 대해서도 얼굴 영역 추출의 정확도를 비교 평가하였다. H 성분과 I 성분에 대한 각각의 얼굴색 정보를 포함한 구간을 탐색하여 최적 구간을 설정한 결과, H 성분의 경우는  $0^\circ \sim 14^\circ$  ( $0 \sim 20$ ), I 성분의 경우는  $-22 \sim -2$  ( $130 \sim 150$ )에서 얼굴색을 표현하는 최적의 구간으로 판명되었다. 이를 얼굴 영역 추출에 적용한 결과, I 성분의 얼굴색 정보를 이용한 방법이 H 성분의 얼굴색 정보를 이용한 방법보다 약 10% 정도 얼굴 영역을 정확히 추출하였다. 또한 최적의 얼굴색 정보와 움직임 정보를 결합한 영상에 대해서도 I 성분을 적용한 방법이 H 성분을 적용한 방법에 비해 약 3% 정도 얼굴 영역을 더 정확히 추출하였다. 한편 얼굴색 정보를 처리하는데 걸리는 시간은 H 성분의 경우 약 0.56초, I 성분을 이용한 경우는 약 0.16초로, I 성분 이용시 처리 시간이 약 3.5배정도 단축됨을 알 수 있다. 따라서 본 논문에서의 실험에 의해 색상 정보를 이용하여 얼굴 영역을 추출할 경우 YIQ 컬러 모델의 I(in-phase) 성분이 HSI 컬러 모델의 H(hue) 성분을 이용하는 것보다 비교적 더 정확한 추출율을 나타내는 것으로 판명되었다.

반면 문제점으로는 조명에 따른 영향 및 간혹 빛의 산란 효과에 의해 얼굴 영역 추출율이 낮아지는 경우가 발생한다는 점이다. 또한 얼굴색을 가진 움직이는 물체가 얼굴 영역과 연결되는 경우에 대한 문제점도 내포하고 있다. 따라서 향후 연구 과제로는 위 문제점에 적절



그림 3 얼굴 영역이 추출된 영상 데이터의 예

히 대응할 수 있는 적응형 알고리즘 개발과 인종, 연령, 성별에 따른 다양한 사용자에 대한 얼굴색 정보의 분포에 대해 좀더 세밀한 연구가 필요하다. 나아가 헤드 제스처 인터페이스에 적용하기 위한 눈, 코, 입 등의 얼굴 특징점을 추출하는 방법에 대해서도 연구가 뒤따라야 한다.

### 참 고 문 헌

- [1] C. Morimoto, Y. Yacoob and L. Davis, "Recognition of Head Gestures using Hidden Markov Models," Proc. of Int'l Conf. on Pattern Recognition, pp.461-465, 1996.
- [2] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proc. of the IEEE, Vol. 83, No. 5, pp.704-740, 1995.
- [3] L. C. De Silva, K. Aizawa and M. Hatori, "Detection and Tracking of Facial Feature by Using Edge Pixel Counting and Deformable Circular Template Matching," Trans. on IEICE Information and System, Vol. E78-D, No. 9, pp.1195-1207, 1995.
- [4] C. L. Huang and C. W. Chen, "Human Facial Feature Detection for Face Interpretation and Recognition," Pattern Recognition, Vol. 25, No. 12, pp.1435-1444, 1992.
- [5] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, No. 10, pp.1042-1052, 1993.
- [6] Y. H. Kwon and N. V. Lobo, "Face Detection Using Templates," Proc. of Int'l Conf. on Pattern Recognition, Vol. 1, pp.764-767, 1994.
- [7] S. Sirohey, "Human Face Segmentation and Identification," Tech. Report, CAR-TR\_695, Center for Automation Research, Univ. of Maryland, 1993.
- [8] T. C. Chang, T. S. Huang and C. Novak, "Facial Feature Extraction from Color Images," Proc. of Int'l Conf. on Pattern Recognition, Vol. 2, pp.39-43, 1994.
- [9] 유태웅, 오일석, "색채 분포 정보에 기반한 얼굴 영역 추출," 정보과학회논문지(B), 제24권, 제2호, pp.180-192, 1997.
- [10] N. Herodotou, K. N. Plataniotis and A. N. Venetsanopoulos, "Automatic Location and Tracking of the Facial Region in Color Video Sequences," Signal Processing: Image Communication, Vol. 14, No. 5, pp.359-388, 1999.
- [11] 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.
- [12] J. C. Terrillon, M. David and S. Akamatsu, "Automatic Detection of Human Faces in Natural Scene Images by Use of a Skin Color Model and of Invariant Moments," Proc. of Int'l Workshop on Automatic Face- and Gesture-Recognition, pp.112-117, 1998.
- [13] L. Stringa, "Eyes Detection for Face Recognition," Applied Artificial Intelligence, Vol. 7, No. 4, pp.365-382, 1993.
- [14] C. H. Lee, J. S. Kim and K. H. Park, "Automatic Human Face Location in a Complex Background using Motion and Color Information," Pattern Recognition, Vol. 29, No. 11, pp.1877-1889, 1996.
- [15] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison Wesley, pp.228-235, 1993.
- [16] R. Crane, A Simplified Approach to Image Processing, Prentice-Hall, pp.17-18, 1997.



박형철

1997년 공주대학교 전자계산학과 졸업(이학사). 1998년 ~ 현재 공주대학교 대학원 전자계산학과 석사과정. 관심분야는 패턴인식, 컴퓨터비전, HCI, 3D 그래픽스.



전병환

1989년 연세대학교 전자공학과 졸업(공학사). 1991년 연세대학교 대학원 전자공학과 졸업(공학석사). 1996년 연세대학교 대학원 전자공학과 졸업(공학박사). 1991년 ~ 1992년 SK컴퓨터통신 주식회사 연구원. 1996년 ~ 1997년 연세대학교 신호처리연구센터 전문연구원. 1999년 ~ 현재 한국정보학회 충청지부 이사. 1997년 ~ 1999년 공주대학교 전자계산학과 전임강사. 1999년 ~ 현재 공주대학교 전자계산학과 조교수. 관심분야는 패턴인식, 컴퓨터비전, 가상현실, 인공지능.