

# 모바일 아바타 생성을 위한 얼굴 추출에 관한 연구

최재영, 황승호, 황보태근, 양영규  
경원대학교 소프트웨어 대학

## A Study on Face Extraction Generation of Mobile Avatar

Jae-Young Choi, Seung-Ho Hwang, Taeg-Keun Whangbo, Young-Kyu Yang  
College of Software, Kyungwon University

### 요 약

오늘날 모바일의 발전과 함께 지속적으로 모바일 콘텐츠들에 대한 연구가 급속히 증가되어 지고 있다. 특히 모바일상에서 아바타나 캐리커처 시스템에 관한 연구가 활발히 진행되고 있으며, 이에 필요한 기반기술인 얼굴 영상의 추출에 대한 연구가 필수적으로 병행 되어지고 있다. 본 논문에서는 이러한 흐름에 맞추어 핸드폰용 카메라에서 사용 가능한 얼굴추출 알고리즘을 제안한다. 얼굴의 추출을 위해서 ACM을 응용하였으며, 이에 대한 초기의 제어점 선택을 위하여 T-S 색상 공간을 분석한 후에 마할라노비스 거리측정을 하였다. 또한 ACM의 외부에너지를 일반적으로 많이 사용하는 기존의 에지를 이용한 방법이 핸드폰용 카메라의 특성상 조명의 영향과 화질의 열화로 의해 정확하게 추출되지 못하는 단점을 보완하고자 얼굴 영역 면적의 변화량 측정 방법을 사용함으로써 조명의 영향에 강인하도록 하였다.

### 1. 서론

인터넷은 90년대 중반부터 급속히 성장하여 현재 실생활에 없어서는 안되는 중요한 역할을 하고 있다. 최근 이러한 인터넷 기술성장, 지능형 핸드폰의 기술 성장과 칼라핸드폰의 발전을 바탕으로 인터넷 콘텐츠를 무선 인터넷에 적용하는 연구가 급속히 발전하고 있다[1]. 특히, 미래형 고부가가치 문화산업으로 부각되고 있는 애니메이션 산업에 대한 연구가 증가되면서 기반 기술인 아바타 생성, 얼굴 영상처리기술에 대한 연구가 활발히 진행 되어지고 있다.

얼굴 영상기술은 크게 얼굴을 인식하는 분야, 얼굴의 특징점을 추출하는 분야, 캐리커처 또는 개인 아바타의 생성 등을 위한 얼굴 영역을 추출하는 분야로 나누어 볼 수 있고, 아바타를 위한 얼굴 영역 추출 분야는 기존의 얼굴 인식이나 얼굴 특징 추출과는 달리 정확한 얼굴 형태를 추출할 수 있어야 한다[2]. 얼굴

추출에 관한 연구는 입력영상의 열화 및 조명의 영향에 민감한 특징이 있으며 특히, 해상도 및 화질이 떨어지는 핸드폰용 카메라에 의해 입력되는 영상에서 더욱 크게 나타나므로 정확한 얼굴 영역의 추출에 어려움이 발생한다[3].

본 논문에서는 핸드폰 카메라를 이용한 얼굴 영역의 정확한 추출을 위해 다른 색상 공간보다 얼굴 영역의 추출에 효율적인 TSL 색상 공간을 이용하여 얼굴 영역의 후보위치를 선정하고, 이를 근간으로 초기의 제어점을 생성하였다. 또한 ACM(Active Contour Model)의 외부에너지를 화소간의 밝기차가 아닌 얼굴 영역의 면적의 변화량을 측정하여 응용하였다. 이에 관한 세부내용으로 2장에서 TSL 색상 공간과 마할라노비스 거리를 이용한 후보 영역의 선정과정을 기술하고, 3장에서 변형된 ACM의 방식으로 보다 정확하게 얼굴을 추출하는 방법을 서술한다. 4장에서는 제안한 방법의 유효성을 평가하고 이를 근간으로 결론을 정리하며, 향후의 발전방향을 제시한다.

본 연구는 2003년도 대학 IT 연구센터 육성지원 사업비의 지원을 받아 연구되었음

## 2. TSL 색상 공간을 이용한 얼굴 영역 탐색

얼굴의 영역을 검출하는 방법은 여러 가지 방법들이 있으나 가장 간단하고 빠른 성능을 보일 수 있는 것은 얼굴 색상 공간을 이용하는 방법이다. 이는 조명의 영향에 의해 미검출되는 영역이 존재하는 단점이 있지만, 얼굴에 근사한 영역의 추출이 가능하므로, 이를 이용하여 ACM의 초기 제어점을 선택하였다[4]. 본 논문에서 얼굴의 정확한 위치 판단은 3장의 응용 ACM에서 하기 때문에 우선적으로 얼굴 영역의 근사 위치 판단을 위하여 TSL(Tint, Saturation, Luminance) 공간과 마할라노비스 거리를 이용하였다.

얼굴 영역의 추출 알고리즘은 그림 1과 같이 구성되어 있다. 입력 영상은 핸드폰 카메라에 의해서 얻어진 영상으로 얼굴이 중앙에 위치하는 영상만으로 제한하였다. 우선 입력영상이 들어오면 핸드폰 연산의 한계성을 고려해 입력 영상들은 다운 샘플링을 하

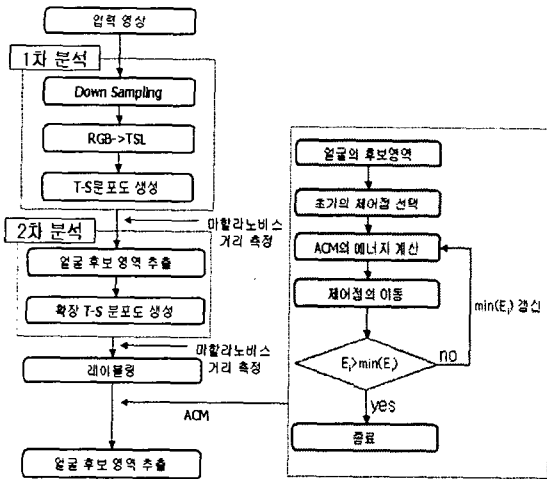


그림 1. 얼굴 영역 추출 알고리즘

여 사용함으로써 연산량을 감소시키고, 화질의 열화와 조명의 영향으로 생기는 국부적 잡음을 제거하여 미검출되는 영역을 최소화 하였다. 다운 샘플링은 세로를 50화소로 고정하고 가로는 입력영상의 가로:세로 비율에 의해 자동으로 설정되게 하였다. 입력된 영상은 핸드폰의 특성상 조명의 영향을 많이 받으며, 영상의 질 또한 일반적인 영상들보다 많이 떨어지는 현상이 나타나게 된다. 이러한 입력 영상은 RGB 색상으로 입력되므로 식 (1)을 이용하여 TSL 색상 공간으로 변환하였다.

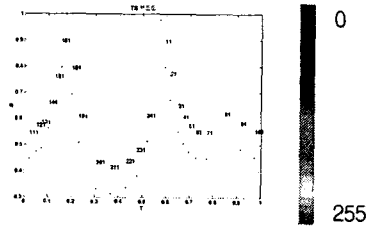
$$T = \begin{cases} \arctan(r'/g')/2\pi + 1/4, & g' > 0 \\ \arctan(r'/g')/2\pi + 3/4, & g' < 0 \\ 0 & g' = 0 \end{cases} \quad (1)$$

$$S = [9/5(r'^2 + g'^2)]^{1/2}$$

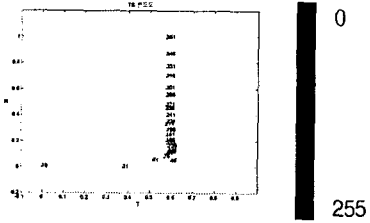
$$L = 0.299R + 0.587G + 0.114B$$

, where  $r' = (r-1/3)$ ,  $g' = (g-1/3)$   
 $r = R/(R+G+B)$ , and  $g = G/(R+G+B)$

변환된 T-S 공간에서의 전체적인 색상의 분포는 그림 2의 (a)와 같고, 단일색상에서의 밝기값 분포는 그림 2의 (b)와 같이 나타난다. 일반적으로 얼굴의 색상은  $0.5 < T < 0.7$  의 범위에서 나타나지만, 조명에 의해 특정한 색상이 강조되는 핸드폰 카메라에서는 얼굴 색상 공간이 달라질 수 있다.



(a) 색상별 T-S 분포



(b) T-S 공간내의 명암의 변화  
 그림 2. T-S 공간 분석

또한, 이러한 얼굴 색상 공간의 특징은 타원형의 구조를 가지고 있으므로 얼굴 중심 영역의 T-S값을 근간으로 마할라노비스 거리에 의한 얼굴 영역을 선정하였다. 마할라노비스 거리측정은 식 (2)와 같고, 이에 의한 얼굴 후보 영역의 추출은 그림 3과 같다.

$$P(X(i,j) | K) = \frac{\exp[-\frac{1}{2}([X(i,j) - m_K]^T C_K^{-1} [X(i,j) - m_K])]}{2\pi |C_K|^{1/2}} \quad (2)$$

$$[\lambda_K(i,j)]^2 = [X(i,j) - m_K]^T C_K^{-1} [X(i,j) - m_K] \quad (3)$$

$X(i,j)$ 를 좌표  $i, j$ 에서의  $[T(i,j) S(i,j)]^T$  벡터로 표현 가능하다고 할 때,  $X(i,j)$ 가 얼굴색 공간  $K$ 에 포함

될 조건 확률함수를 2차원 가우시안 확률 밀도 함수로 나타내면 식 (2)와 같다. 여기서  $C_K$ 를 얼굴색 공간  $K$ 의 공분산이라 하고,  $m_K$ 를  $[m_{\tau_K} \ m_{s_K}]^T$ 로 표현되는 평균 벡터라 가정하면 마할라노비스 거리  $\lambda_K(i, j)$ 는 식 (3)과 같다. 여기서 중심은  $m_K$ 가 되며, 중심축은 공분산  $C_K$ 에 의하여 결정된다.

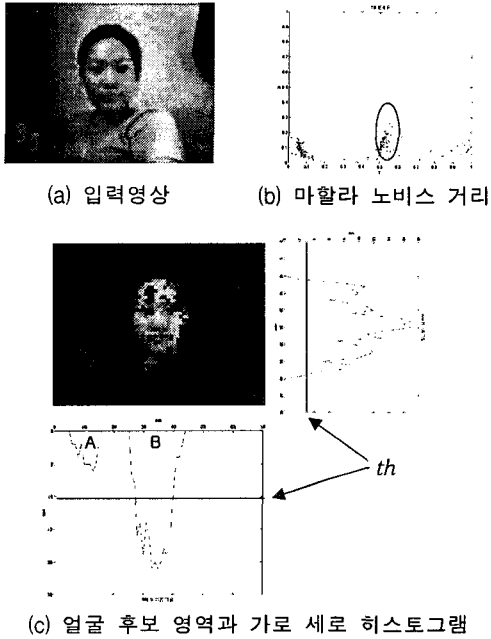


그림 3. 얼굴 후보 영역 추출

본 논문에 쓰인 입력영상들은 각 개인들이 핸드폰 카메라를 이용하여 찍은 영상들을 인터넷에서 수집하였다. 그림 3의 (a)와 같은 영상이 입력되면 식 (3)에 의해 얼굴 영역에 대한 마할라노비스 거리 측정을 하고, 그림 3의 (b)의 타원안에 해당하는 화소를 추출할 수 있으며, 추출된 영역은 그림 3의 (c)와 같이 나타난다. 그림 3에서 누적 히스토그램의 중심부터의 연결성을 판단한 후 A와 B 처럼 분리된 영역이 나타나고, 나타난 영역이 임계값 'th'보다 작은 영역이라면 얼굴 후보 영역에서 제거하여 영역 B만을 선택한다. 하지만, 1차 분석에 의한 영상은 그림 4의 (a), (c)와 같이 얼굴 영역 내의 많은 부분들이 추출되지 못함을 볼 수 있다. 또한 그림 4의 (c)에서의 좌측처럼 잡영이 얼굴 영역으로 판단되는 부분이 나타나는데, 이러한 잡영의 제거를 위한 추가적인 알고리즘이 필요하다.

본 논문에서는 레이블링을 이용하여 얼굴 후보 영역을 추출하였다. 하지만, 얼굴내부의 미 추출된 요소

에 의해서 레이블링시 얼굴 영역이 분할되는 문제를 보완하기 위하여 2차 분석의 단계를 적용하였다. 이는 얼굴 내부의 요소들을 이용해 마할라노비스 거리를 1차 추출보다 20%정도 추가하여 확장하는 방법을 사용하였다. 이러한 연산을 한 후 얼굴 후보 영역을 레이블링 연산하면 보다 정확한 얼굴 후보 영역의 추출이 가능하다. 그림 4의 (a)와 (c)는 1차 추출 후의 영상이고, (b)와 (d)는 2차 추출 후의 영상이다. (a)와 (b)를 비교하면 1차 추출에 의한 잡영이 2차 추출 과정에서 사라졌고, 1차 추출에서는 목 부위가 선택되었던 반면, 2차 추출에서는 제거되었음을 확인할 수 있다. 또한 (c)와 (d)를 비교하면, 1차 추출시 얼굴의 좌측 뺨 부분에서 미 추출된 영역이 2차 추출에 의해서 많이 추출되었다.

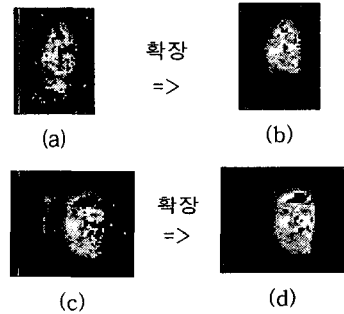


그림 4. 확장 추출된 얼굴 후보 영역

### 3. 응용 ACM

ACM은 찾고자 하는 영상의 주위에 제어점을 근간으로 하는 폐곡선을 생성시킨 후 생성된 제어점들을 관심 객체의 윤곽선에 매칭을 시키는 과정으로 이루어진다. M. Kass가 제안한 이 알고리즘은 물체의 윤곽선을 능동적으로 찾아주기 때문에 영상분할이나 객체 추적과 같은 분야에서 널리 적용되어지고 있다[5].

초기의 제어점의 선택은 얼굴의 후보영역의 최외곽 사각형에 내접하는 타원형으로 선택을 하였고, ACM의 에너지는 식 (4)와 같이 표현된다.

$$E = \int_0^1 (\alpha E_{inner}(v(S)) + (\beta E_{ext}(v(S)))) ds \quad (4)$$

$$S = [v_1, v_2, \dots, v_n]$$

$$v_i = (x_i, y_i)$$

여기서  $S$ 는 제어점의 집합이고,  $v_i$ 는 각각의 제어점들의 좌표를 나타하며,  $E_{inner}$ 는 내부에너지,  $E_{ext}$ 는 외부

에너지를 나타낸다. 일반적인 내부에너지는 식 (5)와 같이 제어점들의 곡선의 성질을 이용하고, 외부에너지는 식 (6)과 같이 영상에서의 화소값들의 변화량을 이용하여 에지에 근접하도록 하는 방법을 사용한다. 이렇게 계산된 에너지  $E$ 를 최소화 하는 방향으로 제어점을 움직이며 윤곽선을 찾아내는 것이다. 하지만, 모바일 영상에서는 조명의 영향이나 화질의 열화에 의해 화소값의 변화량이 일정하지 않게 나타난다.

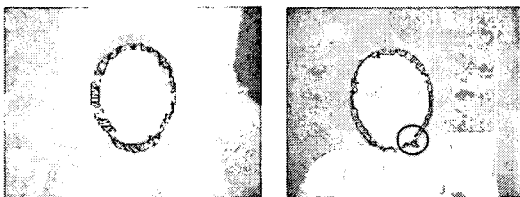
$$E_{inner} = \frac{1}{2} (\alpha(S)|g'(s)|^2 + \beta(S)|g''(s)|^2) \quad (5)$$

$$E_{ext} = \frac{G_{min} - G_i}{G_{max} - G_{min}} \quad (6)$$

그렇기 때문에 본 논문에서의 내부에너지는 일반적인 방법과 동일하게 제어점들의 곡선성을 식 (5)와 같이 사용을 하였으며, 외부에너지는 화소값의 변화가 아니라 얼굴 영역의 면적 변화를 식 (7)을 이용하여 판별하여 사용하였다..

$$\Delta E_{ext} = Area(S_i) - Area(S_{i-1}) \quad (7)$$

$Area(S_i)$ 는 폐곡선  $S$ 의 얼굴 영역 면적을 계산한 것이다. 이렇게 계산 되어진 전체 에너지  $E$ 를 최소화 하는 폐곡선  $S$ 를 이동하게 되는데, 외부에너지에 의한 제어점의 변화는 영상의 중심으로 이동하게 되고, 면적이 얼굴 영역에 일정하게 수렴하였을때 종료하게 하였다. 그림 5에서의 최초의 제어점은 T-S 색상 공간 분석에 의해 추출된 얼굴 후보 영역에 내접하는 타원을 생성하여 선정하였다. 그림 5의 (a)의 선은 최외곽이 초기의 제어점이고, 이들이 에너지에 의해 변형되면서 얼굴 영역으로 잘 정합되어진 것을 확인할 수 있고, 그림 5의 (b)는 하단에 얼굴 영역과 유사한 영역이 존재하기 때문에 원형영역에서처럼 제어점의 이동이 잘못되어졌지만, 대체적으로 얼굴 영역에 거의 유사하게 정합하였음을 알 수 있다.



(a) (b)  
그림 5. 추출된 얼굴 후보 영역

#### 4. 결론

지금까지의 ACM을 이용한 얼굴 영역 추출에 대한 대부분의 연구가 에지를 이용하였으나, 모바일 영상의 특성상 조명 및 해상도의 영향이 크고, 이에 의해 정확한 에지의 추출이 어려운 문제가 있었다. 그 뿐 아니라 ACM의 성능은 제어점의 초기 위치에 많이 의존하므로 이에 대한 초기 위치 선정이 어려운 문제들이 있었다.

이에 본 논문에서는 T-S공간 분석을 이용하여 보다 효율적인 제어점 선정을 하고, 얼굴 영역의 면적 변화량을 외부에너지로 한 ACM을 적용함으로써 얼굴 영역을 정확하게 추출할 수 있었다. 이렇게 추출된 얼굴 영역은 향후 자동 캐리커처 생성에 관한 개별 아바타 자동 생성 연구에 적합하게 사용 가능하며, 이를 위하여 얼굴 구성요소의 특징을 이용한 외부에너지를 적절하게 적용하는 연구 등이 추가적으로 진행되어야 한다.

#### [참고문헌]

- [1] 박재용, 박선영, 박종만 "가상캐릭터 애니메이션 기술," 정보과학회지, 제 19권, 제 5호, pp.24-29, 2001
- [2] 권해인, "히스토그램과 영역분할 기법을 이용한 얼굴추출에 관한 설계 및 구현," 컴퓨터정보통신공학과, 건국대학교, 2002
- [3] R.-L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face detection in color images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706, May 2002
- [4] Vytautas Perlibakas, "Automatic detection of face features and exact face Contour," Pattern Recognition Letters, vol. 24, pp. 2977-2985, 2003
- [5] M. Kass, A. Witkin and D. Terzopoulos, "Snakes : active contour models," International Journal of Computer Vision, vol. 1, no. 4, pp. 321-331, 1998