

T형 구조를 이용한 구성요소 기반 얼굴 탐지

황대동*, 김계영**

*승실대학교 컴퓨터학과

**승실대학교 컴퓨터학과

e-mail: chabert1@nate.com

gykim11@ssu.ac.kr

Component-Based Face Detection using T-type Structure

Dae-dong Hwang*, Gye-young Kim**

*Dept of Computer Science, Soongsil University

**Dept of Computer Science, Soongsil University

요 약

본 논문은 눈, 입 등 얼굴의 구성요소를 이용하여 얼굴을 보다 고속으로 탐지하는 방법을 제시한다. 얼굴 탐지 방법의 기본적인 절차는 피부색을 추출하여 SkinMap을 만들고, 그 내부에서 구성요소의 색상 특징을 이용하여 눈과 입을 검출 후, T형 구조 검증방식을 이용하여 추출한 후보요소들이 눈과 입에 해당하는지를 최종적으로 확인한다. 기존의 얼굴탐지 방식에 많이 사용되던 삼각형 구조 검증방식에 비해 보다 연산량이 적은 T형 구조 검증방식을 사용하고, 불필요한 방법을 제거 및 보완하여 보다 고속으로 얼굴을 탐지할 수 있다.

키워드 : 얼굴 탐지.

1. 서론

얼굴은 영상에서 탐지할 수 있는 여러 개체 중, 사람인지 아닌지를 판별할 수 있는 가장 강력하고 정확한 정보이다. 손과 발의 개수를 보고 몇 명의 사람이 있다고 추측할 수 있지만, 정확한 인원을 알 수 없으며, 그밖에 신체의 다른 부분은 의복으로 인해 특징이 천차만별이고 움직임, 개개인의 차이 등에 의해 형태가 일정하지 않아 쉽게 추출할 수 없다. 하지만 얼굴은 다른 신체부위와는 달리 특징이 쉽게 변하지 않고, 신체에 유일하게 하나만 존재하며, 크기 및 형태도 비교적 일정하기 때문에 영상에서 사람의 위치, 인원 수 등의 정보를 얻어야 하는 방법과 추적, 인식 등의 방법의 전처리 기술로 매우 유용하게 사용된다.

현재까지 연구된 얼굴 탐지 기술은 여러 가지 방법이 있다. 대표적인 방법으로는 Viola와 Jones가 제안한 Adaboost 알고리즘을 이용한 얼굴탐지 방법 [1]이 있으며, 이는 고속의 수행속도와 뛰어난 탐지결과로 인해 상당히 주목받고 있는 방법 중 하나이다. 하지만 이 방식은 학습을 통해 얼굴을 검출하기 때문에 학습되지 않은 예외적인 개체에

대해 좋지 않은 결과를 가져올 수 있으며, 회전된 개체의 검출에 취약하다. Hue와 Abdel-Mottaleb은 얼굴 구성요소 기반의 얼굴 탐지방법 [2]을 제안하였다. 이 방법은 피부영역을 추출하여 해당 피부영역들을 얼굴영역으로 가정하고, 피부영역 안에서 색차성분과 휘도성분을 이용하여 눈, 입 등 얼굴 구성요소를 검출, 조합하여 얼굴을 탐지하는 방법이다. 이 방법은 검출된 피부영역을 얼굴이라고 가정하고 구성요소의 존재 유무로 얼굴임을 판단하기 때문에, 일반적인 의복을 입고 촬영한 영상에서는 좋은 결과를 보이지만, 노출이 심한 영상 및 누드, 음란 영상같이 피부영역이 과다하게 검출되는 경우 오검출 되는 경우가 많다. 따라서 Lin과 Fan에 의해 제안된 삼각형 구조 기반의 얼굴 검증방법 [3]을 사용하여 구성요소들의 위치로 검증하는 방법을 이용하기도 하였는데 이 방법은 연산량이 많아서 구성요소 후보영역이 많이 추출되는 경우 상당한 시간이 소요된다는 단점이 있다. 따라서 본 논문에서는 삼각형 구조 기반의 많은 연산량을 보완한 T형 구조 기반의 얼굴 검증 방법 및 개선된 구성요소 추출 방식, 그에 따른 문제점 해결 방안을 제안한다.

II. 피부영역 추출

피부영역 추출은 Garcia가 제안한 방법 [4]을 사용하였다. 이 방법은 YCbCr 색상공간에서 영상의 명암값인 Y와 색상값인 Cb와 Cr을 이용하여 각각 8가지 방법으로 피부색을 검출한 후 그 중 가장 피부값과 가깝다고 예상되는 값이 임계값 안에 들어오면 피부영역이라고 결정하는 방식으로 명암의 변화에도 비교적 강건하게 피부를 탐지하는 방법이다.

$$Cr \geq \max -2(Cb + 24, -(Cb + 17), -4(Cb + 32), 2.5(Cb + \theta_1), \theta_3, 0.5(\theta_4 - Cb)) \dots (1)$$

and

$$Cr \leq \min(220 - Cb)/6, 4(\theta_2 - Cb)/3$$

$$\theta_1 = \begin{cases} -2 + (256 - Y)/16, & \text{if } Y > 128 \\ 6, & \text{otherwise} \end{cases}$$

$$\theta_2 = \begin{cases} 20 - (256 - Y)/16, & \text{if } Y > 128 \\ 12, & \text{otherwise} \end{cases}$$

$$\theta_3 = \begin{cases} 6, & \text{if } Y > 128 \\ 20 - (256 - Y)/16, & \text{otherwise} \end{cases}$$

$$\theta_4 = \begin{cases} -8, & \text{if } Y > 128 \\ -16 + Y/16, & \text{otherwise} \end{cases}$$

θ 값을 구한 후, 식 (1)을 이용하여 구하고자 하는 화소의 Cr값이 해당 범위 안에 속하면 피부값으로 인정한다. 추출된 비부영역은 레이블링을 통해 각각의 객체로 분류하고, 넓이가 임계값보다 작은 객체는 얼굴영역이 아닌 것으로 가정하고 제거한다. 임계값은 영상의 종류, 목적 등에 따라 틀리겠지만 본 논문에서는 영상 총 넓이의 1/100을 적용하였다.

III. 얼굴요소 후보 추출

본 논문에서는 눈과 입을 얼굴 구성요소로써 사용한다. 눈과 입을 검출하는 방법으로는 Hue가 제안한 YCbCr 색상공간에서의 구성요소 탐지 방법인 EyeMap과 LipMap [2]을 사용한다. EyeMap은 명암값의 대비를 이용한 EyeMapL과 색상값을 이용한 EyeMapC로 이루어져 있으나 EyeMapL은 연산량이 많으며, 예지가 뚜렷하지 않은 저해상도 영상의 경우 잘 추출되지 않는다는 단점이 있기 때문에 본 논문에서는 색상값을 이용한 EyeMapC만을 사용하여 눈을 추출하였다.

$$EyeMapC = \frac{1}{3} \{ (C_b^2) + (\tilde{C}_r)^2 + (C_b/C_r) \} \dots (2)$$

C_b^2 과 $(\tilde{C}_r)^2$, (C_b/C_r) 모두 [0,255] 사이로 정규화 한다. C_b 와 C_r 은 각각 YCbCr 색상공간의 Cb와 Cr값이며, \tilde{C}_r 은 역변환된 Cr 값(i.e., $255 - C_r$)이다. 일반적으로 눈 주변의 화소들은 Cb값이 높고, Cr값이 낮다는 관찰에 근거한 방식이다. 식 (2)를 통해 구한 EyeMapC는 눈 영역은 밝게 나오

고 눈을 제외한 다른 영역은 어둡게 나오는 결과를 보인다. 따라서 EyeMapC 영상을 이진화하고 레이블링하여 눈 후보영역을 추출한다.

입 후보 추출에 사용되는 Lip Map 또한 YCrCb 색상공간에서 추출한다. 이 방법은 입 영역에 속하는 화소들이 다른 영역에 비하여 Cr값이 높고, Cb값이 낮으며, (C_r/C_b) 특징에 대해서는 낮은 반응을 나타내지만, C_r^2 특징에 대해서는 높은 반응을 나타내는 관찰로 만들어 진다.

$$LipMap = C_r^2 \cdot (C_r^2 - \eta \cdot C_r/C_b)^2 \dots (3)$$

C_r^2 과 (C_r/C_b) 는 모두 [0,255] 사이로 정규화 한다. η 는 얼굴영역에 속하는 화소들의 C_r^2 과 (C_r/C_b) 평균의 비율이다. 이렇게 추출된 LipMap 또한 EyeMapC와 마찬가지로 입 영역은 밝게 나오고 입을 제외한 다른 영역은 어둡게 나오며, 이진화 및 레이블링을 통해 입 후보영역을 추출한다.

IV. 얼굴요소 후보 조합

추출된 피부영역에 얼굴요소 후보들을 이용하여 얼굴영역 인지를 검증하게 된다. 얼굴 구성요소인 눈과 입은 일정한 구조를 통해 배치되어 있으며, 이 때문에 기존에는 주로 삼각형 구조 검증방식 [3]을 통해 눈과 입의 배치가 얼굴의 구성요소 배치구조와 흡사한지 판단하여 얼굴인지 아닌지를 검증하였다. 하지만 이 삼각형 구조 검증방식은 정삼각형을 기본구조로 하여 검출하기 때문에 그림 3-(c) 처럼 정면 얼굴이 아닌 30° 이상 회전된 경우 구성요소의 구조가 정삼각형보다 직삼각형에 더 가깝기 때문에 잘 검출하지 못하는 문제가 발생하고, 삼각형수 계산이 최소한 두 번 필요하여 구성요소가 많은 경우 연산시간이 많이 필요하다는 단점이 있다. 따라서 본 논문에서는 삼각형 구조 검증방식을 대신하여 간단한 합연산으로 얼굴을 검증할 수 있는 T형 구조 검증방식을 적용하였다. T형 구조는 그림 2와 같이 두 눈 후보의 중심점 e_1, e_2 와 하나의 입 후보 중심점 m_1 , 그리고 두 눈 후보의 중심점을 연결하는 선 l_1 , 입 후보의 중심점과 l_1 을 수직으로 연결하는 선 l_2 로 이루어져 있다. T형 구조를 적용하기 위해 같은 피부영역 안에 존재하는 두 눈 후보의 중심점을 잇는 선 l_1 을 먼저 산출한 후, 역시 같은 피부영역 안에 존재하는 하나의 입 후보의 중심점과 l_1 을 직교하는 l_2 를 산출한다. 이 때 l_2 가 l_1 에 직교하는 점 p 가 두 눈 후보의 중심점 e_1, e_2 사이에 존재하고, 직교하는 점 p 에서 입 후보영역의 중심점 m_1 까지의 거리 즉 l_2 의 길이가 l_1 의 0.5~1.5배 사이에 해당한다면 얼굴이라고 확정한다. l_1 의 각도를 통해 얼굴의 회전각도를 구하고 식 (4)를 통해 l_2 와의 직교점을 정확히 찾아내므로 회전된 영상에 대해서도 문제없이 적용할 수 있다.

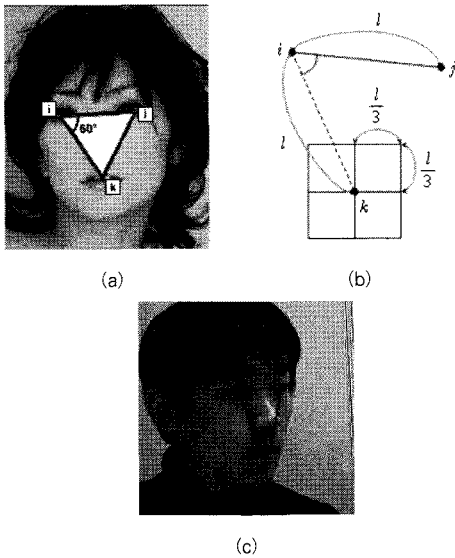


그림 1. 삼각형 구조
검증방식의 예 (a), (b) 및
회전된 얼굴에서의 문제점 (c)

$$x' = x \cos\theta + y \sin\theta, \quad y' = -x \sin\theta + y \cos\theta \quad \dots\dots (4)$$

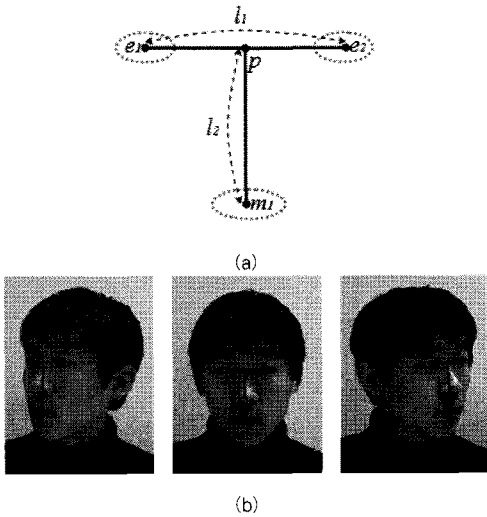


그림 2. 기본 T형 구조(a), 얼굴 회전각과 T형 구조(b)

제한한 방법에서는 눈 후보영역 검출 시 연산속도가 보다 향상될 수 있도록 하기 위해 EyeMapL을 사용하지 않았다. EyeMapL은 눈의 검은자와 흰자의 명암값 대비를 이용하여 눈썹이 눈으로 검출되는 것을 방지하지만 EyeMapL은 연산속도가 느리기 때문에 본 논문에서는 사용하지 않아 눈썹이 눈으로 오검출되는 경우가 발생할 수 있다. 눈썹 포함 눈 후보영역이 e_1, e_2, e_3, e_4 라면 하나의 눈 후보 e_i 와 직선 l_2 의 연장선상을 지나서 다른 눈 후보 e_j 와 연결되는 직선 l_1 은 그림

4와 같이 4개가 발생하게 된다. 이 4개의 l_1 중 그림 4-(d)처럼 가장 길이가 짧은 l_2 가 산출되도록 직교하는 l_1 의 두 눈 후보를 눈으로 판단한다. 이는 눈썹이 상대적으로 눈보다 입에서 멀리 떨어져 있다는 근거에 의한 방식이다.

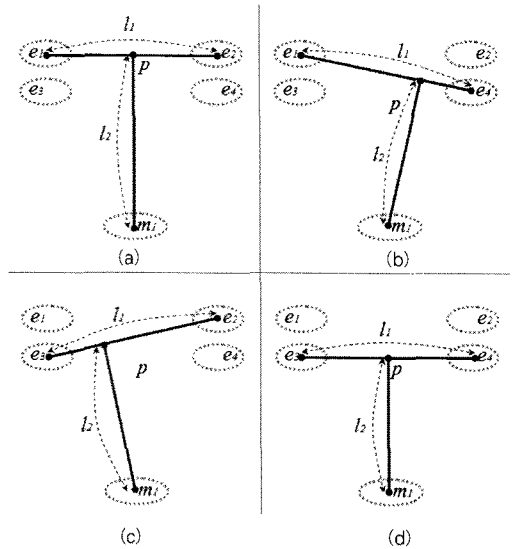


그림 3. 눈 후보영역으로 산출된 눈범영역 제거

V. 실험 및 결과

본 논문에서는 320*240, 640*480, 1024*768 3종류의 해상도 총 109개의 얼굴이 존재하는 80개 영상을 이용하여 정확도와 수행시간을 측정하였으며, 이와 비교하기 위해 EyeMapL을 포함하고 삼각형 구조 검증방식을 이용한 얼굴 탐지기도 동일한 방식을 적용하여 실험하였다.



그림 4. 얼굴탐지 결과

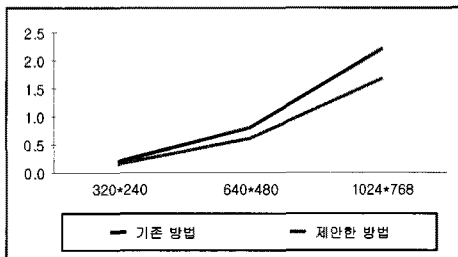
정확도는 얼굴을 정확하게 탐지하였는지를 보여주는 탐지율과 얼굴이 아닌 이미지를 얼굴로 잘못 탐지한 오탐지율로 나누어 측정하였다. 탐지율은 기존방법은 90.82% 제안한 방법은 92.66%로 두 방법 모두 비슷한 비율을 보였으나 제안한 방법이 보다 우수한 성능을 보였으며 30°~40° 정도 회전된 얼굴도 정확하게 검출하였다. 오탐지율은 기존방법은 21.25%, 제안한 방법은 13.75%로 기존 방법이 EyeMapL을 적용하였음에도 불구하고 일부 눈썹이 검출되는 영상의 경우 눈썹을 눈으로 인식하고 탐지하였던 반면에 본 논문에서는 눈썹을 눈으로 오인식 한 경우가 없었으며 주로 피부영역 또는 눈 영역이 제대로 검출되지 않아 오탐지한 경우가 많았다.

표 1. 기존 방법과 제안한 방법의 정확도 비교

방법	탐지율(%)	오탐지율(%)
기존 방법	90.82	21.25
제안한 방법	92.66	13.75

수행시간은 230*240, 640*480, 1024*768 3가지 해상도를 이용하여 109개 영상 중 10개 영상에 대해 측정한 후 평균을 통해 산출하였다. 실험 결과 제안한 방법이 기존 방법에 비해 24% 정도 수행속도가 빨라졌음을 알 수 있었다. 삼각형 구조가 T형 구조보다 계산상 더 복잡하기는 하지만 5% 미만의 수행속도 향상효과를 얻었을 뿐이며 EyeMapL을 사용하지 않고 눈썹영역 제거 방식을 사용한 결과가 수행시간 단축에 큰 영향을 미쳤다.

표 2. 기존 방법과 제안한 방법의 수행시간 비교



VI. 문제점 및 향후연구

본 방법은 구성요소를 이용하여 얼굴을 탐지하는 방법으로 비록 기존 방법에 비해 속도를 향상시켰으나 Adaboost를 이용한 방법 등 다른 방법들에 비해 비교적 느린 속도를 보여 고해상도 영상에서의 실시간 처리에는 적합하지 않다. 또한 얼굴 피부영역과 구성요소의 탐지를 너무 색에 의지해서 검출하기 때문에 조명, 광학기기에 의한 변색 또는 후보정을 통해 색상이 변경된 영상의 경우 잘 탐지하지 못하는 문제가 있

으며 피부 및 구성요소의 검출에 시간이 많이 소요된다는 문제도 있다. 따라서 피부와 얼굴 구성요소를 보다 빠르고 강건하게 검출하는 연구를 통해 피부, 구성요소들을 검출하는 시간을 줄이고 강건하게 검출할 수 있다면 T형 구조 검증방식을 이용한 얼굴 탐지 방법은 특별한 학습과정이 없이도 간단하게 얼굴을 탐지할 수 있는 좋은 방법이 될 것으로 보인다.

참고문헌

- [1] P. Viola, M.J. Jones. "Robust Real-Time Face Detection", International Journal of Computer Vision 57(2), pp.137-154, 2004.
- [2] R.L. Hsu, M. Abdel-Mottaleb, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, v.24, n.5, May 2002.
- [3] C. Lin, K.C. Fan, "Triangle-based approach to the detection of human face", Pattern Recognition Society, v.34, p.1271-1284, 2001.
- [4] C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis", IEEE Trans. Multimed. 1 (1999), pp. 264 - 277.