컬러 기반 영상에서 눈동자 템플릿을 이용한 얼굴영상 추출

최지영, 김미경, 차의영

부산대학교

A Face Detection using Pupil-Template from Color Base Image

Ji-Young Choi, Mi-Kyung Kim, Eui-Young Cha*

Pusan National University

E-mail : smil75@nate.com

요 약

본 논문에서는 얼굴을 검출하기 위해서 컬러영상에서 눈동자 템플릿을 사용한 얼굴을 검출하는 방법에 관해 제안하였다. 전체 시스템의 구성은 크게 피부색 모델에 의한 피부 영역을 검출하고 검출된 피부 영역에 타원을 적용하여 얼굴영역을 찾은 뒤 템플릿을 사용하여 눈동자를 추출하여 정규화 된 얼굴을 검출하는 단계로 이루어 졌다. 특히 타원을 적용할 때와 눈동자 템플릿을 적용할 때는 모멘트를 사용하였으며, 눈동자를 추출할 때 벡터를 기반으로 하기 때문에 사이즈에 제한 없이 자유로운 변형이 가능할 뿐만 아니라 눈동자 템플릿을 사용하면 눈동자의 위치와 크기를 동시에 얻어낼 수 있었다. 마지막으로 이렇게 검출된 얼굴을 기울어진 각도만큼 회전시켜 정규화 된 얼굴을 검출 할 수 있었다.

ABSTRACT

In this paper we propose a method to detect human faces from color image using pupil-template matching. Face detection is done by three stages. (i)separating skin regions from non-skin regions; (ii)generating a face regions by application of the best-fit ellipse; (iii)detecting face by pupil-template. Detecting skin regions is based on a skin color model. we generate a gray scale image from original image by the skin model. The gray scale image is segmented to separated skin regions from non-skin regions. Face region is generated by application of the best-fit ellipse is computed on the base of moments. Generated face regions are matched by pupil-template. And we detection face.

키워드

얼굴 검출(face detection), 가장 적합한 타원(best-fit ellipse), 눈동자 템플릿(pupil-template)

I. 서 론

생체인식이란 개개인마다 다른 생체정보를 인식하는 학문 또는 기술로써 얼굴인식은 대표적인 생체인식 분야이다. 특히나 보안관련 시스템의 요구와 시장이 급격하게 요구됨에 따라 생체인식 기술 분야의 연구가 더욱 활발히 이루어지고 있다. 본 논문에서는 얼굴인식의 입력영상이 되는 얼굴을 검출하기 위해서 컬러영상에서 눈동자 템플릿을 이용하여 얼굴을 검출하는 방법을 제안하고 있다.

본 논문에서 전체 시스템 구조는 그림1과 같다.

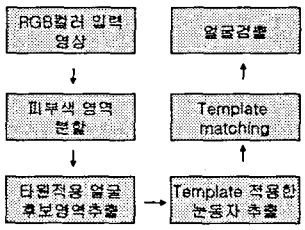

그림 1. 전체 시스템 개요도

II. 본 론

2.1 YCbCr컬러 모델을 이용한 피부색 모델

여러 실험을 통해서 다양한 인종의 다양한 피부색깔이라 할지라도 컬러 공간에서 긴밀한 결합을 갖는 클러스터를 형성한다는 것이 증명되었다[1]. 이것은 인종에 관계없이 사람피부의 컬러배합은 유사하다는 것을 의미하는데 이러한 사람의 고유 특징인 피부색을 이용해서 얼굴의 대략적인 위치를 검출 할 수 있다. 인간의 고유한 피부색깔의 주요한 변화 요인은 밝기(luminance) 변화에 기인하기 때문에 이 밝기성분인 휘도성분을 필터링한 모델을 적용할 필요가 있다. YCbCr컬러 모델은 Y(luminance)성분이 CbCr평면에 거의 영향을 미치지 않고 피부컬러 영역이 CbCr평면에서 N(m, C)인 가우시안 분포로 나타내지어 얼굴검출에 이용하기 적당하다. 본 논문에서도 Y성분을 배제한 CbCr영역에 피부색영역을 나타낸 후 3차원의 가우시안 분포를 사용한 피부색 영역을 통해 피부색 모델을 구축하고 얼굴 영역을 검출하였다.

YCbCr은 컬러 정보로부터 광도를 분리하는 또 하나의 컬러 공간이다. 광도는 Y로 기호화되고 푸른 정보와 붉은 정보는 CbCr로 기호화 된다. 컬러를 기본으로 한 피부 영역과 피부 아닌 영역을 분리하기 위해서 조명의 변화와 다른 피부색의 사람들에게도 적응할 수 있는 신뢰성 있는 피부색 모델이 필요하다[2]. N(m , C)인 가우시안 분포로 나타내어 질 수 있는 피부색 분포 특징을 이용하여 YCbCr 컬러 정보로 변환된 입력 영상에서 피부색 영역만 추출하는 피부색 모형을 생성할 수 있다. 그림2는 다양한 인종의 사람들에게서 추출한 컬러 이미지로부터 얻어진 피부색 샘플들은 잡음의 효과를 줄이기 위해 Loss-Pass Filter를 적용했으며 다음과 같이 주어진다.

$$\frac{1}{9}\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix} \qquad (1)$$

그리고 주어진 피부 영역의 데이터들로부터 얻어진 Cb성분과 Cr성분의 평균과 표준편차를 이용하여 가우시안 확률 밀도 함수를 구하여 3차원의 CbCr 공간상에 나타낸 것이다. N(m, C)과 확률밀도 함수 식은 다음과 같다.

$$m = E\{x\} \qquad (2)$$

$$C=E\{(x-m)(x-m)^T\} \qquad (3)$$

$$P(r\ b) = \exp[-0.5(x-m)^T C^{-1}(x-m)] \qquad (4)$$

$$x = (r,b)^T \qquad (5)$$

따라서 피부컬러 모델은 컬러이미지에서 그레이 스케일 이미지로 변환되는데 각 픽셀의 그레이 값은 피부가 속해 있는 픽셀의 확률을 보여준다. 그리고 그레이 스케일 이미지는 피부영역인지 아닌지를 보여주는 이진화(thresholding)된 영상으로 변환 된다.

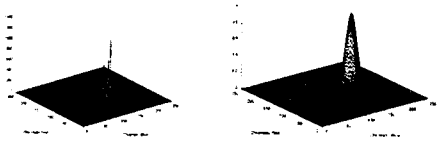

그림 2. 피부색 영역의 가우시안 분포

2.2 피부색 영역 분할

입력된 RGB영상에서 YCbCr로 변환한 모든 픽셀들을 식 (4)와 (5)를 기본으로 피부후보 영역을 결정하고 이렇게 결정된 영역이 피부색 영역이면 255 gray value, 그렇지 않으면 0 gray value로 나타내어 피부색이 검출되어진 256 그레이 영상을 얻는다. 그림3의 실험 결과에서도 보았듯이 모든 피부 영역은 피부 영역 아닌 것 보다 더 밝음을 보여준다. 피부 영역에서 다른 인종의 다른 피부를 가진 사람들의 피부는 다른 확률을 가지기 때문에 고정된 이진화 값으로 그 영역을 찾기는 쉽지 않다. 그래서 적응(adaptive) 이진화 과정을 통한 최적의 이진화 값을 얻어야한다. 적응적 이진화 과정은 전체 영상의 히스토그램을 이용하는 것이 아니라 영상의 일부분에 대한 히스토그램을 가지고 그 일부분만을 위한 임계값을 계산하기 때문에 효과적으로 피부색 영역과 피부 아닌 영역을 분할하며 다른 인종의 다른 피부색 사람들에게도 좋은 결과를 가져온다.

그림 3. 원영상에서 추출된 피부색 영역

2.3 타원적용 얼굴영역 검출

분할된 피부색 영역에 사람의 얼굴이 형태학적으로 타원 형태를 띈다는 사실을 이용하여 본 논문에서는 모멘트를 사용한 타원을 적용함으로써 얼굴영역을 검출한다[3]. 여기서 사용되는 타원은 중심이 $(\bar{x}, \bar{y})$ 이고, 방향은 θ, 길이는 단축 a, 장축 b로 정의한다.

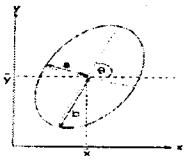

그림4. 타원의 변수

타원의 방향 θ는 최소 관성모멘트(moment of inertia)에 의해 계산되어 질 수 있다.

$$\theta = \frac{1}{2} \cdot \arctan\left(\frac{2\mu_{1.1}}{\mu_{2.0} - \mu_{0.2}}\right) \quad (6)$$

$\mu_{i,j}$는 연결성분(connected component)의 중심 모멘트(central moments)이고, 최적의 타원에서의 장축과 단축의 길이도 관성모멘트에 의해서 계산되어 질 수 있다. 방향이 θ일때 최소 관성모멘트 $I_{\min}$과 최대 관성모멘트 $I_{\max}$는 다음과 같다.

$$I_{max} = \sum_{(x,y) \in C} [(x - \bar{x})\sin\theta - (y - \bar{y})\cos\theta]^2 \quad (7)$$

$$I_{min} = \sum_{(x,y) \in C} [(x - \bar{x})\cos\theta - (y - \bar{y})\sin\theta]^2 \quad (8)$$

길이 a축과 b축의 결과는 다음과 같다.

$$a = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{(I_{\max})^3}{I_{\min}}\right]^{1/8} \quad (9)$$

$$b = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{(I_{\min})^3}{I_{max}}\right]^{1/8} \quad (10)$$

이렇게 계산된 타원의 요소들을 기본으로 얼굴 후보 영역의 수들을 감소시키는 것이 가능하다. 즉, 얼굴 후보 영역에서 타원의 적용은 장축과 단축사이의 관계와 타원의 방향으로 타원의 표준을 결정하여 적용함으로써 수행된다. 그리고 남아있는 피부 영역 후보들에 그것의 최적의 타원 적용에 의해 얼마나 연결성분이 잘 접근하였는지 평가해야한다. 이러한 목적으로 사용되는 평가함수는 다음과 같다.

$$V = \frac{\sum_{(x,y)\in E}((1 - b(x,y)) + \sum_{(x,y)\in C/E} b(x,y)}{\sum_{(x,y)\in E} 1} \quad (11)$$

$$b(x,y)\begin{cases} 1 & \text{단},(x,y)\in C \\ 0 & otherwise \end{cases} \quad (12)$$

그림 5. 최적의 타원 적용 결과

이 평가함수 V는 타원 밖의 연결성분의 점들과 타원 안의 홀(hole)수에 의해 최적의 타원과 연결성분의 관계를 결정한다. 즉 평가함수 V의 값이 작을수록 타원에 가깝고 클수록 타원과 멀어진다. 이 평가 함수로 타원 안의 점들의 수에 잘못된 점들의 수의 비율이 계산되며, 이 비율에 있어서 경험적인 임계값을 기본으로 타원의 알맞은 연결성분을 선택하고 피부 영역에서 얼굴의 후보들이 고려되어진다. 이렇게 고려된 얼굴의 후보 영역에서 얼굴의 특징적인 부분을 이용해서 최종적인 얼굴을 검출하게 된다[4]. 그림 5는 타원을 이용해서 검출된 얼굴 영역을 보여준다.

2.4 눈동자 추출

앞의 단계에서 찾아진 얼굴후보 영역에 특징성분인 눈동자를 추출함으로써 얼굴인식에 필요한 최종적인 얼굴을 검출 할 수 있다. 본 논문에서는 눈동자를 추출하기 위해서 템플릿 매칭을 사용하였으며 눈동자 추출을 위한 알고리즘은 다음과 같다.

[단계1] 피부색 모델을 통해서 분할된 피부 영역에 타원을 이용한 얼굴 영역을 추출하여 그 후보 영역에 명암값의 수직 수평 투영을 통한 최종적인 눈동자를 찾기 위한 후보영역을 추출한다.

[단계2] 좌측눈 영역과 우측눈 영역 템플릿매칭 적용하며 다음과 같은 수식을 이용한 템플릿을 사용하여 눈동자 후보영역에 각각의 템플릿을 적용시킨다.

$x^2 + y^2 \le r$
x = 1, 2, ···, N1, (수평인텍스)
y = 1, 2, ···, N2, (수직인텍스)
r = 1, 2, ···, n (반지름)

[단계3] 3x3영역의 픽셀단위를 수행하며 9개의 픽셀 중 가장 작은 명도값을 가진 픽셀이 결과 영상의 명도 값으로 결정되는 침식연산을 사용하여 눈동자에 반사된 빛을 제거한다.

[단계4] 템플릿을 적용하여 최종적인 눈동자를 추출한다.

먼저 눈동자를 추출하기 위해 그림6과 같이 수직 수평투영을 통한 후보영역을 검출한다.

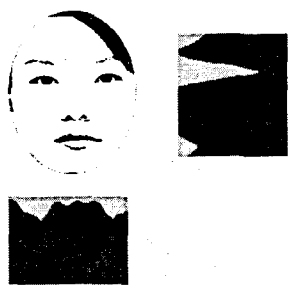

그림 6. 명암값의 수직 수평투영

이렇게 얻어진 전처리 과정으로 눈동자에 반사된 빛을 제거하기 위해 침식연산을 사용한다. 침식연산은 기본적인 3x3마스크의 형태이다[5][6]. 그리고 본 논문에서 사용되어진 템플릿은 단계2의 수식을 사용하였다.

III. 실험 및 결과분석

본 논문에서는 총 240개의 JEPG이미지를 사용하였고, 그 이미지에 포함된 눈의 총개수는 800개이다.

그리고 수식의 x, y는 각각 이미지의 수평, 수직 인덱스이며 r은 반지름이다. 여기서 얻어진 특징 벡터 값을 이용하여 최적의 눈 후보를 찾아내게 된다.

X : 수평성분의 특징벡터
Y : 수직성분의 특징벡터
R : 반지름 성분의 특징벡터
S : 반지름이 같은 템플릿 내에서 모멘트 값
V : 반지름이 서로 다른 템플릿 내에서 모멘트 값(2차모멘트 분산)

사람의 눈은 흰자와 검은자로 구성되어지기 때문에 최적의 눈동자의 크기를 얻었을 때 명암의 변화율이 가장 낮아서 분산을 사용하게 된다. 본 논문에서는 V값으로 2차 모멘트(분산)를 사용하였다. 그리고 단순히 같은 반지름을 갖는 템플릿 중에서 분산 값이 가장 작은 템플릿을 선택하는 것이 아니라, r_{i-1}의 분산과 r_i의 분산차가 가장 클 때의 원을 최적의 원으로 보았다. 그림7은 템플릿의 반지름을 수렴시켜 가면서 S값을 비교 하여 값이 가장 큰 지점을 표시한 것이다.

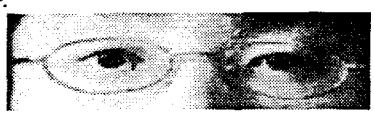

그림 7. 탐색과정

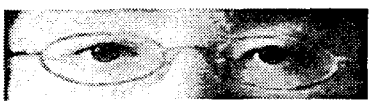

그림 8.추출된 눈동자의 위치와 크기

그림8은 각 반지름별 최적 후보끼리 분산차를 구하여 분산차가 가장 큰 최종 최적해를 표시하고 있다. 즉, 눈동자의 위치와 크기를 동시에 얻어 낼 수 있다는 것을 알 수 있다. 이렇게 얻어진 눈의 좌표를 이용하여 이미지를 다시 회전시킨다. 그림9는 원영상에서 눈동자를 추출한 결과 영상이다.

그림 9. 템플릿을 사용한 눈동자 검출

본 논문의 눈동자 템플릿 매칭을 이용한 얼굴 검출에 관한 검출률은 표1, 2와 같다.

표1. 얼굴영역 추출 성공률

JPEG이미지	성공	실패
240	207(86.25%)	33(13.75%)

표2. 눈동자 추출 성공률

눈의 총개수	VERY GOOD	GOOD	BAD
800	412(51.5%)	286(35.75%)	102(12.75%)

IV. 결론

본 논문에서는 컬러 영상에서 눈동자 템플릿을 사용한 얼굴검출 시스템을 수행하였다. 본 논문의 얼굴검출 과정에서도 영상의 배경색이 부분적인 피부색일 때는 얼굴검출에 큰 영향을 주지 않지만 전체적으로 피부색과 유사한 배경은 얼굴이 배경과 함께 피부색으로 인식되어 타원을 찾지 못하는 경우가 있었으며, 눈을 감고 있는 상태에서는 눈동자 템플릿을 이용한 눈동자를 찾을 수가 없었다. 그리고 타원을 적용해서 얼굴 영역을 찾고 템플릿 매칭을 하였기 때문에 타원을 못 찾은 경우의 눈동자 추출은 배제 되었다.

본 논문의 결과에서 알 수 있듯이 향후에는 전체적으로 피부색과 유사한 배경색에도 안정적으로 얼굴과 특징성분을 추출할 수 있는 시스템이 연구되어 져야 겠다.

참 고 문 헌

[1] J. Cai, A. Goshtasby, "Detection Human Faces in Color Image", Image and Vision Computing, pp.63-75, 18, February 1999

[2] Jie Yang and Alex Waibel, "A Real-Time Face Tracker", CMUCS Technical Report .

[3] A. K.Jain, "Fundamentals of Digital Image Processing", Prentice Hall, 1989

[4] Karin Sobottka, Ioannis Pitas, "Extraction of Facial Regions and Features Using Color and Shape Information",in IEEE Proceedings of ICPR'96

[5] 장동혁, "디지털 영상처리의 구현", 정보게이트, pp200-208, 2001

[6] 민경필, 한태화, 전준철, "2차원 얼굴 영상을 이용한 3차원 얼굴 모델링 및 메쉬 최적화", proceedings of the 6th KSII Fall Conference(3권 2호), pp.418-421, Nov.23, 2002