

TSL 색상 정보와 기하학적 구조를 이용한 얼굴 포즈 추정에 관한 연구

김성환, 최재영, 김낙빈
경원대학교 소프트웨어대학

A Study on Facial Pose Estimation using TSL Color Information and Geometrical Structure

Sung-Hwan Kim, Jae-Young Choi, Nak-Bin Kim
College of software, Kyungwon University.

요 약

본 논문은 컬러 입력 영상에서 검출된 얼굴 영역 내의 홀(hole)들간의 기하학적 구조를 이용하여 포즈를 추정하는 방법을 제시한다. 얼굴 영역 검출에서는 특징값 기반의 알고리즘 중 피부색 분포를 이용하는 방법을 적용하며, 이 때 발생하는 조명에 의한 열화를 제거한다. 본 논문에서는 TSL 색상 모델을 사용하고, 조명에 의해 너무 밝게 표현되는 부분의 피부값을 조정함으로써 조명에 대한 보정을 실시한다. 그런 다음, 얼굴 영역 안에서 찾은 홀을 피부영역이 아닌 얼굴 구성요소(양눈, 입)로 가정하여, 후보 구성요소들의 기하학적 구조를 이용해 다양한 포즈의 입력 영상에 대한 포즈를 추정한다. 추정된 값은 향후 다양한 포즈에 대한 특징점 추출이나 얼굴 인식에 활용될 수 있다.

1. 서 론

최근 들어 일상생활에서 기계를 통한 인증 및 인식의 폭이 확대되어가고 있는 가운데 기존의 인증 도구로 사용되어 왔던 비밀번호에 대한 보안 침해 사례가 늘어나고 있다. 그래서 등장하게 된 인증/인식 방법이 인간의 생물학적 특징을 이용한 생체인식(Biometrics) 분야이며, 이는 지문인식을 필두로 홍채인식, 음성인식, 정맥인식, 얼굴인식 등으로 그 응용분야가 넓어지고 있다. 더욱이 최근 생체인식에 관한 실용화가 수행되면서 좀더 편하고 안정적인 생체인식에 관한 연구가 증대되고 있는 추세이다.

그 중 얼굴인식 기술은 지문인식이나 홍채인식과는 달리 자신의 신체 일부를 인식장치에 직접 접촉시키지 않아도 되고, 생체정보의 획득에 있어서 강제성이 적어 사용자로 하여금 거부감과 불편함을 주지 않는다는 장점을 지니고 있다. 얼굴인식 기술은 이러한 비접촉성, 비강제성, 편리성 등의 특징들로 인해 신원확인 분야 외에도 사용자가 모르는 사이에 자연스럽게

정보를 획득해야 하는 지능형 무인감시 등의 분야에 중요하게 활용될 수 있다. 하지만, 이러한 획득방법도 실시간으로 변화되는 순간적인 현장 영상을 포착하기 때문에 접근해오는 얼굴 영상의 촬영 각도에 따라 얼굴 영상의 입력이 달라지며 이를 판단하여 보정하는 추가적인 기술이 필요하다.

따라서, 본 논문에서는 컬러 정지 영상에서 포즈가 다양한 얼굴들의 포즈를 추정하기 위한 알고리즘을 제안한다. 제안한 알고리즘은 크게 2단계로 구성되는데 1단계는 얼굴 검출단계, 2단계는 포즈 추정 단계로 구성된다. 1단계는 먼저 입력 영상이 컬러 영상이므로 얼굴 영역을 추출하기 위한 기본적인 방법으로 많이 쓰이는 색상 정보를 이용하는데, 그 중 피부색 분포에 적합한 TSL 색상모델을 적용하였다[1]. 그런 다음, 검출된 영역 내에서 4방향 탐색을 통해 홀(hole)을 찾고, 조명 보정을 하여 피부영역을 확장한다. 2단계는 포즈 추정 단계로서 다양한 각도의 입력 영상에 대한 포즈를 추정하는데, 1단계에서 추출된 얼굴 영역에서

얼굴 구성요소(양눈, 입)를 찾는다. 이렇게 찾은 구성요소의 중점을 연결한 삼각형과 전체 얼굴 영역의 기하학적 구조를 통하여 다양한 각도의 영상에 대한 포즈를 추정한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문과 관련된 기존의 연구들을 기술하며, 3장에서는 본 논문에서 제안한 내용인 TSL 색상모델을 통해 얼굴 영역을 검출하고 다양한 각도의 입력영상에 대한 포즈를 추정하는 방법을 설명한다. 4장에서는 실험 결과를 기술하고, 마지막으로 5장에서 결론 및 향후 연구과제에 대해 논하도록 한다.

2. 관련 연구

본 장에서는 제안하는 방법과 관련된 기존의 얼굴 검출과 포즈추정의 방법을 소개한다.

2.1 얼굴 검출

얼굴 검출은 인식 시스템에서 처리할 얼굴부분의 데이터만을 제공해주므로 전체 시스템에서 아주 큰 비중을 차지하며, 독립적인 분야로서 많은 응용분야를 가지고 있어서 연구의 필요성이 높아지고 있다.

초기의 연구들은 T. Sakai 등이 연구한 단순한 템플릿 매칭기법을 시작으로 발전되었다[2]. V. Govindaraju의 연구에서는 얼굴의 비례를 이용하였는데[3], 얼굴의 타원형 특성을 응용한 이 방법은 오늘날까지도 Hough 변환 등의 방법으로 사용되고 있다. 하지만 이러한 방법들은 탐색 및 연산 시간이 매우 오래 걸린다[4].

일반적으로 손쉽게 얼굴을 탐지하는 방법은 색상모델을 이용하는 방법인데, 피부색을 이용하여 얼굴 영역을 미리 찾으면 보다 쉽고 정확하게 특징들을 추출할 수 있으므로 본 논문에서도 이 방법을 사용한다.

색상모델은 매우 다양한데, H. Yao[5]는 RGB 색공간을 YUV 색공간으로 변환하여 피부색을 모델링하였으며[5], Jie Yang 과 Karin Sobotka는 각각 정규화된 RGB 영상과 HSI 색상 모델을 이용하여 얼굴 영역을 검출하였다[6,7].

2.2 포즈 추정

S. Gong은 다각도로 회전된 얼굴의 모델 데이터를 만들어 입력 영상과 프로토타입 데이터 베이스와의 유사도를 계산하여 포즈를 추정하였으며[8], J. Huang는 일반적 얼굴 모델에 기반한 SVM(Support Vector Machine)을 이용하여 얼굴의 위치를 추적하였다[9]. 하지만, 이러한 방법들은 각 회전각에 대한 모델 데이

터를 만들기 위한 훈련 영상이 매우 많이 필요할 뿐만 아니라 유사성 측정을 위한 수단을 잘 모델링 해야하는 어려움이 있다.

J. Heinzman은 템플릿을 매칭 방법을 사용하여 얼굴 안의 기하학적 정보의 이동을 측정하였으며, 이를 머리 제스처 인식에 사용하였다[10].

또한, 동영상의 경우 각 프레임간의 차이에서 발생하는 움직임 벡터를 이용하여 얼굴의 포즈를 추정하는 연구도 진행되고 있다.

3. 제안한 알고리즘

그림1은 제안하는 알고리즘의 전체 흐름도를 나타내는데, 먼저 TSL 색상 모델로 얼굴 영역을 검출하고, 홀을 탐색하여 조명에 의해 너무 밝게 표현되는 피부값을 조정함으로써 조명 보정을 실시한다. 그런 다음, 피부 영역이 아닌 홀들 중에서 얼굴 구성요소를 선정하여, 구성요소와 얼굴 영역간의 기하학적인 구조를 통하여 포즈를 추정하게 된다.

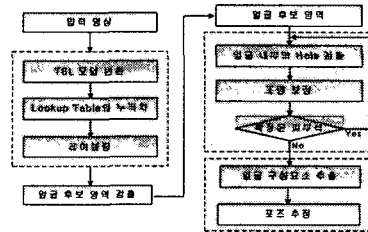


그림1. 제안 알고리즘의 전체 흐름도

3.1 얼굴 검출

본 논문에서 얼굴 후보 영역을 추출하기 위해 기본적으로 많이 사용되는 피부색을 이용한 색상 모델 방법을 사용하였는데, 색상 모델로는 피부색 분포에 적합한 TSL 색상 모델을 사용하였다. 입력된 RGB 영상을 TSL 색상 모델로 변환하는 방법은 식(1)과 같다[1].

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

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

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

$$r' = (r - 1/3), g' = (g - 1/3)$$

$$r = R/(R + G + B), g = G/(R + G + B)$$
(1)

식(1)을 사용하여 한국인의 피부 샘플을 학습시킨 후, 각각의 T-S 값에 대한 누적분포 록업 테이블을 생성한다. 생성된 누적분포 록업 테이블의 상위 90% 이상의 누적치를 통하여 얼굴 영역을 검출하고, 검출된 얼굴 영역에 대한 잡음 제거를 위해 레이블링을 시행한다. 그림2의 (b)가 얼굴 검출 결과를 나타낸 것이고, (c)는 레이블링을 통해 잡음이 제거된 영상이다.



그림2. 얼굴 검출과 레이블링

잡음제거를 위한 방법으로 모폴로지나 레이블링이 있는데 본 논문에서는 모폴로지를 시행하면 검출된 홀들이 메워지기 때문에 적합하지 않으므로 레이블링을 시행하였으며, 가장 큰 영역을 선택함으로써 얼굴 영역을 확보할 수 있다.

3.2 홀(hole) 검출

일반적으로 얼굴 구성요소를 추출하기 위해서는 얼굴의 눈과 입과 같은 수평 방향 정보를 이용하지만, 얼굴이 틀어진 경우는 적합하지 못하기 때문에 본 논문에서는 검출된 얼굴 영역 내부의 홀을 탐색하여 얼굴 구성요소로 가정하게 되는데, 그 방법은 그림3과 같이 4방향으로 탐색을 하여 3방향 이상이 막혀 있다면 홀로 판단하게 된다.

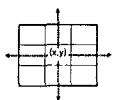


그림3. 4방향 탐색



그림4. 홀(hole) 검출 결과

그림4의 (a)는 4방향 탐색하여 3방향이 막혀 있기 때문에 홀로 판단된 경우이고, (b)는 2방향만 막혀 있기 때문에 홀로 판단할 수 없는 경우이다.

3.3 조명 보정

보정시 조명을 분석하기 위해 TSL 중 식(1)의 L을 사용한다. 얼굴 영상의 중심 좌표를 기준으로 밝기값이 평균 이상에 해당하는 각도는 조명의 영향을 받은 부분이 된다. 식(2)는 각 각도에 해당하는 조명의 평균값을 산출하는 수식인데, 여기서 각도별 조명의 양

을 측정하였으면, 산출된 각도 범위에 해당하는 얼굴 영역의 홀들에 대한 피부값을 조정하여 조명 보정을 시행하는데, 일반적으로 시행하는 영상 전체에 대한 조명 보정이 아니라 조명의 영향을 받은 부분만을 보정하게 된다.

$$L(\theta) = \frac{\sum f(x,y)}{N_\theta} \quad s.t. \quad f(x,y) \in Lface$$

$$\theta = \arctan[(f_y - CL_y)/(f_x - CL_x)/\pi] \times 180 \quad (2)$$

CL: 얼굴색 영역 Lface의 무게 중심
N_θ: 각도θ에 해당하는 Lface의 화소수

그림5의 (b)는 조명 방향의 각도를 분석한 것이고, (c)는 영상에 대한 조명의 방향을 나타낸다. 결과적으로 (d)의 레이블링 결과와 (e)보정 결과를 비교해 보면 조명 방향의 홀들이 보정된 것을 확인 할 수 있다.

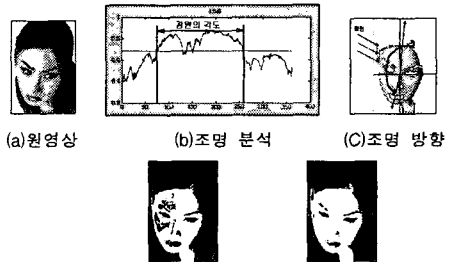


그림5. 조명 보정 결과

3.4 얼굴 구성요소 추출

조명 보정 결과로 검출된 얼굴 영역에서 얼굴 구성요소(양눈, 입)를 추출해야 하는데, 먼저 입의 경우는 얼굴 검출과 마찬가지로 방법으로 TSL 모델을 이용하여, 입 영역만을 검출하여 검출된 입 영역과 유사한 위치의 홀을 입으로 추출한다. 양눈의 경우는 그림6과 같이 얼굴 영역의 홀들 중에서 군집이 큰 홀들을 검출하여, 이미 검출된 입 영역의 양 끝점을 지나는 직선에 수직인 직선으로 L-part와 R-part를 구분한다.

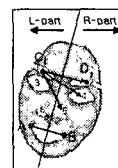
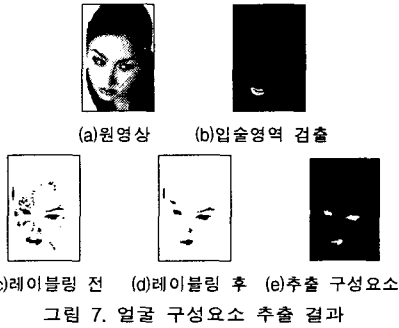


그림6. 얼굴 구성요소 후보

각 part의 홀과 홀의 거리와 각도를 동시 연산하여 식(3)의 조건에 맞는 홀들을 양눈으로 가정하며, 식(3)에서 A, B는 입의 양끝점이고, C_i, D_j는 각 part의 홀들이다.

$$\begin{aligned} & \bullet \overline{AB} \approx \overline{C_i D_j} \mid s.t \text{ distance}(\overline{AB}, \overline{C_i D_j}) < \epsilon \\ & \bullet \overline{AB} \parallel \overline{C_i D_j} \end{aligned} \quad (3)$$

그림7의 (b)가 입술 영역을 검출한 결과이고, (c)는 미세한 홀들의 레이블링 전 모습, (d)는 레이블링 후의 군집이 큰 6개의 홀과 입영역을 나타낸 것이다. (e)는 최종 얼굴 구성요소인 양눈과 입을 선정한 결과이다.



3.5 포즈 추정

선정된 얼굴 구성요소와 얼굴 영역의 기하학적 구조를 통하여 입력 영상에 대한 포즈를 추정하게 되는데, 입을 기준으로 턱밑 부분은 제외한 순수한 얼굴 영역의 중심좌표와 구성요소들의 중심을 꼭지점으로 하는 삼각형의 중점 사이의 기하학적 구조를 통하여 포즈를 추정한다. 얼굴 영역의 중심좌표에서 삼각형의 중점으로서의 벡터가 표시되는데, 이 벡터가 추정된 포즈값을 의미한다.



그림8. 포즈 추정 결과

그림8의 (a)는 얼굴 구성요소를 연결한 삼각형의 중점이고, (b)는 십자선으로 표시한 턱밑을 제외한 순수 얼굴 영역의 중심좌표를 표시한다. 그리고, (c)는 얼굴

영역의 중심좌표에서 삼각형의 중점까지의 벡터를 나타내는 포즈 추정 결과이다. 여기서 벡터의 방향은 얼굴 포즈의 방향이고, 길이는 방향으로의 이동치를 의미하며 그 길이를 분석하여 좌우회전(yaw)과 상하회전(tilt)의 각도를 산출한다.

4. 실험 결과

본 장에서는 제안한 알고리즘을 바탕으로 구현 및 실험한 결과를 보이도록 하겠다. 실험을 위한 데이터는 다양한 포즈의 영상으로 인터넷을 통하여 획득하였고, 그림9는 영상의 다양한 포즈를 정량화하기 위하여 좌우회전(yaw)과 상하회전(tilt)상의 각도를 측정한 것이다.

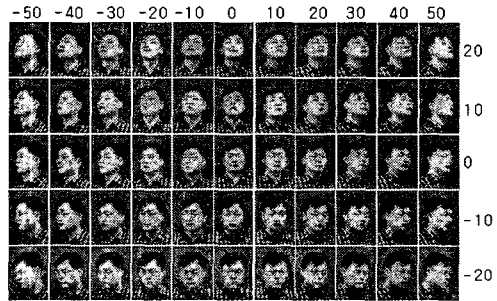


그림9. 각도별 얼굴의 다양한 포즈

표1은 알고리즘의 각 단계별 결과 영상을 나타낸 것인데, 얼굴 검출에 있어서 1번과 2번 영상에서는 얼굴 검출 후 조명 보정을 통해 피부 영역이 확장되었고, 3번 영상은 분리되지 못했던 입술 부분이 조명 보정을 통해 보정된 것을 볼 수 있다. 1번, 2번 영상에서의 입술검출은 색상 정보의 부족으로 충분히 검출되지 못했는데, 이는 입술색 학습의 데이터 수를 늘림으로써 오류를 수정할 수 있다.

표1. 알고리즘의 각 단계별 결과

	입력 영상	얼굴 검출	조명 보정	입술 검출	구성요소 추출	포즈 추정 벡터	포즈 추정치
1							방향 : 우측/아래 좌우 : 19 도 상하 : -7 도
2							방향 : 좌측/위 좌우 : -13 도 상하 : 6 도
3							방향 : 우측/아래 좌우 : 16 도 상하 : -3 도

이렇게 추출된 구성요소와 얼굴영역의 기하학적 구조를 통해 포즈 추정 벡터가 구해지고, 이 벡터를 통해 포즈의 틀어진 방향과 그 방향에 대한 각도가 산출된다. 1번의 경우는 영상을 보는 사람의 입장에서 우측 아래 방향을 향하고 있으며, 벡터의 길이를 통해 우측으로 19도, 아래로 7도의 각도가 산출되며, 2번의 경우는 반대로 좌측 위 방향을 향하고 있으며, 좌측으로 13도, 위로 6도의 각도가 산출된다. 3번의 경우는 1번과 마찬가지로 우측 아래 방향을 향하고 있으며, 우측으로 16도, 아래로 3도의 각도가 산출된다. 이와 같이 검출된 얼굴 영역내에서 추출된 구성요소들과 얼굴 영역의 기하학적 구조를 통해 얻어진 벡터를 분석하여 입력 영상에 대한 포즈 추정치를 얻어 낼 수 있다.

5. 결 론

본 논문에서는 얼굴 검출 방법으로 색상 모델을 이용하는 방법을 제시하였다. 색상 모델을 이용할 경우 고려해야할 점은 적절한 색상 모델의 선택과 조명에 의한 영향을 가능한 제거해 주는 것이다. 이에 제안한 방법에서는 전처리로서 TSL 색상 누적분포를 이용하여 얼굴 검출을 시행하고, 광원의 각도에 따른 조명에 대한 영향을 제거하여 주었다. 그리고, 얼굴 구성요소로 가정되는 홀들 중에서 얼굴 구성요소인 양눈과 입을 선정하여 연결한 삼각형과 순수 얼굴 영역의 중심 좌표와의 기하학적 구조를 이용하여 다양한 각도의 입력 영상에 대한 벡터가 산출되는데, 이 벡터의 방향이 포즈의 방향을 의미하며, 벡터의 길이는 이동치를 의미한다. 구해진 벡터를 분석해 좌우 및 상하의 회전 각도를 추정할 수 있었다.

향후 이렇게 추정된 포즈값을 정규화된 기본 템플릿에 적용하여 가변 템플릿을 생성하고, 이를 이용해서 얼굴의 특징점을 좀더 정확히 추출하는 연구가 진행되어야 할 것이다. 또한, 이렇게 추출된 특징점은 얼굴 인식이나 3차원 모델링 및 애니메이션 등과 같은 곳에 적용이 가능하다.

[참고문헌]

[1] Jean-Christophe Terrillon, Mahdad N. Shirazi, Hideo Fukamachi and Shigeru Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in proceedings Fourth IEEE

International Conference on Automatic Face and Gesture Recognition, 2000, pp 54-60.

[2] T. Sakai, M. Nagao, and S. Fujibayashi, "Line Extraction and Pattern Recognition in a Photograph," Pattern Recognition. vol. 1, pp.233-248, 1969.

[3] V. Govindaraju, "Locating human faces in photographs," International Journal of Computer Vision, vol. 19, no. 2, pp.129-146, 1996.

[4] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, pp.1042-1052, 1993.

[5] H. Yao and W. Gao, "Face locating and tracking method based on chroma transform in color images," in signal proceedings international conference on, vol. 2, pp.1367-1371, 2000.

[6] Jie Yang, Weier Lu, and Alex Waibel, "Skin-Color Modeling and Adaptation," Proceedings of ACCV'98, pp.687-694, 1998.

[7] Karin Sobottka, Ioannis Pitas, "Segmentation and Tracking of Faces in Color Images," Proc.eedings of the 2nd International Conference on Automatic Face and Gesture Recognition, pp.236~241, 1996.

[8] S. Gong, E. Ong, and S. McKenna, "Learning to associate faces across views in vector space of similarities to prototypes," In British Machine Vision Conference, vol. 1, pp54-64, 1998.

[9] J. Huang, X. Shao, and H. Wechsler, "Face pose discrimination using support vector machines," Proc. of the IEEE International Automatic Face and Gesture Recognition, pp.194-199, 1998.

[10] Jochen Heinzmann, "Real-time human face tracking and gesture recognition," Univ. of Wollongong, Master's Thesis, 1996.