

2족 보행 로봇을 이용한 얼굴 검출 및 추적 시스템

김재현, 정도준, 김향준
경북대학교 컴퓨터공학과

e-mail : { jhkim, djjung, hjkim }@ailab.knu.ac.kr

Face Detection and Tracking System using 2-legged Walking Robot

Jae Hyun Kim, Do Joon Jung, Hang Joon Kim

Dept. of Computer Engineering, kyung-pook National University

요 약

본 논문에서는 카메라가 장착된 2족 보행 로봇을 이용한 얼굴 검출 및 추적 시스템을 제안한다. 제안된 시스템은 PCA(Principal Component Analysis) 기반의 시스템으로서 얼굴을 검출하기 위해 먼저, 스킨칼라 정보와 모션 정보를 사용하고, 그 이후에 PCA 를 사용하여 스킨칼라 영역에서 실제 얼굴이 있는지를 검증 한다. 새로 검출된 얼굴과 이전에 추적되는 얼굴 사이의 동일성은 Eigenspace 상에서의 Euclidian distance 를 사용하여 검증한다. 2족 보행 로봇이 얼굴을 추적하기 위해서는, 검출된 얼굴 영역이 카메라 스크린 중심 영역에 계속 유지되도록 로봇의 움직임을 조절해 간다. 제안된 시스템은 움직임이 많고, 조명 변화나 배경의 변화가 심한 환경에서도, 얼굴을 잘 검출하고 추적 하였으며, 다른 2족 보행 시스템이나 인간과 로봇의 상호작용을 위한 제스처 인식 시스템으로의 확장도 가능하다.

1. 서론

바퀴를 이용해서 굴러가거나 몸을 이용해서 기어가는 모바일 기계들은 간단한 일들만을 할 수 있고, 움직일 수 있는 환경에 제약을 받게 된다. 움직일 수 있는 환경의 제약 때문에 ASIMO 와 SPR-4X 등과 같은 휴머노이드 2족 보행 로봇들이 만들어지고 있다. 지능적이고 상호작용이 가능한 기계들을 설계하는 목적 중에 하나는 사용자의 존재를 알아차리는 것이다. 이러한 이유로 2족 보행 로봇을 사용해서 얼굴을 검출하고 추적하는 것은 의미 있는 문제들이다. 또한 이러한 환경에서는 로봇 위에 달린 카메라의 움직임과 사용자의 움직임이 둘다 존재하기 때문에 영상이 매우 동적이므로 문제 해결이 쉽지 않다.

일반적으로, 보는 관점에 따라 추적 방법은 두 그룹으로 분류할 수 있다. 어떤 사람들은 추적 방법을 인식기반 추적과 모션기반 추적으로 분류하고 또 다른 사람들은 에지 기반 추적과 영역기반 추적으로 분류한다 [1].

인식기반 추적은 물체 인식 기술에 기반을 두며 추적 시스템의 수행 능력은 인식 방법의 효율에 의해 결정된다. 모션기반 추적은 모션검출 기술에 의존하고 있으며, 이 기술들은 옵티컬 플로우 방법과 모션 에너지 방법으로 나뉠수 있다. 에지기반 방법은 연속적인 영상에서 에지들을 추적하는 방법으로 물체의 경계선 부분에 중요성을 두고 있다. 그러나 이 방법들은 칼라 또는 조명의 변화로부터 영향을 받게 된다. 왜냐하면 추적된 물체의 경계선은 칼라나 조명으로 인해 큰 변화를 보여줄 뿐만 아니라, 백그라운드 이미지에 강한 에지들이 있을 경우에는 신뢰성 있는 결과를 제공하기가 어렵기 때문이다. 최근 연구와 관련된 대부분의 방법들은 Kass 등 여러 사람의 노력에 의해 나온 방법들이다 [2]. 이런 연구 방법들은 얼굴을 추적할 때 배경이미지에 잡음이 존재하면 얼굴이 아닌 부분을 추적하는 경향이 있었다. 예를 들면 손과 팔 등이 있다.

본 논문에서는 카메라가 장착된 2족 보행 로봇으로부터 들어오는 움직임이 많고 조명 변화나 배경의 변화가 심한 영상에서 PCA 기술을 사용하여 얼굴을 검출하고 추적하는 시스템을 제안한다. 제안된 시스템은 그림 1 에서 보듯이

얼굴 검출과 얼굴 추적의 두 단계로 구성된다. 두 연속 프레임의 차영상과 스킨칼라 정보를 사용하여 얼굴 후보들을 검출하고, 검출된 얼굴 후보가 실제 얼굴인지 PCA 를 사용하여 검증한다. 그 이후에, 검출된 얼굴이 이전에 추적되던 얼굴과 동일한지를 eigen-technique 를 사용하여 검증한다.

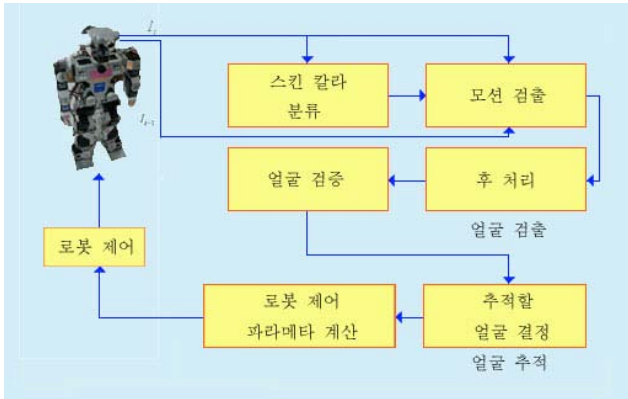


그림 1. 제안된 시스템의 구조

2. 얼굴 검출

이 장에서는 제안된 시스템에서 사용된 얼굴 검출 기술에 대해서 소개한다. 우리는 얼굴 검출의 정확성을 높이기 위해 잘 알려져 있는 스킨칼라모델[3]과 PCA[4,5]기술을 함께 사용하였다.

2.1 스킨 칼라 분류

스킨칼라로 얼굴 같은 픽셀들을 분류하는 것은 얼굴을 검출하고 추적하는데 신뢰성 있는 결과를 제공한다. 대부분의 비디오 카메라로부터 획득한 RGB 표현에서는 칼라 뿐만 아니라 밝기 값을 포함하고 있다. 이러한 칼라 공간은 스킨칼라로 얼굴 같은 픽셀들을 분류하기 위한 좋은 공간은 아니다. 하지만 밝기는 칼라 성분을 빛의 밝기에 의해 나눔으로써 제거할 수 있고, 사람 얼굴의 칼라 분포는 chromatic 칼라 영역에서 작은 영역에 군집하게 된다 [6].

얼굴 조각으로부터 획득된 스킨 칼라 픽셀들의 chromatic r 과 chromatic g 값들은 $CrCg$ 공간상에서 타원 모양으로 군집을 이룬다. 그림 2(a)는 샘플 얼굴 조각들이고, (b)는 실험에서 사용한 MPC-C30 CCD 카메라의 스킨 칼라 위치를 나타내고 있다. 스킨칼라 픽셀 덩어리의 영역은 간단한 멤버쉽 함수인 한쌍의 2 차함수로 나타낼 수 있다 [7]. 그러나 이 실험에서 스킨 칼라 픽셀 덩어리들의 형태는 타원과 비슷한 모양을 나타내었으므로, 우리는 다음과 같은 타원의 방정식으로 멤버쉽 함수를 결정하였다.

$$S = \begin{cases} 1, & \frac{(x'-c_x)^2}{a^2} + \frac{(y'-c_y)^2}{b^2} < 1, \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

여기서 $C_x = 0.382$, $C_y = 0.326$, $a = 0.063$, $b = 0.026$,

$\theta = 0.261$ (라디안)은 $CrCg$ 공간에서 스킨칼라 픽셀 덩어리로부터 계산하였다.

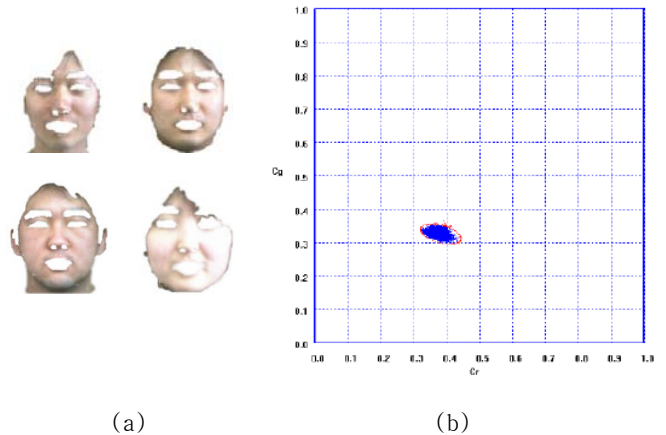


그림 2. (a) 샘플 얼굴 조각들
(b) $CrCg$ 공간에서의 MPC-C30 카메라의 스킨칼라 위치

2.2 모션 검출

스킨칼라는 얼굴 검출에 많이 사용되는 특징이지만, 스킨칼라가 사람의 얼굴 영역 뿐만 아니라 백그라운드 영역에서도 나타날 경우에는 스킨칼라 만으로 안정되게 얼굴을 검출하기가 어렵다. 이런 결점은 모션 정보를 사용함으로써 효과적으로 극복할 수가 있다. 간단히 말해 스킨칼라를 분류한 후 단지 그 영역에서만 모션이 있는지를 고려하는 방법이다. 결과 적으로 스킨칼라모델과 모션정보가 결합되어 얼굴영역과 배경영역으로 나뉘어진 2 진영상을 획득 할수 있다. 이 2 진 이미지는 다음과 같이 정의 된다.

$$M_t(x, y) = \begin{cases} 1, & I_t(x, y) \in S_t \ \& \ |I_t(x, y) - I_{t-1}(x, y)| > \theta \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

여기서 $I_t(x, y)$ 와 $I_{t-1}(x, y)$ 는 현재 프레임과 이전 프레임 픽셀들의 밝기이다. S_t 는 현재 프레임의 스킨칼라 픽셀들의 집합이다. 그리고 θ 는 적응 문턱 기술을 사용하여 계산된 문턱 값이다 [8]. 후 처리로 M_t 이미지를 모폴로지 연산과 연결 성분 분석을 사용하여 간단화 시킨다.

2.3 PCA 를 사용한 얼굴 검증

연속적인 영상에서 관심 얼굴 영역을 추적하는 것은 어려운 일이다. 왜냐하면 연속되는 영상에서는 움직이는 물체들이 많이 있기 때문이다. 게다가, 움직이는 물체가 얼굴인지 아닌지의 검증도 필요하기 때문이다. 우리는 얼굴 검증 문제를 위해 eigenspace 에서 각 영역의 웨이트 벡터를 사용한다. N 크기의 이미지는 N 차원의 이미지 공간에서 한 점으로 표현될 수 있고, 이때 각각의 차원은 한 픽셀에 대응된다. 게다가 우리는 특징공간에서 차원을 감소하기 위해 N 차원 얼굴 후보 이미지를 저차원 특징 공간(eigenspace 또는 얼굴 공간이라고도 함)으로 투영시켰다 [4,5].

Eigenspace 를 간단히 말하면 다음과 같다. M 개의 학습 이미지는 얼굴 공간을 구성하기 위해 $I_1, I_2, I_3, \dots, I_M$ 로 표현하고, 이들 입력 이미지의 평균 이미지는 $A = 1/M \sum_{i=1}^M I_i$ 로 구한다. 각 학습이미지와 평균 이미지와의 차이 벡터인

$\Phi_i = I_i - A$ 들로 이루어진 새로운 벡터 집합 $Y = [\Phi_1, \Phi_2, \dots, \Phi_M]$ 를 구한다. 얼굴 이미지들의 분포를 최적으로 묘사하기 위한 M 개의 직교 벡터를 만들기 위해서, 공분산 행렬은 다음과 같이 계산 되어진다.

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = YY^T \quad (4)$$

C 는 $N \times N$ 차수 행렬이다. $N \times N$ 차수 행렬에서 eigenvectors 와 eigenvalues 를 결정하는 것은 어렵다. 그러므로, 계산상의 편리함을 위해 C 의 eigenvectors 를 찾는 대신에 $[Y^T Y]$ 의 eigenvectors v_k 와 eigenvalues λ_k 를 계산한다. 계산되어진 v_k 와 λ_k 를 이용하여 베이스 집합 u_k 를 다음과 같이 계산한다.

$$u_k = \frac{Y \times v_k}{\sqrt{\lambda_k}} \quad (5)$$

λ_k 의 값에 따라 의미 있는 M' 개의 eigenvector 가 선택되어지고, 이를 이용해 학습 얼굴 이미지의 특징 벡터 $W_i = [w_1, w_2, \dots, w_{M'}]$ 를 다음과 같이 계산 한다.

$$w_k = u_k^T \Phi_i, \quad k = 1, \dots, M. \quad (6)$$

후보 얼굴 영역이 실제 얼굴 이미지 인지 아닌지를 검증하기 위해 후보 얼굴 영역을 학습된 eigenspace 에 투영시킨다. 투영된 영역들은 얼굴 덩어리 중심과 얼굴이 아닌 덩어리 중심과의 최소 거리를 사용하여 얼굴 인지 검증 되어진다(7).

$$\min(\| W_k^{candidate} - W_{face} \|, \| W_k^{candidate} - W_{nonface} \|), \quad (7)$$

여기서 $W_k^{candidate}$ 는 학습된 eigenspace 에서 k 번째의 후보 얼굴 영역이고, W_{face} 와 $W_{nonface}$ 들은 얼굴 덩어리와 얼굴이 아닌 곳의 중심 좌표이며, 거리 측정을 위해서는 Euclidean distance 방법이 사용되었다.

3. 얼굴 추적

새롭게 검출된 얼굴들중에서, 다음 이미지 시퀀스에서 추적할 얼굴은 eigenspace 에서 거리 측정 방법에 의해 결정된다. 얼굴을 추적하기 위해, 추적된 얼굴 특징 벡터들과 새롭게 검출된 얼굴 k 의 특징 벡터 사이의 Euclidean distance 는 다음과 같이 계산된다.

$$obj = \arg \min_k \| W_{old} - W_k \|, \quad k = 1, \dots, K. \quad (8)$$

얼굴 영역이 결정된 후에 검출된 얼굴 영역의 중심과 카메라 스크린 중심과의 거리는 다음과 같이 계산된다.

$$dist_t(face, screen) = Face_t(x, y) - Screen\left(\frac{height}{2}, \frac{width}{2}\right), \quad (9)$$

여기서 $Face_t(x, y)$ 는 t 시간에 검출된 얼굴 영역의 중심이고 $Screen(height/2, width/2)$ 는 $screen$ 의 중심이다. 얼굴 크기와 얼굴이 있는 방향 그리고 거리 벡터를 사용하여 카메라

의 상하, 좌우 움직임을 제어한다. 카메라의 제어는 로봇의 움직임을 제어하여, 검출된 얼굴 영역이 화면의 중심 영역을 벗어나지 않도록 하는 것이다. 우리는 미리 정의된 로봇의 액션 “앞으로 걷기”, “뒤로 걷기”, “오른쪽으로 움직이기”, “왼쪽으로 움직이기”, “오른쪽으로 방향 틀기”, “왼쪽으로 방향틀기” 를 사용하여 로봇의 움직임을 제어하였다.

4. 실험 결과

실험은 잡음과 조명 상태가 일정하지 않은 연구실 환경에서 하였다. 그림 3 은 제안된 시스템의 셋업과 인터페이스 구조를 보여준다.

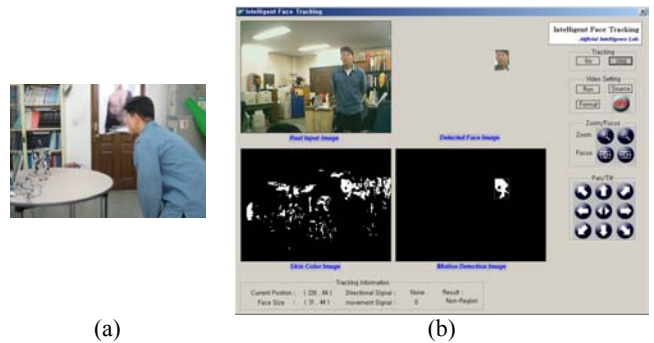


그림 3. 제안된 시스템의 구성 (a) 셋업 (b) 인터페이스

시스템은 얼굴 추적을 위해 로봇을 제어 한다. 제어된 결과들은 로봇의 액션으로 나타나게 된다. 사용된 로봇 KHR-1 은 장착된 서보 컨트롤 보드인 RCB-1 에 의해 다양한 각도로 움직일 수 있다. RCB-1 은 그림 4(b)에서 보여진 HeartToHeart1.0 소프트웨어에 의해 조정 되어진다. 우리는 로봇을 조정하기 위해 HeartToHeart1.0 을 사용하였다. 먼저, 우리는 HeartToHeart1.0 의 모션 생성기를 사용하여 로봇의 액션을 만들고, 그 이후에 모션 제어를 사용하여 각각의 액션에 번호를 할당하고 액션을 RCB-1 에 저장 하였다. 이렇게 하여, 로봇 제어기는 액션 번호가 주어지면 로봇을 제어 하게 된다.

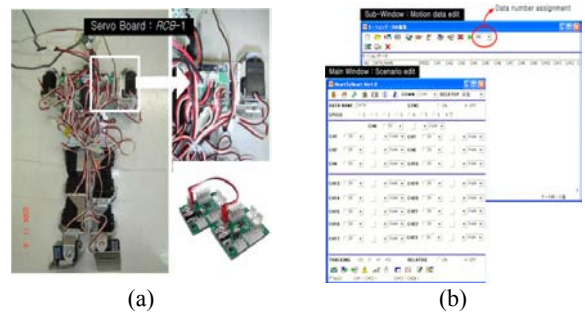


그림 4. 로봇 제어기: (a) 컨트롤 보드: RCB-1 (b) 사용자 그래픽 인터페이스(GUI): HeartToHeart1.0

이 실험에서, 사용자는 복잡한 배경이 있는 환경에서 로봇 앞에 서있게 되고, KHR-1 의 머리 부분에 장착된 카메라로부터 이미지를 얻게 된다.

성능 평가를 위해, 제안된 시스템은 7 명의 서로 다른 사용자에게 의해 테스트 되었고, 학습 이미지는 15 명의 각각 다른 각도에서의 얼굴이미지 16 장씩과, 팔과 손등 얼굴이 아닌 이미지 50 장이 사용되었다. 그림 5 에서는 eigenspace

구성에 사용된 얼굴 학습 이미지 중에서 일부를 보였다.



그림 5. eigenspace 구성에 사용된 학습 이미지의 샘플

실험을 분석해본 결과, 평균 얼굴 검증률이 83.1%로 나타났다. 표 1 에서는 얼굴을 얼굴로 검증한 비율과, 얼굴이 아닌것을 얼굴이 아닌것으로 검증한 비율을 나타내고 있다.

$$\text{Face verification rate} = \frac{\text{Number of correctly verified faces}}{\text{Number of images verified as true face}} \quad (10)$$

$$\text{Non-face verification rate} = \frac{\text{Number of correctly verified non_faces}}{\text{Number of images verified as non_face}} \quad (11)$$

표 1. 얼굴 검증률

사용자들	얼굴	얼굴 아님	합계
사용자 1	81.9 %	78.3 %	80.1 %
사용자 2	85.3 %	77.5 %	81.4 %
사용자 3	86.2 %	83.2 %	84.7 %
사용자 4	83.6 %	82.2 %	82.9 %
사용자 5	86.3 %	83.6 %	84.9 %
사용자 6	85.1 %	81.5 %	83.3 %
사용자 7	84.3 %	84.6 %	84.4 %
합계	84.6 %	81.5 %	83.1 %

그림 6 은 제안된 시스템의 결과를 보여준다. (a)는 얼굴 추적의 결과들이고, (b)는 얼굴이 검출된 시점들이다. 맨 윗줄은 입력 이미지들이고 중간 줄은 입력 이미지들의 스킨 칼라 영역들이고, 맨 아랫줄은 추적되는 얼굴 영역이다.

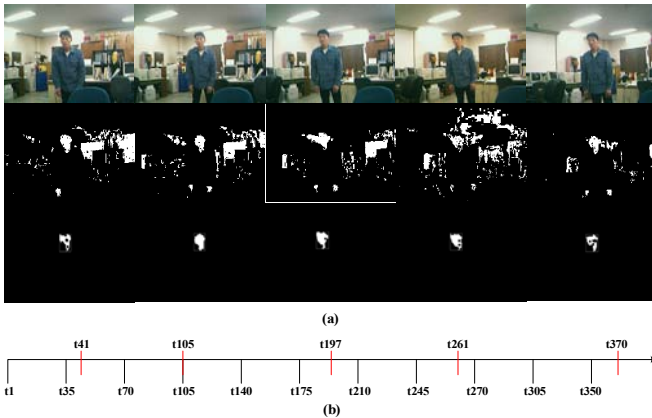


그림 6. 제안된 시스템의 결과 (a) 얼굴 추적 결과 (b) 얼굴 검출 시점들

5. 결론

본 논문에서는 카메라가 장착된 2 축 보행 로봇을 이용한 얼굴 검출 및 추적 시스템을 제안하였다. 제안된 시스템은 실시간으로 얼굴 검출과 얼굴 추적이라는 주요 두 단계를 수행하도록 구성되었다. 입력 비디오 영상에 대해, 칼라와 모션 정보 그리고 PCA 를 사용하여 얼굴 영역을 검출하였다. 검출된 얼굴 영역이 카메라의 중심 영역을 벗어나지 않도록 2 축 보행 로봇을 조정하였다. 제안된 시스템은 잡음이 있는 환경에서도 안정성이 있음을 실험 결과를 통해 알 수 있었다. 제안된 시스템은 인간과 로봇의 상호 작용을 위한 제스처 인식 시스템이나, 다른 2 축 보행 로봇 시스템 등에도 확장이 가능하다. 향후 연구 과제로는 다른 큰 2 축 보행 로봇 환경에서 제안된 시스템을 적용하고 좀더 많은 실험을 해보는 것이다.

참고문헌

[1]Shearer, K., Wong, K.D., Venkatesh, S.: Combining multiple tracking algorithms for improved general performance. Pattern Recognition, Vol. 34 (2001) 1257-1269

[2]Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Computer Vision, Vol. 1 (1996) 321-331

[3]Hsu, R.L., Abdel-Mottaleb, M., Jain, A.K.: Face detection in color images. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 (2002) 696-706

[4]Turk, M., Pentland, A.: Face Recognition Using Eigenfaces. IEEE Proceedings CVPR '91, (1991) 586 -591

[5]Guan, A.X., Szu, H.H.: A local face statistics recognition methodology beyond ICA and/or PCA. International Joint Conference on Neural Network, (1999) 1016 -1021

[6]Yang, J., Waibel, A.: A Real-Time Face Tracker. IEEE Workshop on Application of Computer Vision, (1996) 142-147

[7]Soriano, M., Martinkauppi, B., Huovinen, S., Laaksonen, M.: Adaptive skin color modeling using the skin locus for selecting training pixels. Pattern Recognition, Vol. 36 (2003) 681-690

[8]Habibi, N., Moini, A., Burgess, N.: Automatic Thresholding for Change Detection in Digital Video. SPIE Proceedings on Visual Communications and image Processing, Vol. 4067 (2000) 133-142