

# 조명 변화에 강건한 피부색 검출을 사용한 게임 인터페이스

장상수<sup>o</sup>, 박혜선, 김은이<sup>l</sup>, 김향준  
경북대학교 컴퓨터공학과 인공지능연구실  
{ ssjang<sup>o</sup>, hspark, hjkim }@ailab.knu.ac.kr  
eykim@konkuk.ac.kr<sup>l</sup>

## Game Interface using Robust Skin Color Detection

Sang Su Jang, Hye Sun Park, Eun Yi Kim<sup>l</sup> and Hang Joon Kim

Dept. of Computer Engineering, Kyungpook National Univ., Daegu, Korea  
Dept. of Internet and Multimedia Engineering, Konkuk Univ., Seoul, Korea<sup>l</sup>

### 요 약

최근 사용자의 제스처를 이용한 게임 시스템에 대한 연구가 많은 관심을 받고 있다. 사용자의 얼굴 및 손의 움직임을 이용하여 게임을 제어하기 위해서는 복잡한 배경 및 조명에 강건한 얼굴 및 손 영역의 추출이 필수적이다. 본 논문에서는 조명 변화에 강건한 피부색 검출을 이용한 게임 인터페이스를 제안한다. 이를 위해 제안된 시스템은 다음의 두 단계로부터 얼굴 및 손 영역을 추출한다. 먼저, 피부색과 유사한 물건들을 제거하기 위해 배경 영상과 현재 영상의 차영상으로부터 전경 물체를 추출한다. 그 다음, 조명에 의한 깜박임이나 잡음을 줄이기 위해서 SCT 알고리즘을 이용하여 전경 물체 영역 안에서 피부색 영역만을 정확하게 검출한다. 추출된 얼굴 및 손의 움직임으로부터 얻어지는 제스처는 은닉마르코프 모델을 사용하여 인식된다. 복잡한 환경에서 실험한 결과, 제안된 시스템은 정확한 피부색 영역 검출을 제공하고 이를 통한 보다 정확한 인식률을 제공할 수 있다는 것이 증명되었다.

### 1. 서론

최근 사용자의 얼굴, 또는 손 등의 제스처를 이용하여 컴퓨터 게임 및 로봇을 제어할 수 있는 인터페이스의 개발에 대한 연구가 활발히 진행되고 있다. Freeman과 Weissman[1]은 손을 추적하여 TV를 제어하는 시스템을 제안하였고, Yang et al.[2]는 제스처를 사용하여, 원격으로 로봇을 제어하였다.

이러한 시스템들은 사용자의 얼굴과 양손의 위치 좌표를 사용하여 시스템을 제어하기 때문에, 사용자의 얼굴 및 손을 정확히 추출하는 작업이 중요하다. 하지만 실제 환경에서의 복잡한 배경 때문에 사용자의 얼굴과 양손을 정확하게 검출하는 것은 매우 어렵다. 예를 들어 배경 영상에는 사용자의 피부색 영역과 유사한 색깔 값을 가진 물건이 많이 존재할 뿐만 아니라, 조명에 의한 깜박임, 또는 잡음이 포함되어 있다.

본 논문에서는 이러한 문제점을 해결할 수 있는 제스처 기반의 게임 인터페이스를 제안한다. 강건한 얼굴 영역의 추출을 위해, 제안된 시스템은 얼굴 및 손 영역을 추출한다. 먼저 피부색과 유사한 물건들을 제거하기 위해 배경 영상과 현재 영상의 차영상으로부터 전경 물체를 추출한다. 조명에 의한 잡음을 줄이기 위해서, SCT 알고리즘을 이용하여 전경 물체 영역 안에서 피부색 영역만을 정확하게 검출한다. 추출된 얼굴 및 양손의 움직임으로부터 얻어지는 제스처는 은닉 마르코프 모델을 사용하여 인식된다. 복잡한 환경에서 실험한 결과, 제안된 시스템은 정확한 피부색 영역 검출을 제공하고 이를 통한 보다 정확한 인식률을 제공할 수 있다는 것이 증명되었다.

본 논문은 다음과 같이 구성된다. 제 2 장에서는 제안된 인터페이스에 대해 간략하게 소개하고, 3 장에서는 강건한 피부색 검출 방법을 자세히 설명한다. 그리고, 4 장은 실험 결과를 보여주고, 마지막으로 5 장에서 결론을 맺는다.

### 2. 제안된 게임 인터페이스

본 논문에서 소개하고 있는 게임 인터페이스는 입력 장치인, 카메라와 영상을 처리하는, 프로세스 그리고 출력장치인, 프로젝터로 구성되어 있다. 인터페이스의 핵심 부분인, 프로세스는 특징 추출, 포즈 분류, 제스처 인식, 제스처 모핑, 게임 제어의 모듈로 구성된다. 게임 인터페이스 구성도는 그림 1과 같다.

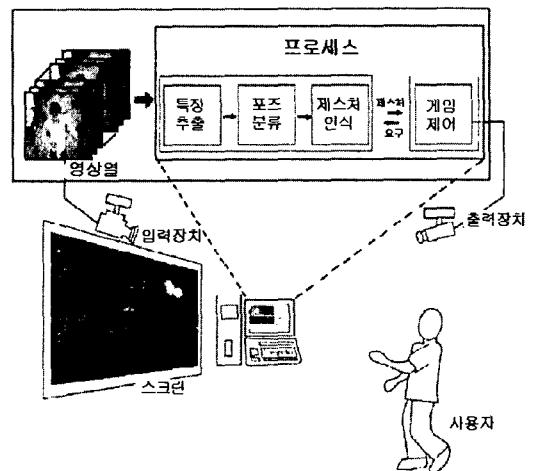


그림. 1 인터페이스 구성도

2.1 특징 추출 모듈

사용자의 제스처를 인식하기 위해서 특징을 추출하는 단계이다. 현재 입력 영상에서 배경을 제거하고, 피부색 영역을 검출하여, 잡음을 제거하고, 머리와 양손의 2 차원 좌표 값을 검출한다. 각 모듈의 동작은 다음 단원에서 자세히 소개된다.

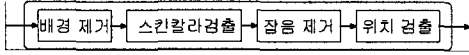


그림 2 특징 추출 모듈 단계도

2.2 포즈 분류 모듈

포즈 분류모듈은 현재의 프레임이 23 개의 포즈 중에서 어떤 포즈인가를 결정하는 단계이다. 각 프레임에서 추출한 특징 벡터들과 미리 정의한 포즈 심볼 테이블을 유클리드 거리값으로 비교하여 차이가 가장 근소한 포즈로 분류한다.

2.3 제스처 인식 모듈

연속된 포즈 심볼로부터 제스처를 인식하는 단계이다. 제스처를 인식하기 위해 새로운 구조의 은닉 마르코프 모델 (hidden Markov Model: HMM) 을 이용한다. 사용된 HMM 은 13 개의 제스처를 하나의 HMM 에 모두 정의되어 있다. 제안된 HMM 은 연속된 포즈 심볼을 사용하며, 각 제스처와 관련된 경로를 구별하기 위한 몇 개의 특징을 가지는 특정 포즈를 가진다. 자세한 구조는 [3] 에서 소개된다.

2.4 게임 제어 모듈

제스처 인식기에서 판단한 제스처 명령어를 Quake II 게임의 명령어로 변환하기 위한 단계이다. Quake II 게임 내의 입력 처리 모듈인 gamex86.dll 을 수정하여 사용하였다.

3. 얼굴 및 손 영역 검출

먼저 피부색과 유사한 물체들을 제거하기 위해, [4] 에서 제안한 방법을 사용한다. 사용된 방법은 배경 영상과 현재 영상의 차이 영상으로부터 전경물체를 추출한다. 추출 방법은 식 (1), (2) 와 같다.

$$FG_i(x, y) = \begin{cases} 1 & |BG_i(x, y) - F_i(x, y)| \geq \theta \\ 0 & otherwise \end{cases} \quad (1)$$

식 (1)에서,  $BG_i$  와  $F_i$  는 배경 모델의 명도와 현재 프레임의 명도를 나타낸다.  $FG_i$  는 전경물체로만 구성된 이진 영상이다. 임계값  $\theta$  는 20 이다. 초기 배경 영역  $BG_0$  과 현재 배경 영역  $BG_i$  을 사용하여 배경 모델을 업데이트한다.

$$BG_{i+1} = (1 - w)BG_0 + wBG_i, \quad (2)$$

식 (2)에서,  $BG_{i+1}$  은 업데이트 된 배경 모델이며, 가중치 요소  $w$  는 0.2 이다.

그 다음 조명에 의한 깜박임이나 잡음을 줄이기 위해, [5] 에서 제안한 SCT 알고리즘을 이용하여 전경 물체 영역 안에서 피부색 영역만을 정확하게 검출한다. SCT 알고리즘은 식 (3)-(5) 에 따른다.

$$L = \sqrt{R^2 + G^2 + B^2}, \quad (3)$$

$$\angle A = \cos^{-1} \left[ \frac{B}{L} \right], \quad (4)$$

$$\angle B = \cos^{-1} \left[ \frac{R}{L \sin(\angle A)} \right], \quad (5)$$

$L$  은 명도 값이다. 만약 어떤 색깔이 RGB 공간에서의 임의의 한 점이 있다면, 그림 3의 (a) 와 같이 SCT 를 구성하는 사면체 내부에 존재하게 된다. 임의의 점의 벡터 크기는  $L$  이고, 벡터  $L$  과 파란색 축의 차이각도는  $\angle A$  이며, RG 평면에서 벡터  $L$  의 투영선과 빨간색 축 사이의 각도는  $\angle B$  이다. 변환에 의해 결정된 색깔 공간의 결과는 (b)처럼 삼각형으로 표현할 수 있다. 이때, (c)와 같이 색깔은 벡터가 지나가는 삼각형 내에 점으로 정의 될 수 있다.

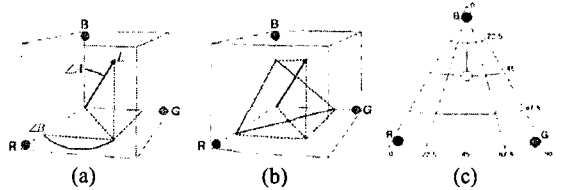


그림 3. SCT 모델의 구성도

SCT에서의 살색 영역을 찾기 위해 그림 4와 같은 100장의 살색 영역 데이터를 사용하였고, 사용된 데이터는 조명에 따른 피부색의 변화와 사람마다의 피부색 차이점을 고려하여 수집하였다.  $L, \angle A, \angle B$  값은 각각 170~330, 1.001~1.18, 0.82~0.97로 나타났다.



그림 4. 사용된 피부색 이미지

$$\angle A, \angle B, L = \begin{cases} 1 & \angle A \cap \angle B \cap L \text{ skin-color region} \\ 0 & otherwise \text{ non skin-color region} \end{cases}, \quad (6)$$

스킨 칼라모델은 식 (6)에 따라 생성된다. 실험을 통해 추출한 피부색 영역들의  $\angle A$  와  $\angle B$ , 그리고  $L$  의 값이 모두 포함하면 피부색 영역으로 판단하고, 하나라도 포함하지 않으면 피부색 영역이 아닌 것으로 판단한다.

위의 방법으로 검출된 얼굴 및 손 영역으로부터 모폴로지 연산을 통해, 검출된 영역 안의 작은 구멍들을 채우고, 연결 성분 분석 (connected components analysis)을 사용하여, 검출된 영역을 라벨링함으로써, 사용자의 머리와 양손의 2D 위치 좌표값을 검출한다. 이때, 라벨링은 다음과 같은 5 가지의 휴리스틱한 방법들을 사용한다.

- 라벨 영역의 크기가  $m$  보다 커야 선택 한다.
- 가장 큰 라벨은 머리로 추정한다.
- 두 번째 세 번째 큰 것은 손으로 추정한다.
- 손은 항상 교차 되지 않는다는 것을 가정한다.
- 영역의 무게 중심은 영역의 위치를 사용된다.

4. 실험 결과

제안된 인터페이스는 사용자의 얼굴 및 손의 움직임을 이용하여, Quake II 게임을 제한한다. 실험 환경은, 입력장치로써 320\*240의 RGB 영상을 입력받아 처리할 수 있는 디지털 비디오 카메라를 사용하였고, 사용된 카메라는 게임 사용자와 1.75m 떨어진 곳에서 2.55m 위에 55°로 내려보도록 설치되었다. 영상처리 장치는, 2.7GHz PC에서 Visual C++ 6.0를 사용하였다. 제안된 인터페이스는 1초당 10-11프레임을 처리할 수 있었고, 게임을 실시간으로 진행할 수 있었다.

제안된 인터페이스가 복잡한 배경 및 조명에도 강건한지 여부를 증명하기 위해서, 일반적으로 피부색 검출에 사용되는 CrCg 방법과 본 논문에서 소개한 SCT 방법을 비교해 보았다. 그림 4는 CrCg와 SCT의 피부색 영역의 크기를 비교한 결과 그림이다.

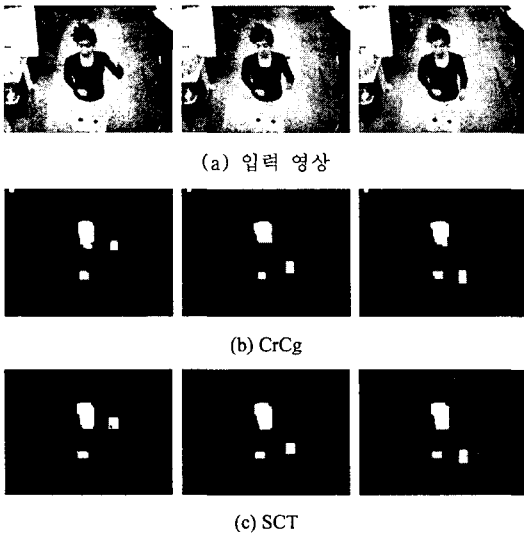


그림 4. CrCg와 SCT의 피부색 영역

실험 결과, CrCg는 그림 5의 (b)와 같이, 조명에 의한 깜박임이나 잡음에 민감하게 반응하기 때문에 매 입력 프레임마다 피부색 영역의 크기가 계속 변동하였다. 그러나, 그림 4의 (c)에서 보면 알 수 있듯이, 제안된 논문에서 사용된 SCT 알고리즘을 이용한 결과, 매 입력 프레임마다 거의 일정한 크기의 피부색 영역을 검출할 수 있었다.

표 1에서는 CrCg와 SCT의 피부색 영역 검출시 크기 변화량은 비교한 것이다.

표 1. 피부색 영역 크기의 평균 변화량

CrCg	SCT
68.251	29.573

또 CrCg의 경우는 그림 5의 (b)에서 볼 수 있듯이, 조명 상태에 따라 검출코져 하는 피부색 영역이 갈라졌다. 그러나 SCT를 이용해 본 결과, 그림 5의 (c)와 같이 피부색 영역을 잘 검출할 수 있었다.

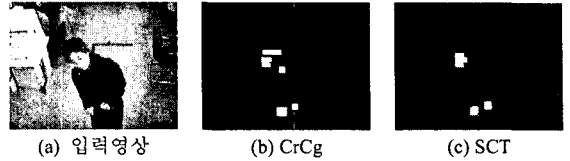


그림 5. 조명에 따라 달라지는 피부색 영역

결과적으로 그림 6에서 보듯이, 제안된 게임 인터페이스가 조명 변화에 강건한 피부색 검출을 함을 증명할 수 있었다.



그림 6. 피부색 검출 과정

그림 6 (a)는 입력 영상이고 (b)는 (a)로부터 전경물체를 배경으로부터 검출하지 않고 단지, CrCg로 피부색을 검출한 결과이다. (c)는 전경물체를 배경을 제거하여 전경물체 검출 방법을 사용한 후, CrCg로 피부색을 검출한 결과이다. 마지막으로 (d)는 전경물체 검출한 뒤, SCT 알고리즘을 적용한 결과이다.

실험은 5명이 13개의 제스처를 각각 10번씩 수행하여 얻은, 약 15,000개의 이미지 데이터를 사용하였다. 표 2는 피부색에 의한 오류의 수를 나타낸다. CrCg보다는 SCT가 오류가 적음을 알 수 있었다.

표 2. CrCg와 SCT 오류수

CrCg	SCT
381	275

5. 결론

본 논문에서 강건한 피부색 검출방법을 이용한 비전 기반의 인터페이스가 제안되었다. 제안된 시스템은 복잡한 배경 및 조명하에서도 사용자의 움직임을 정확하게 검출할 수 있으며, 정확한 인식률을 제공하였다.

참고 문헌

- [1] W.T. Freeman, and C.D. Weissman, "Television control by hand gestures", In Proceedings of International Workshop of Face and Gesture Recognition, pp. 179-183, 1995
- [2] S. Iba, JM.V. Weghe, C.J.J. Paredis,P.K. Khosla, "An architecture for gesture based control of mobile robots", Intelligent Robots and Systems, Vol. 2, pp. 850-857, 1999.
- [3] H.S. Park, E.Y. Kim, S.S. Jang and H.J. Kim, "An HMM Based Gesture Recognition for Perceptual User Interface", Pacific-Rim Conference on Multimedia(PCMM2004),LNCS, to be published,2004, 12.
- [4] F.S. Chen, C.M Fu, C.L. Huang "Hand gesture recognition using a real-time tracking method and hidden Markov models" Image and Vision Computing, Volume 21, Issue 8, 1 August 2003, Pages 745-758
- [5] Hyams, J.; Powell, M.W.; Murphy, R. "Cooperative navigation of micro-rovers using color segmentation," Computational Intelligence in Robotics and Automation, IEEE International Symposium on , 8-9 Nov. 1999, pp. 195 - 201