

## 피부색 정보 및 역운동학을 이용한 인체 움직임 추정

박정주\* 김주혜<sup>0</sup>,\* 김명희\*,\*\*

\*이화여자대학교 컴퓨터학과

\*\*이화여자대학교 컴퓨터 그래픽스/가상현실 연구센터  
{ ddun21, angela98<sup>0</sup> }@ewhain.net, mhkim@ewha.ac.kr

### Human Body Postures Estimation Using skin color information and Inverse Kinematics

JungJu Park\* JuHaye Kim<sup>0</sup>,\* Myoung-Hee Kim\*,\*\*

\*Dept. of Computer Science & Engineering Ewha Womans University

\*\*Center for Computer Graphics & Virtual Reality (CCGVR), Ewha Womans University

#### 요 약

본 논문에서는 CAVE<sup>TM</sup>-like 시스템에서 역운동학을 이용하여 사용자의 인체 움직임을 실시간으로 추정하기 위한 기법을 제안한다. 사용자를 둘러싼 스크린으로 투사되는 빛에 인하여 매순간 변화하는 배경을 포함하는 영상으로부터 사용자 영역을 추출하기 위하여 적외선 반사 영상을 이용하였다. 이를 이용하여 추출된 사용자 컬러 영상에 프로젝션 기반 가상환경의 특성을 고려한 컬러 모델을 적용하여 사용자의 얼굴과 손 영역을 추출하였다. 위치 추정 단계를 통하여 다음 프레임에서의 관심영역 위치를 미리 예측하고 추출된 사용자의 손과 얼굴 위치 정보를 말단장치로 이용하고, 관절모델과 운동학적 제약조건을 기반으로 역운동학적인 방법을 통해 사용자의 동작을 추정하였다. 제안 기법에서는 별도의 마커나 도구를 사용하지 않기 때문에 비침입적이며 사용자 움직임의 제약을 최소화할 수 있기 때문에 보다 인간 중심적인 인터랙션을 위한 기반 기술로 이용될 수 있다.

#### 1. 서 론

가상현실(virtual reality)이라는 개념이 처음 등장하고 HMD(head mounted display)를 이용하여 이러한 개념이 실제로 실현된 이 후부터 지금까지 다양한 형태의 가상현실 시스템 및 관련 기술이 꾸준히 개발되고 있다[1]. 또한, 컴퓨터 그래픽스 기술 발전, 프로세싱 성능 향상 및 가격 하락, 다양한 센싱 장비가 개발됨에 따라 가상환경이 널리 보급되고 있으며, 여러 대학이나 연구소에서는 가상환경 기술을 군사, 의료, 건축, 교육, 훈련, 우주, 항공 산업, 엔터테인먼트 등과 같은 응용 분야에 접목시켜 다양한 어플리케이션을 개발하고 있다. 가상현실이 일반화 되어 대중적으로 활용되기 위해서는 가상객체와 사용자 간의 쉽고 직관적인 상호작용이 필수적이다. 특히, 가상현실의 주요 요소인 몰입감 측면에서 볼 때, 자연스럽게 사용자 중심적인 인터랙션의 개발이 무엇보다 중요하다고 할 수 있다. 이를 위하여 어플리케이션에 따라 적합한 인터랙션 방법들이 다양하게 개발되고 있지만, 대부분 완드나 3차원 마우스와 같은 특정 디바이스를 이용하고

있다. 이러한 디바이스들은 처음 사용하거나 익숙하지 못한 사람들은 이용하는데 불편함을 느끼게 된다. 또한, 가상현실 기술에 익숙한 사람이라 할지라도 어플리케이션을 개발할 때마다 그 환경에 맞는 인터랙션 방법을 익혀야 하는 불편함이 있다. 자연스러운 인터랙션을 위해서는 특정 디바이스를 사용하지 않고, 손이나 사용자의 신체 움직임 등을 이용하여 일상생활에서 하는 동작과 유사한 방식으로 상호작용 할 수 있어야 한다. 가상공간에서의 사용자 인터랙션은 입력장치의 위치나 움직임을 측정하는 트래킹을 통해 얻어진 정보를 기반으로 한다. 흔히 사용되는 마그네틱 방식이나 초음파 방식의 트래킹은 케이블로 연결된 디바이스를 이용하여 인터랙션하거나 사용자의 고정된 자세만을 인식할 수 있다. 사용자의 팔이나 다리의 움직임을 트래킹하기 위해서는 사용자의 몸에 장치를 부착하거나, 무거운 장비 혹은 슈트를 입어야 한다. 이러한 방법들은 케이블에 의해 사용자의 행동과 움직일 수 있는 범위에 제약을 줄 뿐만 아니라 사용자로 하여금 부자연스러움과 불편함을 느끼게 하는 문제점이 있다. 비전 트래킹 방식은 카메라를 사용함으로써 사용자의 이동성 문제를 해결하고 사용자의 머리와 손이나 손가락 위치를 추적하여 간단한 제스처나 포인팅과 같이 자연스러운

\* 본 연구는 정보통신부 대학정보통신연구센터(ITRC) 육성지원사업 지원에 의해 수행되었음.

인터랙션이 가능하게 해준다. 하지만 카메라를 이용한 방법은 케이블로 연결되어 있지 않아 사용자의 이동성은 보장하지만 여전히 활성센서[2]나 비활성센서[3,4,5]를 몸이나 손가락에 부착하거나 마커가 부착되어 있는 슈트를 입어야 하는 불편함이 있다는 문제점이 있다.

본 연구에서는 몰입형 가상현실 장비인 CAVE™-like 시스템 안에서 사용자의 피부색 컬러 정보를 기반으로 마커와 케이블 없이 사용자의 움직임을 실시간으로 추적하고 단순화 시킨 인체 모델과 역운동학적 방법을 기반으로 사용자의 관절 움직임을 측정하고, 제한적이지만 사용자의 동작을 찾아 낼 수 있다.

## 2. 사용자 트래킹

본 연구에서는 몰입형 가상환경인 CAVE™-like 시스템과 같이 동적배경인 환경에서 사용자의 피부색 영역을 추출하고 추출 영역에 대한 라벨링을 통해 각 영역을 트래킹하였다.

### 2.1 컬러모델을 이용한 피부영역 추출

몰입형 가상환경에서 일반적인 어플리케이션에 의해 제공되는 밝기는 보통 머신 비전에서 사용하기에는 낮은 광도이다. 이를 보완해주기 위해 상단에 흰색 조명을 설치하였다. 사용자 피부색 영역을 추출하고 트래킹하기 위해서는 동적배경으로부터 사용자의 컬러정보를 획득할 수 있어야 한다. 동적배경(그림 1)으로부터 사용자의 영역을 분리하기 위해서, 적외선 반사 영역을 획득하고 배경 제거 알고리즘을 통해 사용자 영상을 빠르고 정확하게 획득하였다[6]. 사용자 컬러정보를 획득하기 위해 추출된 적외선 반사영상과 쌍을 이루는 컬러 카메라 영상을 매핑시켜 동적배경이 제거된 사용자 영역(그림 2)을 추출하였다[7].



그림 1 동적배경 제거 전



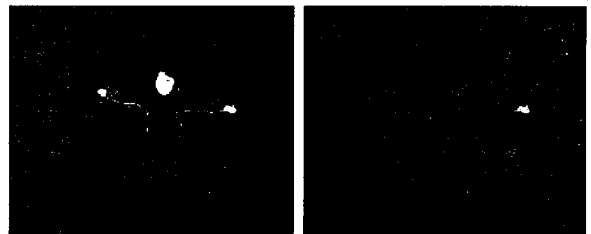
그림 2 동적배경 제거 후

사용자의 피부색 영역인 양손과 얼굴을 추출하기 위해 적합한 컬러모델로 HSV컬러모델을 사용하였다. HSV 컬러모델은 색조(Hue), 채도(Saturation), 명암(Value)의 세 가지 형태로 구성되어 있으며 다른 컬러모델보다 밝기에 대한 증속성이 작고, 지속적으로 검출이 가능하기 때문에 트래킹하기 위해 필요한 예측정보를 효율적으로 얻을 수 있다. 선택된 HSV 컬러모델에서 단순히 색상만으로, 혹은, 채도만으로 피부색 영역을 정확하게 추출해 낼 수 없기 때문에 두개의 값을 혼합한 경계값을 통해 추출하였다. 검출된 영역은 피부색과 비슷한 픽셀들로 추출하고자 하였다. 검출된 영역은 피부색과 비슷한 하는 관심영역의 수보다 더 많이 추출되거나, 추출 실패 등의 노이즈가 발

생하게 된다. 이를 방지하고 정확하게 관심영역을 추출하기 위해 추가적인 이미지 프로세싱을 최대각도와 최소각도를 정의해줌으로써 실제 사람이 표현할 수 없는 동작이나 움직일 수 없는 관절각도를 제거 하도록 한다.

### 2.2 영역 라벨링 및 트래킹

각 영역을 구성하는 개개의 픽셀들을 하나의 영역으로 묶어내고, 인접되어 있고 연결되어 있는 모든 픽셀에는 동일한 라벨을 부여하고 다른 연결 성분은 서로 다른 라벨을 부여하도록 한다. 추출된 사용자의 얼굴, 양손 영역은 이차화된 블럽(Blob)형태(그림 3(a))로 나타내어진다. 영역을 쉽게 구분해주기 위해 연결 성분 분석을 통해 추출된 영역을 라벨링(그림 3(b))하였다. 추출된 영역은 매 프레임마다 적당한 블럽을 찾아내는 단계와 다음 프레임에서 블럽이 나타날 위치를 예측하는 두 단계로 이루어진 멀티블럽 트래킹 방법을 통해 추적한다. 이를 통해 추출된 양손과 얼굴영역을 시간의 흐름에 따라 놓치지 않고 지속적으로 추적할 수 있도록 하였다.



(a) 이차화 영상 (b) 라벨링된 영상  
그림 3 피부색 영역 추출

## 3. 역운동학을 통한 사용자 움직임 추정

본 연구에서는 트래킹을 통해 얻어진 사용자의 손과 얼굴영역 위치정보를 말단장치로 규정하고, 중간관절은 역운동학적으로 통하여 관절각을 계산하였다. 이를 통해 사용자 전신 움직임을 추정하였다.

### 3.1 인체모델 단순화

사람의 인체는 206개의 뼈, 89개 관절, 189의 자유도를 가지고 있다. 하지만 이러한 정보를 모두 나타내려면 굉장히 복잡한 연산을 필요로 한다. 이는 연산 속도를 낮추는 원인이 되어 인체 움직임을 빠르게 파악하기 힘들다. 대부분의 어플리케이션은 사용자의 모든 관절 집합값을 필요로 하지 않기 때문에 특정 분할 수준에 맞춰 인체 모형을 생성하였다. 본 연구에서 사용하는 영상은 높이 2.4m인 CAVE™-like 시스템 상단에 부착되어 있는 카메라에서 얻는 것으로, 시스템 공간을 모두 포함하고, 사용자 전신을 캡처하기 때문에 손목이나 발과 같은 부분의 작은 움직임을 정확하고 자세하게 찾을 수 없다. 또한, 본 연구에서는 손, 발 움직임과 같은 미세한 움직임까지 포착하는 것에 중점을 두지 않고 사용자의 전체적인 움직임을 추정해내는 것으로 손목은 하완 부분으로, 발은 다리 부분으로 취급하도록 하였다. 사람의 움직임을 파악하는 프로세스는 시간이 많이 걸리는 부분으로 허용되지 않는 관절 움직임을 빠르게 찾아내고 보완해주기 위해서는

관절각도의 운동학적 제한조건(kinematic constraint) 설정이 필요하다. 이는 관절의 최대각도와 최소각도를 정의 해줌으로써 실제 사람이 표현할 수 없는 동작이나 움직일 수 없는 관절각도를 제거하고 유효한 값을 계산해 낼 수 있다. 본 연구에서는 10개 관절, 16개 분절, 16 자유도를 갖는 관절모델(articulated model)로 인체를 단순화 시켰다.

3.2 주요 관절각 계산

앞에서 멀티플렉스 트래킹을 통해 얻어진 양손과 얼굴의 위치 값을 말단장치(end-effector)로 한다. 역운동학 방법 [8]을 통해 팔꿈치와 무릎 위치를 계산하고 각 관절이 갖는 모든 자유도 대해 회전각을 구한다. 각 관절 제한값을 적용함으로써 찾는 공간을 줄여준다. 역운동학을 통해 얻어진 인체모델을 실루엣(그림 4)과의 매칭을 통해 모델을 보정해주도록 한다.

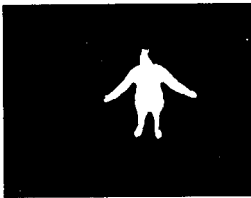


그림 4 실루엣 영상

4. 실험결과

그림 5는 CAVE™-like 시스템 내부에서 설치된 흑백 카메라와 컬러 카메라를 이용하여 동작배경을 제거하고, 동작배경이 제거된 영상에서 HSV 컬러모델을 사용하여 관심영역을 추출 및 추적하여 역운동학을 통한 사용자의 움직임을 추정한 결과영상이다. 결과영상을 통해 관심영역인 사용자의 양손과 얼굴의 위치를 추적하고 움직임에 따른 관절의 위치와 동작을 추정할 수 있음을 알 수 있다. 실험환경은 Intel Pentium-IV CPU 3.0GHz, 1GB RAM) 이며, 영상의 해상도는 320X240이다.

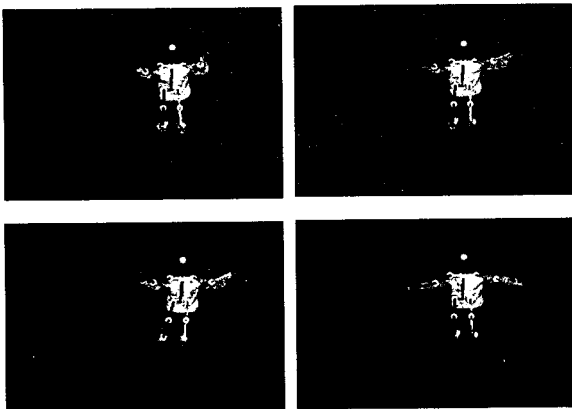


그림 5 사용자 움직임 추정 결과

5. 결론 및 향후 연구

본 연구에서는 몰입형 가상환경에서 사용되고 있는 인터랙션 방법이 가지고 있는 문제점을 해결하고, 사용자에 대한 동작정보를 획득할 수 있음으로써 보다 자연스러운 인터페이스를 위한 정보를 제공해줄 수 있도록 하였다. 이를 위해, 동적으로 변하는 환경에서 조명 변화에 민감하지 않은 컬러모델을 이용하여 마커를 사용하지 않고 사용자의 피부색 영역을 추출 및 트래킹하였다. 추출된 사용자의 손과 얼굴의 위치 정보를 단순화 시킨 인체모델과 역운동학적인 방법을 통해 사용자 동작을 추정하였다.

더 많은 카메라를 사용하여 자기 가림 현상(self occlusion)이 일어나는 문제를 해결할 수 있도록 한다. 단순화 시킨 인체모델을 보다 자세하게 정의하고, 운동학적 제한점을 보다 유연하게 정의하여 다양한 동작을 추정할 수 있도록 한다.

참고문헌

[1] Slater M, and Usho M, "Body Centered Interaction in Immersive Environments", Artificial Life and Virtual Reality, John Wiley and Sons, 1994.  
 [2] G. Welch, G. Bishop, L. Vicci, S. Brumback, K. Keller, and D. Colucci, "The hiball tracker: High-Performance wide-area tracking for virtual and augmented environments." Proc. VRST, pp.1-10, 1999  
 [3] Hyoson Kim, et al. "3D Modeling with Hand gesture interaction in a semi-immersive environment"  
 [4] A. Hogue, M. Robinson, M. R. Jenkin, R.S. Allison "A vision-based head tracking system for fully immersive display", International Immersive Projection Technologies Workshop, Eurographics Workshop on Virtual Environments (2003), 2003  
 [5] Thomas B. Moselund, Moritz Storrang and Erik Granum, "Vision-Based User Interfacing with a Virtual Environment", Proc. DANKOMB 2000, Aalborg University, Denmark, 2000  
 [6] 이선민, 박지영, 김명희, "혼합현실환경을 위한 능동 적외선 기반 동작배경 제거", 2004년도 한국정보과학회 가을 학술발표논문집 Vol3. 31, No.2, 2004  
 [7] 이선민, 박지영, 김명희, "스테레오 흑백-칼라 영상 쌍을 이용한 사용자 추출", 2004년도 한국정보과학회 가을 학술발표논문집 Vol3. 31, No.2, 2004  
 [8] D.Tolani, A. Goswami, and N. Badler, "Real time inverse kinematics techniques for anthropomorphic", Graphics Models 62(5), Sept. 2000