

## 혼합현실환경을 위한 능동 적외선 기반 동적배경 제거\*

박지영<sup>0,\*</sup> 이선민\* 김영희<sup>\*,\*\*</sup>

\*이화여자대학교 컴퓨터학과

\*\*이화여자대학교 컴퓨터 그래픽스/가상현실 연구센터

{lemie<sup>0</sup>, blue, mhkim}@ewha.ac.kr

### Dynamic Background Subtraction Based on Active Infrared for Mixed Reality Environment

Jiyoung Park<sup>0,\*</sup> Seon-Min Rhee\* Myoung-Hee Kim<sup>\*,\*\*</sup>

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

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

#### 요 약

사용자의 실영상을 가상세계에 합성하여 보여주는 형태의 혼합현실(mixed reality) 기술에서 입력영상으로부터의 사용자 추출은 필수적인 단계이다. 본 연구에서는 물체로부터 반사되는 적외선 정보를 이용하는 능동 적외선(Active Infrared) 기반 동적배경 제거 기법을 제안한다. 이 기법에서는 대형 디스플레이 환경 내에 설치된 적외선 광원 및 적외선 투과 필터를 장착한 흑백카메라를 이용하여 획득된 적외선 반사 영상에 배경 제거 알고리즘을 적용함으로써 사용자 영역을 추출한다. 제안 기법을 이용하면 CAVE™-like 시스템과 같이 정적배경을 보장하지 못하는 환경에서도 고가의 장비 없이 빠르고 정확하게 사용자를 추출할 수 있기 때문에 프로텍션 환경 내에 존재하는 사용자를 가상세계와 함께 합성하여 보여줄 수 있다.

#### 1. 서론

최근, 가상세계와 실사를 결합하여 보여줌으로써 사용자에게 극대화된 몰입감을 제공하는 혼합현실(mixed reality)에 관한 연구가 활발하게 진행되고 있다[1,2]. 특히 사용자의 모습을 촬영하여 실시간으로 가상세계와 합성하여 보여질 수 있도록 하는 비디오 아바타 생성 기법은 사용자에게 보다 친숙한 혼합현실 환경을 제공한다는 점에서 그 중요성이 점차 커지고 있다. 비디오 아바타를 생성하기 위해서는 촬영된 영상으로부터 사용자를 추출하기 위한 기법이 필수적으로 요구 되며 이를 위한 다양한 시도가 이루어지고 있다. 뿐만 아니라, 구름, 나뭇잎의 움직임, 빛, 디스플레이 장치 등에 의한 환경적인 변화 요소를 포함하는 영상에서 관심 객체를 추출하는 동적 배경에서의 객체 추출기법은 비디오 감시 시스템(video surveillance system)[3], 사용자 인터페이스, 물체 인식 등의 다양한 분야에서 중요한 이슈가 되고 있다[4].

비디오 영상에서 사용자를 추출하기 위해 현재까지 개발된 방법은 크게 네 가지 종류로 나누어 볼 수 있다. 크로마키(chroma keying)은 방송 분야에서 널리 사용되고 있지만 블루스크린과 같이 극도로 제한된 환경을 전제로 하며, 배경과 동일한 색상의 옷을 착용할 수 없기 때문에 일반적으로 사용되기 어렵다[5]. 배경제거(background subtraction) 기법은 학습된 참조영상을 이

용하여 입력영상에서 배경을 제거하고 관심영역만을 추출해내는 기법이다. 이 기법은 정적 배경이 전제되어야 하므로 대형 디스플레이와 같이 빛과 색이 수시로 변하는 환경이 배경상에 존재할 경우에는 정확한 추출결과를 보장하지 못한다. Depth keying은 영상에서 깊이를 측정하여 전경(foreground)에 해당하는 영역을 추출하는 방식으로, 사용자가 앞으로 움직일 경우 추출결과가 정확하지 않다. 최근에는 적외선 정보를 이용하여 사용자를 추출하는 기법들[6]이 시도되고 있으나 이러한 기법들은 열화상 카메라(thermal vision camera)와 같은 고가의 특수카메라를 이용해야 하며, 저온의 일반 객체의 움직임은 잡아낼 수 없다는 단점이 있다.

본 연구에서는 물체로부터 반사되는 적외선 정보를 이용하는 능동 적외선(active infrared) 기반 배경 분리 기법을 제안한다. 이 기법에서는 적외선 투과 필터(가시광선 차단)가 장착된 흑백카메라를 이용하여 적외선 광원으로 부터 나오는 빛을 반사한 사용자를 촬영하게 되므로 가시광선 변화로 인한 동적배경은 정적배경으로 처리될 수 있다. 이와 같은 방식으로 촬영된 적외선 반사 영상에 대하여 배경제거 알고리즘을 적용함으로써 CAVE™-like 시스템과 같이 대형 디스플레이를 이용하는 혼합현실환경에서 고가의 장비 없이 빠르고 정확한 사용자 추출 결과를 제공할 수 있다.

\* 본 연구는 부분적으로 정보통신부 대학정보통신연구센터(ITRC) 육성지원사업과 한국과학재단의 해외공동연구과제 지원에 의해 수행되었음.

## 2. 적외선 반사 영상 획득

본 연구에서는 실사 기반 사용자 영상과 가상세계를 함께 보여주기 위한 디스플레이 장비로 CAVE™-like 시스템을 사용하였으며, 스크린에 투사되는 영상으로 인한 배경 변화를 차단하기 위하여 다음과 같은 환경에서 적외선 반사 영상을 획득하였다.

### 2.1 하드웨어 구성

능동 적외선 방식을 이용한 사용자 추출 시스템은 적외선 광원과 적외선 투과 필터를 장착한 흑백카메라로 구성된다. 적외선 투과 필터는 그림 1[7]에서처럼 가시광선 영역(400-700nm)의 파장대의 빛은 차단하고, 주로 적외선 파장대역(700-2500nm)의 빛만 통과시킨다. 따라서, 광원으로부터 방출되어 사용자 몸에 반사된 적외선만이 필터를 통과하게 되고, 프로젝터로부터 스크린에 투사되는 가시광선 파장대의 빛은 필터에 걸려져 카메라로 입력되지 않는다. 본 연구에 사용된 적외선 광원은 Computar사의 IR75 Infrared Lamp (750nm, 2개)이며, 적외선 투과 필터는 Edmund Optics 사의 사각형태의 FILTER OPTCAST IR이다. 적외선 광원과 필터의 외형은 그림2와 같으며, 흑백카메라는 Point Grey Research사의 Dragonfly를 사용하였다.

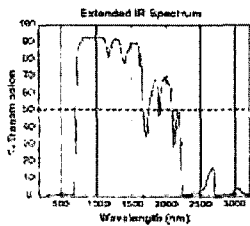


그림 1 적외선 투과 필터의 흡수 파장 대역

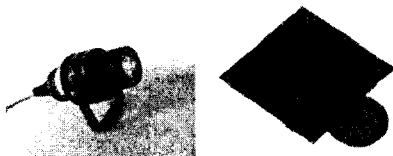


그림 2 적외선 광원(좌) 적외선 필터(우)

### 2.2 적외선 광원 및 카메라 위치

입력영상으로부터 사용자 영역을 쉽게 분리하기 위해서는 적외선 빛을 카메라 뷰잉 볼륨에 해당하는 영역 내에 골고루 분산시켜야 한다. 또한 적외선 빛이 사용자에게 반사되어 바닥에 그림자를 생성하지 않도록 광원의 위치를 설정해야 한다. 본 연구에서 디스플레이 환경으로 사용된 CAVE™-like 시스템은 정면, 좌/우측면, 바닥면의 총 4면의 스크린으로 구성되어 있다. 위의 고려사항을 만족시켜 최적의 입력영상을 얻기 위하여 적외선 광원을 좌, 우측면의 바닥에 설치, 그림자 생성을 감소시켰으며, 중심을 향하도록 함으로써 방출된 적외선이 반대쪽 스크린에 반사되어 골고루 분산될 수 있도록 하였다. 카메라

는 정면과 좌측면 또는 우측면이 만나는 코너의 상단에 위치시켜 사용자의 전신 촬영을 가능하게 하고 디스플레이면에 영향을 주지 않도록 하였다. 그림 3은 CAVE™-like 시스템에 구축된 영상획득 환경을 보여준다.

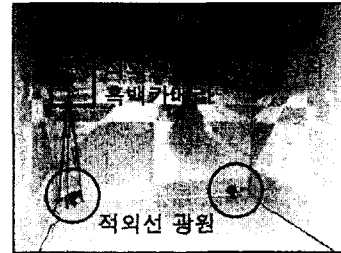


그림 3 적외선 영상 획득 환경

## 3. 적외선 영상에서의 사용자 추출 기법

본 연구에서는 위와 같은 환경에서 촬영된 적외선 반사 영상을 배경제거 알고리즘에 적용하여 사용자 영역을 빠르고 정확하게 추출하였다.

### 3.1 적외선 반사 영상의 특성

적외선 투과 필터를 통해 얻어지는 흑백영상(그림 4(b))은 일반 흑백영상(그림 4(a))과는 달리 사용자 영역이 다른 영역보다 상대적으로 밝게 나타나게 되며, 주변 환경, 특히 스크린에 디스플레이 되는 빛은 필터에 의해 걸러지므로 카메라에 입력되지 않는다. 따라서 적외선 반사 영상은 정적배경을 전제로 하는 배경분리 알고리즘에 적용할 수 있다.



(a) 적외선 필터 장착 전 (b) 적외선 필터 장착 후

그림 4 흑백카메라 영상

### 3.2 배경제거를 통한 객체 추출

배경제거는 전체리 단계인 배경검출 단계, 배경분리 단계, 후처리 단계로 구성된다.

배경검출 단계에서는 사용자가 존재하지 않는 배경 영상의  $n$ 개 연속 프레임이 저장되고 각 픽셀 위치의  $n$ 개 값에 대한 평균, 표준편차가 계산된다.

배경분리 단계에서는 사용자를 포함한 연속 영상이 입력되며 각 픽셀값과 배경검출 단계에서 계산된 동일 위치상의 통계값을 식 (1)을 이용하여 비교함으로써 변화 여부를 판별한다. 식(1)에서  $p$ 는 해당 픽셀,  $v$ 는 픽셀값,  $\mu$ 와  $\sigma$ 는 각각 습득영상에 대해 사전에 계산된 각 픽셀위치의 평균 및 표준편차이고  $k$ 는 상수이다. 입력영상의 픽

셀값( $v_p$ )과 습득 영상들의 평균값( $\mu_p$ )의 차이가 습득 영상들의 표준편차( $\sigma_p$ )의 k배 보다 크면 해당 픽셀은 변화가 일어났다고 간주되어 사용자 영역에 포함된다.

$$\text{If } (v_p - \mu_p) > k * \sigma_p \text{ then } p \text{ is foreground} \quad \text{식(1)}$$

후처리 단계에서는 팽창(dilation)과 미디언 필터를 적용하여 배경분리 단계를 거친 영상의 잡음을 제거하고 사용자 영역 및 실루엣을 보다 자연스럽게 나타내게 한다.

**4. 실험결과**

그림 5(a)는 CAVE™-like 시스템 내부에 설치된 흑백 카메라로 촬영한 영상이며, 그림 5(b)와 같이 적외선 투과 필터를 통과하여 사용자 추출 작업에 사용된다. 배경 습득이 완료되고 난 후, 사용자를 포함한 입력영상을 배경분리 및 후처리 단계에 적용시키면 그림 6과 같은 결과 영상을 얻을 수 있다. 결과영상을 통해 사용자의 그림자가 완전히 제거되었으며 후처리로 인하여 추출 영역의 윤곽선이 매우 자연스러워졌음을 알 수 있다. 또한 스크린 상에 콘텐츠가 투사되어 배경이 변화하는 환경에서도 사용자를 정확하게 추출할 수 있었다.

실험환경은 Dell Workstation PWS650(Intel Xeon™ CPU 2.80GHz, 2.79GHz, 2GB RAM)이며, 영상의 해상도는 640X480이다. 사용자 추출 단계에 소요된 시간은 약 0.026s/frame이다.



(a) 일반 흑백영상



(b)적외선 필터를 통과한 흑백영상  
그림 5 실험 영상



그림 6 사용자 추출 결과

**5. 결론 및 향후 연구**

본 연구에서는 실시간의 사용자 영상이 가상세계와 함께 보여지는 혼합현실환경을 구축하기 위하여 빠르고 정확한 사용자 추출 기법을 제안하였다. 적외선 광원 및 적외선 투과 필터를 사용하여 얻어진 영상을 이용함으로써 동적 배경을 포함하는 영상에 배경제거 알고리즘을 적용할 수 있었으며, 사용자 영역을 빠르고 정확하게 추출할 수 있었다. 뿐만 아니라, 고가의 특수 카메라 장비가 아닌 일반 CCD 카메라를 이용한다는 것 또한 제안된 기법의 강점이다. 추출된 영역의 컬러 정보는 컬러 카메라의 추가 설치로 획득될 수 있으며, 이를 통해 사용자 전신의 컬러 영상 추출이 가능하다.

이 기법을 이용하면 대형 디스플레이가 배경상에 존재하는 경우에도 배경제거 알고리즘을 적용할 수 있기 때문에 CAVE™-like 시스템과 같은 환경에 콘텐츠가 투사되는 동적배경에서도 사용자를 효과적으로 추출하고 동시에 트래킹 및 인터랙션 또한 가능하게 할 것으로 예상된다.

**참고문헌**

[1] Sang-Yup Lee, Ig-Jae Kim, Sang C Ahn, Heedong Ko, Myo-Taeg Lim, Hyoung-Gon Kim, " Real Time 3D Avatar for Interactive Mixed Reality," In Proceedings of the ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI'04), Singapore, 16-18 June 2004.

[2] O. Schreer and P. Kauff, "An immersive 3D video-conferencing system using shared virtual team user environments," in Proceedings of ACM Collaborative Environments(CVE 2002), Germany, Sept./Oct. 2002.

[3] T.B. Moeslund and E. Granum, " A Survey of Computer Vision-Based Human Motion Capture," J. Computer Vision and Image Understanding, vol. 81, no. 3, pp. 231-286, 2001.

[4] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, " Wallflower: Principles and practice of background maintenance," in Proc. ICCV ' 99, pp. 255- 261, 1999.

[5] Wojciech Matusik, " Image-Based Visual Hulls" , Master's thesis, MIT, 2001.

[6] Yasuda, K., Naemura, T., Harashima, H., " Thermo-Key: Human Region Segmentation from Video," IEEE Computer Graphics and Applications, Vol. 24, No. 1, pp. 26-30, 2004.

[7] Edmund Optics Inc.  
<http://www.edmundoptics.com/>