

## 모션기반 인터랙티브 증강현실 시스템을 위한 연기자 위치 자동 정합 기법

양기선, \*남승진, \*\*김창현

KBS 기술연구소, \* KBS 기술연구소, \*\*고려대학교

ksyang@kbs.co.kr, \*sjnam@kbs.co.kr, \*\*chkim@korea.ac.kr

### Auto-Registration Of The Actor Position For The Motion-Based Interactive Augmented Reality System

Ki-Sun Yang, \*Seung-Jin Nam, \*\*Chang-Hun Kim  
KBS, \*KBS, \*\*Korea University

#### 요 약

본 논문은 HD 해상도의 고화질 방송카메라(회전과 줌이 가능한)와 ‘키넥트’ 같은 범용 모션인식 카메라를 연동하여 방송용 모션기반 증강현실 어플리케이션을 구현하기 위해서, 연기자의 위치를 가상 월드 좌표로 자동 변환시키는 방법을 제안한다. 방송환경에서 키넥트와 같은 모션인식카메라(RGB-D 카메라)를 사용하기 위해서는 거리와 조명의 영향을 많이 받아 연기자와 가까운 천장이나 카메라 밑처럼 그 설치 위치가 제약되곤 한다. 이때, 연기자와 그래픽의 정확한 합성을 위해, 모션인식 카메라로부터의 연기자 추적 정보를 월드 좌표계로 변환해야 하며, 이것을 수작업으로 하기에는 변환의 정확도가 많이 떨어지고 시간도 많이 소요된다. 본 논문에서는 체크박스 패딩 기반의 스테레오 카메라보정 알고리즘을 사용하여 모션인식카메라와 방송용 카메라간의 상관관계(이동, 회전)를 찾고, 이 값과 월드 좌표계 상의 방송 카메라(하드웨어 센서에 의해서 추적됨)정보(이동, 회전)를 이용하여 임의 위치에 설치한 모션인식 카메라로부터 추적된 연기자 위치정보를 월드 좌표계 상의 위치 값으로 자동 변환시키는 방법을 제안한다.

Key word: motion, interaction, augmented reality, application, registration, position, kinect

#### 1. 서론

오래전부터 가상 환경에서 인터랙션을 구현하기 위한 많은 시도가 있어 왔다. 일반적으로 가상 공간의 연기자는 그래픽과의 인터랙션을 위해서 특별한 도구를 사용하여 가상정보를 제어하였다. 이 도구는 인식률을 높이기 위해서 특별한 트랙볼을 사용하거나 내부에 자이로 같은 별도의 센서를 부착하기도 하였다. 그러나, 최근에 키넥트[1]와 같은 모션인식카메라가 일반화 되면서 별도의 장치를 사용하지 않더라도 사람의 모션을 통해서 그래픽과의 상호 작용이 가능하게 되었다.

키넥트와 같은 모션인식카메라를 외부카메라와 함께 사용한 응용 사례[2][3]를 보면, 대부분 초점거리가 고정된 범용 카메라를 사용하였으며, 고화질 카메라의 움직임(pan, tilt, rotate)은 가상 카메라의 움직임과 연동되지 않는다. 실사, [2]처럼, 초점거리가 조절되는 카메라를 사용하더라도 최종적으로는 FOV(Field Of View)를 고정해서 사용하였다. [3]에서는 모션인식카메라의 ‘Tilt’ 정보를 실제 바닥의 깊이 정보를 추출하여 찾기도 하였다. 따라서, 회전 정보를 찾기 위해서는 모션인식카메라를 이용하여 바닥의 일부 깊이 정보를 반드시 추출하여야 하며, 이로 인하여 카메라의 설치 위치에 제약이 따른다. 모두 고화질 카메라와 모션인식카메라를 고정된 위치에서 사용하여 증강현실을 구현할 수 있었다. 모션 센서를

사용하는 대부분의 어플리케이션은 낮은 해상도의 깊이 영상과 조명과의 거리에 다른 적외선 센서의 인식률 저하 문제로, 방송카메라와 연기자간에 거리가 다소 먼 거리(3~5m)에서 촬영하는 방송같은 프리미엄 콘텐츠 제작 환경에서 사용하기에는 제약이 따른다. (TV station에서는 방송 카메라와 연기자간의 거리가 모션인식카메라가 인지하는 거리 보다 먼 거리에서 촬영이 이루어지는 것이 일반적이다.) 따라서, 이때는 모션인식카메라를 인식률이 좋은 위치로 이동시킬 필요가 있다(단, 카메라 앵글에 걸리지 않는 연기자와 가까운 거리). 그러나, 이 경우에 모션인식카메라의 위치는 방송 카메라의 위치와 크게 달라지게 때문에 결과적으로 연기자 위치를 월드 좌표계 상의 그래픽위치와 수동으로 자연스럽게 정합시키기에는 매우 번거로우며, 조정을 위해 많은 시간을 필요로 한다.

본 논문에서는 TV station에서 모션인식카메라를 자유로운 위치에 설치하여 연기자의 추적 정보를 월드 좌표계 상의 위치로 자동 변환시키는 방법을 제안한다. TV station의 가상 스튜디오는 기본적으로 월드 좌표계 원점을 기준으로 가상 스튜디오의 가상환경과 실제환경이 정확하게 일치하도록 보정되어 있다. 우리는 월드 좌표계 상에 위치한 방송용 카메라의 위치와 회전 정보를 카메라에 부착된 센서로부터 실시간으로 알 수 있다. 본 논문에서는 이 정보를 기반으로 모션인식 카메라의 위치와 방향을 찾고, 연기자의 특징점을 월드 좌표계 기준으로 변환하는 방법을 제안한다. 이렇게 함으로써 모션인식

카메라의 위치를 임의로 위치 시킬 수 있어, 방송 카메라와 거리에 상관 없이 원하는 위치에 설치 하여 실시간으로 자동 변환함으로써 연기자의 월드 좌표계 상의 위치를 찾을 수 있다.

본 논문은 2 장에서 시스템 구성, 3 장에서 사용자 인터페이스 및 그 기능을 설명하고, 4 장에서 카메라보정 및 위치 변환에 관해서 자세히 설명한다. 5 장에서는 본 시스템을 이용하여 가상 시나리오에 따라 수행한 실험내용과 모션기반 인터랙티브 일기예보 어플리케이션을 개발하여 본 시스템을 검증하였다. 6 장에서 결론을 말한다.

## 2. 시스템 구성 및 가상환경 시스템

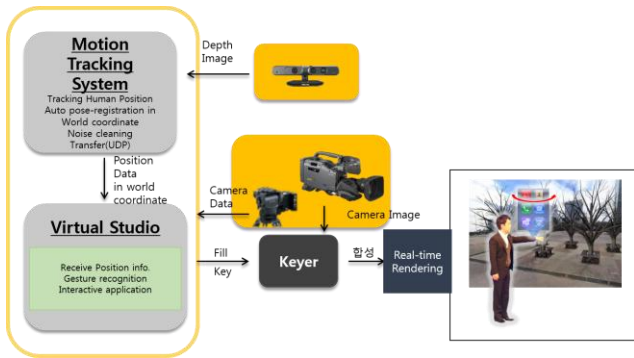


그림 1. 시스템 구성

그림 1. 처럼, 가상 스튜디오용 모션기반 인터랙티브 시스템의 구성은 다음과 같다. 크게 모션 추적 시스템 부와 가상스튜디오 부로 나뉘며, 방송 카메라 SONY HD Camera, Camera Sensor, Keyer, Virtual Studio Software (VizArtist), 모션인식 카메라(Primesense) [4], Motion Tracking System 으로 구성된다. 모션인식 카메라로부터 추적된 연기자의 위치 정보는 월드 좌표계로 보정되고, 노이즈가 필터링되어 Virtual Studio 툴로 전송된다. 이 정보는 Virtual Studio 에서 월드 좌표기준으로 인터랙티브 어플리케이션의 제스처 인식을 위해서 사용된다. 가상 스튜디오에서는 그래픽 영상의 Key 신호와 Fill 신호를 카메라 영상 신호와 함께 Keyer 에서 합성하여 최종 영상을 송출한다.

## 3. 사용자 인터페이스

그림 2. 의 모션 추적 시스템부의 사용자 인터페이스는 다음과 같다. 1 번은 서버 주소와 포트 번호를 입력하여 동시에 두 시스템에서 정보를 수신할 수 있으며, 보내는 시간 주기도 설정가능 하다. 수신되는 그래픽 합성 플랫폼은 한국의 TV Station 에서 많이 사용되는 VizArtist 와 BrainStorm(Virtual Studio)에서 모두 가능하도록 하였다. 2 번은 모션인식 카메라 설정부분이다. 카메라 on/off(Use Depthcam), 깊이영상과 RGB 영상 뷰포인트 정합(Set Depth Viewpoint), 노이즈 필터링(Smoothing), 손 이벤트 인식(Threshold), 추적할 연기자 설정(User), 어깨 추적 보완기능을 지원한다. 3 번은 원하는 추적 포인트를 선택하여 전송할 수 있으며, 4-1 은 수동으로 모션인식 카메라의 위치와 방향을 설정할 수 있다. 4-2 는 자동으로 모션인식 카메라를 설정할 수 있도록 하였다. (자세한 내용은 3.카메라 보정 및 위치 자동 보정 참고)

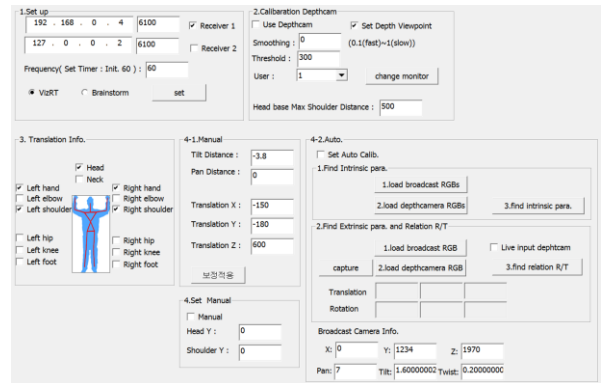


그림 2. 사용자 인터페이스

## 4. 연기자 위치 정보 변환

### 4.1 방송카메라 및 모션인식카메라 내부,외부 파라미터 추출

우리는 월드 좌표계 상의 방송 카메라의 위치와 방향을 알고 있음을 전제로 하였다. 실제로 방송용 카메라의 트래킹 센서를 통하여 카메라의 월드 좌표계 상의 위치와 방향이 가상 스튜디오 툴 (VizArtist [7])로 실시간으로 전송된다. 이때에는 널리 알려진 OpenCV 의 Camera Calibration [8]을 사용하여, 7x5 의 체크보드 패턴을 이용하여 각 카메라의 내부/외부 파라미터들을 추출하였다.

### 4.2 두 카메라간 상관관계 추출

4.1 에서의 두 카메라 각각의 내부 파라미터와 외부 파라미터를 사용하여 OpenCV [6]의 StereoCalibrate () 함수를 사용하여 상관 관계 파라미터 R,T(방송카메라 기준 모션인식 카메라의 회전과 위치정보)를 찾을 수 있다. Stereo Calibrate () 함수는 두 패턴 평면간의 에피폴라 평면 [5]을 해석하여, 객체 좌표에서 3D 점 P 가 주어지면, 각각의 카메라에서 단일 카메라 보정을 수행하여 점 P 를 카메라 좌표로 변환하는 방법으로 R,T 를 얻을 수 있게 한다.

### 4.3 모션인식카메라 월드 좌표계 상 위치 찾기

R,T 를 찾으면, 방송 카메라 좌표계를 기준으로 한 모션인식카메라와의 위치(T)와 방향(R)에 관한 상관관계를 찾은 것이다. 방송카메라의 좌표계에서 원점이 방송카메라의 위치와 방향을 의미하며,  $O_{2\_depthcam}$  은 T 에 의해 찾은 모션인식카메라의 방송카메라 기준 상대위치이다.  $O_{w\_2\_depthcam}$  는  $O_{2\_depthcam}$  를 월드 좌표계 상의 원점 기준으로 바꾼 위치이다. 다음으로 방송카메라의 위치를 월드 좌표계의 원점으로 하고 그 상관 정보 T 에 따라 모션인식카메라 역시 월드 좌표계 기준으로 변환 시킨다. 우리는 이미 방송카메라의 월드상의 위치 정보와 방향 정보( $R_{1\_broadcast}$ )정보를 하드웨어 트래킹 장비로부터 알고 있으므로, 변형된 좌표점 ( $O_{w\_1\_broadcast} = Origin(0,0,0)$ ,  $O_{w\_2\_depthcam}$ )을 방송 카메라의 방향정보  $R_{1\_broadcast} = (R_{x\_tilt}, R_{x\_pan}, R_{x\_twist})$ 에 의해 회전변환 후

방송카메라 위치 정보  $T_{w\_1\_broadcast}$  에 의해 이동시켜 모션인식카메라의 월드 좌표계 상의 위치를 찾는다

#### 4.4 추적점 월드 좌표계 상 위치 찾기

모션인식카메라의 월드 위치를 알고 있으므로, 다음으로는 모션인식카메라의 회전 정보를 찾는 것이다. 모션인식카메라의 회전정보는 앞에서 찾은 두 카메라간의 회전에 관한 정보  $R=(R_x, R_y, R_z)$ 로부터 찾을 수 있다. 모션인식카메라 좌표계에서 추적중인 특징점을  $P_{depthcam} = (P_{depthcam\_x}, P_{depthcam\_y}, P_{depthcam\_z})$ 이라고 하면, 이 점을 월드 좌표계 상의 원점기준 좌표로 이동시켜, 모션인식카메라의 월드 좌표계 상의 위치로부터 이동행렬  $T_{w\_depthcam}$  를 찾는다. 결과적으로 찾고자 하는 월드 좌표계 상의 특징점의 위치는 모션인식카메라와 방송카메라간의 상대 회전정보  $R$  에 의해 회전 후  $T_{w\_depthcam}$  만큼 이동시켜 찾을 수 있다.

### 5. 실험 및 방송 적용

#### 5.1 실험

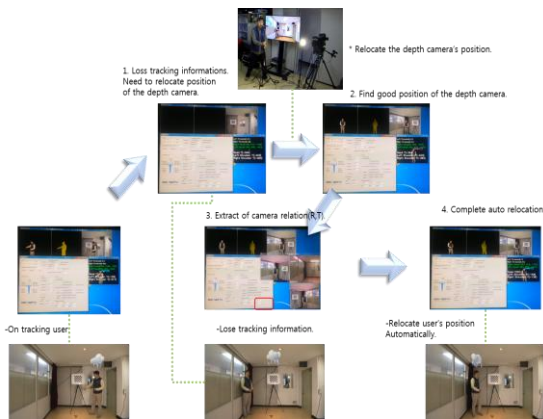


그림 3. 모션인식카메라의 설치 위치 변경 후, 연기자의 월드 좌표계 상의 자동 위치 변환 과정

우리는 먼저 가상 스튜디오의 방송 카메라를 정확하게 월드 보정 후(월드 좌표계 상의 원점기준 가상카메라와 연동됨)에 제안된 방법으로 변환된 연기자 머리의 월드 좌표 위치를 기준으로 일정 오프셋 위에 항상 구름이 따라다니도록 실험 어플리케이션을 개발 하였다(그래픽 구름의 위치 정보(CGx, CGy,CGz)는 연기자 머리의 위치 정보(CGx, CGy+offset, CGz)와 같음). 따라서, 카메라의 위치 변동이 발생할 경우, 보정을 하지 않으면, 연기자와 그래픽의 위치는 틀어지게 된다. 본 실험은 실제 방송 제작에서 발생할 수 있는 시나리오를 가정하여 수행되었다. 시나리오는 다음과 같다.

1. 연기자 추적 시작 - 월드 좌표상의 위치 변환 초기화
2. 연기자 추적 실패 상황 발생 (예, 모션인식카메라가 움직이거나 연기자 위치가 추적영역에서 벗어남)
3. 모션인식카메라 재 설치(연기자 보이도록 조정)

4. 방송카메라 기준 모션인식카메라의 상대위치 및 회전 재 보정(R-T 찾고, 위치 변환)

5. 변환 완료, 연기자 월드 좌표 위치 추적 시작

그림 3 에서와 같이 모션인식 카메라의 위치를 변경 하였더라도 연기자를 따라 다니는 구름의 위치가 잘 따라 다남을 확인할 수 있다.

### 6. 결론

우리는 방송 카메라의 월드 좌표계 상의 위치 정보를 기준으로 모션인식카메라에 의한 연기자 위치(추적중인 모든 골격정보)를 월드 좌표계 상의 위치로 자동으로 변환하였고, 본 시스템을 사용하여 ‘인터랙티브 일기예보 방송프로그램’에 적용하여 잘 정합됨을 확인하였다. 가상 스튜디오에서 방송 카메라는 연기자의 전신을 보여주기 위해서 연기자와 다소 떨어진 위치(2~4m 이상)에 설치 되며, 모션인식카메라는 방송 카메라의 앵글에 걸치지 않으면서 연기자의 추적이 용이한 근거리에서 설치하기 위해서 천장에 설치하는 것이 바람직하였다. 그리고, 변환 후에 더 정확한 정합을 위해서는 약간의 오프셋 조절을 필요로 하였다. 그렇지만, 본 논문에서 제안하는 방법을 사용함으로써, 모션인식카메라의 설치 위치가 자유로워졌으며, 배우의 동선에 따라 자유롭게 모션인식카메라의 설치가 가능하고, 연기자와 그래픽 정합을 위해 초기 위치 조정하는 데에 소요되는 작업시간을 개선할 수 있었다. 향후에 연기자의 추적 영역을 넓히기 위해서 다수의 모션인식 카메라를 사용하는 것을 고려 중에 있다.

### 참고 문헌

- [1] Kinect for Window.<http://www.microsoft.com/en-us/kinectforwindows/meetkinect/default.aspx>
- [2] Giovanni, S. ; Choi, Y.C., Huang, J., Khoo, E.T., Yin, K., Virtual Try-On Using Kinect and HD Camera. International conference 5th Motion in Game, Rennes, France, 2012.
- [3] Jaemin Soh 5et al., Automatic registration of a virtual experience space with Kinect. Multimedia Tools And Applications, 2013.
- [4] Primesense. Available from : <http://en.wikipedia.org/wiki/Primesense>. Oct 11 2012.
- [5] Epipolar geometry theory. Available from: [http://en.wikipedia.org/wiki/Epipolar\\_geometry](http://en.wikipedia.org/wiki/Epipolar_geometry). Dec 25 2012.
- [6] OpenCV. Available from: <http://opencv.org/>
- [7] Viz Artist tools by VizRT. Available from: [http://www.vizrt.com/products/viz\\_artist/](http://www.vizrt.com/products/viz_artist/)
- [8] Camera calibration with OpenCV, Available from: [http://docs.opencv.org/trunk/doc/tutorials/calib3d/camera\\_calibration/camera\\_calibration.html](http://docs.opencv.org/trunk/doc/tutorials/calib3d/camera_calibration/camera_calibration.html)