

증강현실 기반의 공간 상호작용을 위한 깊이 카메라 적용

Introducing Depth Camera for Spatial Interaction in Augmented Reality

윤경담, Kyungdahm Yun*, 우운택, Woontack Woo†

요약 기존의 증강현실 기반 상호작용 기법은 상호작용 주체의 원활한 추적을 위해 마커 기반의 패들이나 데이터 글러브 등의 추가적인 입력장치에 의존하거나 제한적인 3차원 입력만을 허용하는 경우가 많았다. 본 논문에서는 대상 개체 주변 공간의 물리적 점유를 감지하여 사용자의 입력을 해석하는 비접촉식 수동형 공간 상호작용 기법을 제안한다. 제안된 방법은 깊이 카메라로 구성된 증강현실 환경에 가상의 공간 센서를 배치하여 사용자의 입력을 공간적으로 해석한다. 그 결과, 추가적인 입력장치의 착용이 필요하지 않으며, 입력 형태에 대한 제약을 최소화하고, 증강된 영상 내의 상호작용 객체 간에 정확한 거리감을 제시한다. 향후 이러한 기법은 미니어처 AR 시스템과 같은 동적 콘텐츠 전시 플랫폼의 인터페이스에 적용될 것으로 기대된다.

Abstract Many interaction methods for augmented reality has attempted to reduce difficulties in tracking of interaction subjects by either allowing a limited set of three dimensional input or relying on auxiliary devices such as data gloves and paddles with fiducial markers. We propose Spatial Interaction (SPINT), a non-contact passive method that observes an occupancy state of the spaces around target virtual objects for interpreting user input. A depth-sensing camera is introduced for constructing the virtual space sensors, and then manipulating the augmented space for interaction. The proposed method does not require any wearable device for tracking user input, and allow versatile interaction types. The depth perception anomaly caused by an incorrect occlusion between real and virtual objects is also minimized for more precise interaction. The exhibits of dynamic contents such as Miniature AR System (MINARS) could benefit from this fluid 3D user interface.

핵심어: *Spatial Interaction, Depth Camera, Augmented Reality*

본 연구는 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소육성사업의 연구결과로 수행되었음.

*주저자: 광주과학기술원 U-VR 연구실; e-mail: kyun@gist.ac.kr

†교신저자: 광주과학기술원 U-VR 연구실; e-mail: wwoo@gist.ac.kr

3.1 계층적 공간 구조

증강공간 S_i 는 그림 2 처럼 상호작용 계층(interaction layer) I_i 와 활성 계층(activation layer) A_i 로 구성된다.

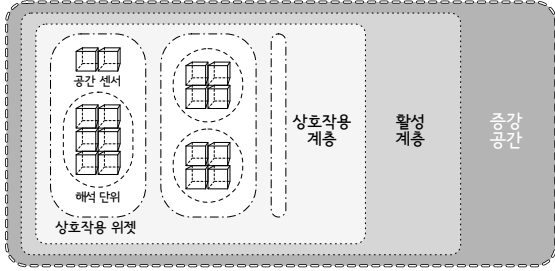


그림 2. 증강공간의 계층적 구조.

상호작용 계층 상호작용 주체 \mathcal{S} 와의 물리적인 접촉을 인식하는 가상의 공간 센서(space sensor)들이 배치되어 있다. 공간 센서 c 는 작은 상자 형태를 취하고 있으며 물리적인 침투 유무에 따라 $\{F, O\}$ 중 한 가지 상태를 가진다. F 는 점유되지 않은 상태(free), O 는 점유된 상태(occupied)이다. 공간에 대한 점유(occupancy)가 발생하면 연동되어 있는 상호작용 단위 $U_{ij} = \{u | u \in W_{ij} \in W_i\}$ 들에게 정보가 전달되고, 이를 바탕으로 현재 입력된 상호작용 주체의 행동이 해석된다. 상호작용 단위들은 다양한 형태의 상호작용 위젯 $W_i = \{w | w \in I_i \in S_i\}$ 로 조합되어 사용자에게 제시된다.

활성 계층 증강공간 S_i 의 외곽을 차지하며 공간 내부의 상호작용 계층 I_i 에 접근하기 위한 통로 역할을 한다. 다수의 증강공간에 의해 전체 장면이 복잡해지는 것을 방지하기 위해 모든 증강공간들은 물리적인 접촉이 불가능하고 시각적으로는 투명한 상태로 초기화된다. 임의의 상호작용 주체 \mathcal{S} 가 활성 계층 A_i 로 침투하면 해당 공간 전체가 활성화되어 사용자의 시야에 등장하며, 이때부터 내부의 상호작용 단위들이 공간의 점유 상태에 대한 관측을 시작한다.

3.2 상호작용 해석 단위

공간 센서들의 집합 $C_{ijk} = \{c | c \in U_{ijk} \in U_{ij}\}$ 는 임의의 수량과 형태로 조합되어 상호작용 해석 단위(interaction unit) U_{ij} 를 구성한다. 각 단위들은 소속되어 있는 공간 센서들이 전달해주는 점유 정보를 바탕으로 상호작용 주체의 행동에 대한 개별적인 해석을 시도한다. 방향 추출기(direction extractor) U_{DE} , 선택 지연기(delayed selector) U_{DS} , 경로 인식기(path detector) U_{PD} 등의 단위들이 고안되었다.

3.3 상호작용 위젯

상호작용 위젯 w_{ij} 는 해석 단위들의 집합 $U_{ij} = \{u | u \in W_{ij} \in W_i\}$ 를 임의로 조합하여 구성하며, 사용자의 입력에 대한 직접적인 인터페이스 역할을 담당한다.

수직 슬라이더 위젯 세로 방향으로 배치된 공간 센서들을 방향 추출기에 등록하여 구성한다. 추출된 방향 정보에 의해 위젯의 지시자가 1차원 축을 따라 GUI의 스크롤 바처럼 이동하며, 이때의 출력 신호는 지시자의 상대적인 위치이다.

선택 버튼 위젯 특정 영역에 군집한 공간 센서들을 지연 선택기에 등록하여 구성한다. 상호작용 주체가 해당 위젯에 상주하여 발생한 점유율과 점유 시간이 특정 조건을 만족하는 경우, 연동되어 있는 개체에 대한 선택 신호가 발생된다.

제스처 인식 위젯 특정 제스처의 형상으로 연결된 공간 센서들을 경로 인식기에 등록하여 구성한다. 상호작용 주체가 해당 경로를 따라 일정 시간 내에 순차적으로 이동을 완료하면 제스처가 인식된다. 각 인식기는 한 가지 종류의 제스처만을 인식하며, 다수의 제스처를 인식하기 위해 위젯 내부의 상호작용 단위들이 중첩되도록 배치할 수도 있다.

3.4 깊이 카메라 적용

깊이-거리 보정 마커에 의한 추적을 가정하는 경우, 시간 t 에 영상 속 마커 영역 M 에 대한 깊이 값의 평균 $\bar{d}_{depth}(t)$ 를 계산한 다음, 모델-뷰 변환 행렬에서 획득한 카메라와 마커 간의 거리 $\bar{d}_{distance}(t)$ 를 이용하여 깊이-거리 보정 모형 $d_{distance} = \alpha d_{depth} + \beta$ 의 계수 α 와 β 를 추정한다. 실시간으로 수행되는 본 추정 과정에는 적절한 크기($n = 500$) 이상으로 수집한 표본에 대해 최소 제곱법을 활용한다.

공간 점유 검사 공간 센서의 점유 상태는 깊이 지도 D 상의 각 점 p_{screen} 에서 증강공간 S 의 모든 공간 센서 c 에 대한 선분-메쉬 충돌 검사를 수행함으로써 확인한다. 효율성을 위해 공간 내부의 객체들은 k -D Tree 로 정렬한다.

객체 차폐 검사 가상 객체가 실제 물체의 전면에 위치한 것과 같은 시각적 표현의 오류로 인해 발생하는 부정확한 거리감 문제를 해결한다. 가상 객체의 영상버퍼 F 에 기록된 픽셀 조각의 깊이 d_{object} 와 깊이 지도상 동일 지점에서 획득한 깊이 d_{scene} 을 비교한다.

4. 실험 결과 및 분석

4.1 비접촉식 수동형 상호작용

제안된 기법의 비접촉식 수동적인 특성을 확인하고, 기존의 마커 기반 접촉식 능동형 기법과 성능을 비교하기 위한 실험을 수행하였다. 우선 그림 3의 환경에서 40 mm 크기의 마커가 부착된 패들을 추적하여 각 프레임 간의 마커 위치에 따른 방향 벡터를 계산한다. 그리고 동일한 상황에서 5×5×5 형태로 배치된 15 mm 크기의 공간 센서를 이용하는 방향 추출기 U_{DE} 를 이용해 계산된 벡터와 비교한다.

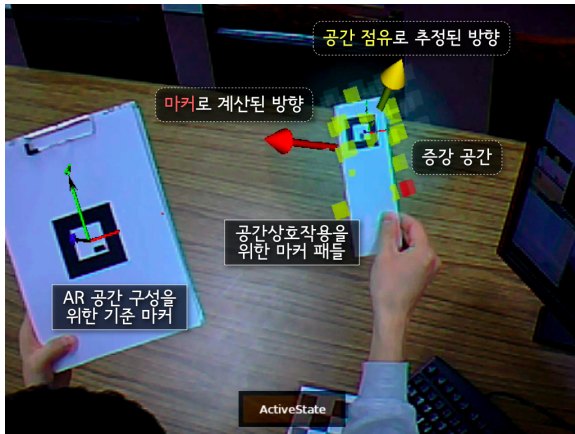


그림 3. 기존의 능동형 추적 기법과 제안된 수동형 추적 기법의 성능 비교를 위한 실험 구성.

실험 결과, X축, Y축, Z축 방향에 대한 상관관계는 각각 0.7467, 0.6422, 0.5074로 높은 수준이었으며, 방향 전환 등의 전반적인 경향을 유사하게 추정해냈다는 것을 그림 4에서 확인하였다.

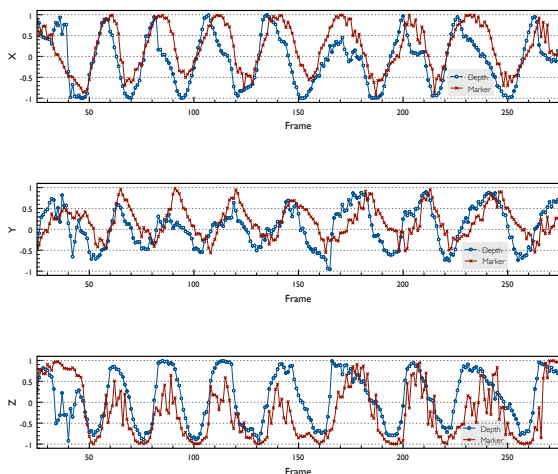


그림 4. 마커로 계산한 방향 벡터와 공간적으로 해석한 벡터의 비교. 상단부터 각각 X축, Y축, Z축의 결과.

4.2 유연한 상호작용 입력

상호작용 주체의 종류에 의한 제약을 최소화하고 사용자가 원하는 형태의 입력을 지원하기 위해서는 범용적인 공간 해석이 가능해야 하며, 다양한 형태의 입력에 대해서도 동일한 공간 점유 형상을 유지한다는 것을 전제로 한다. 이를 확인하기 위해 그림 5의 환경에서 동일한 부피의 물체 표면을 4가지 서로 다른 물리적 특성의 재질로 변경하며 공간 점유도 검사를 실시하였다.



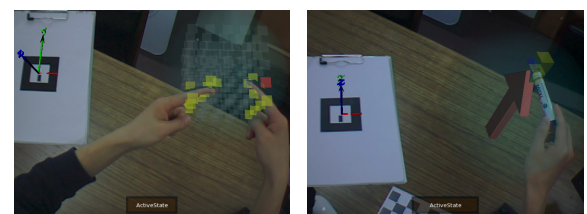
그림 5. 다양한 물체의 공간 점유 비교를 위한 실험 구성.

표 1의 결과에 의하면 대부분의 물체들은 유사한 공간 점유도를 보였으며, 반사가 심한 비닐의 경우에만 이들보다 낮은 값을 기록하였다. 이것은 ZCam에서 발사된 적외선 신호가 교란되어 해당 영역에 대한 깊이 값을 제대로 측정하지 못하기 때문에 발생하는 현상으로, 이러한 특이성만 고려하면 상호작용 주체의 선택에 큰 제약이 없음을 확인하였다.

표 1. 서로 다른 종류의 물체에 의한 공간 점유도.

재질	종이	플라스틱	비닐	모직
공간 점유도	0.5237	0.4790	0.1403	0.5891

그림 6은 동일한 증강공간에서 다양한 상호작용 주체를 이용하여 상호작용하는 장면이다. 기존의 상호작용 기법들이 미리 정해진 주체만을 인식했던 것에 비해 임의의 물체를 사용할 수 있기 때문에 보다 유연한 상호작용이 가능하다.

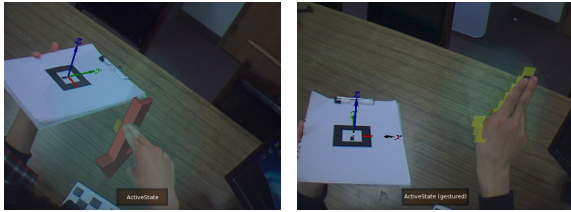


(a) 양손.

(b) 필기구.

그림 6. 임의의 상호작용 주체 활용.

그림 7은 공간 정보의 유연한 해석에 근거하여 구성할 수 있는 상호작용 위젯들의 사례를 보여준다.



(a) 수직 슬라이더. (b) 제스처 인식.

그림 7. 공간 해석 방식에 따른 다양한 상호작용 위젯 구성.

4.3 사실적인 차폐 표현

증강현실 기반 상호작용에서 가상 객체와 실제 객체 간의 차폐 현상과 이에 의거한 거리감(depth perception)은 사용성에 큰 영향을 미치는 요소이다 [10]. 깊이-거리 보정의 유무에 따른 거리감의 차이를 비교하기 위하여 그림 8의 환경에서 마커 전방 300 mm 지점에 가상 객체를 증강하고, 그 주변에서 움직이는 실제 객체를 통해 차폐 현상이 발생하는 위치를 측정하였다. 보정된 환경에서는 그림 9(a)처럼 299 mm 지점에서 차폐가 발생하여 정확한 거리감을 제공하였지만, 보정되지 않은 환경에서는 그림 9(b)처럼 훨씬 더 전방의 388 mm 지점이 되어야 차폐가 일어났다. 즉, 보정이 된 환경에서는 증강현실 공간 좌표계에 의거하여 실제 객체와 가상 객체 간의 정확하고 사실적인 차폐 현상을 표현할 수 있다는 것을 확인하였다.



그림 8. 깊이-거리 보정 유무에 따른 차폐 표현의 성능 비교를 위한 실험 구성.

5. 결론 및 향후 연구

본 논문에서는 깊이 카메라 기반의 공간 상호작용 기법을 제안하였으며, 실험을 통해 비접촉식 수동형 인터페이스의



(a) 보정 상태의 차폐 성공. (b) 미보정 상태의 차폐 실패.

그림 9. 보정 유무에 따른 차폐 표현 결과 (299 mm 지점).

성과 상호작용 입력의 유연성, 보정에 의한 사실적인 차폐 표현 능력을 확인하였다. 향후에는 이러한 공간 상호작용 기법을 미니어처 AR 시스템 등의 상용 플랫폼에 적용하여 직접적인 사용성 평가를 진행할 계획이다.

참고 문헌

- [1] W. Woo, N. Kim, K. Wong, and M. Tadenuma, "Sketch on dynamic gesture tracking and analysis exploiting vision-based 3D interface," *Proceedings of SPIE*, vol. 4310, pp. 656–666, Dec 2000.
- [2] D. Hong and W. Woo, "A 3D vision-based ambient user interface," *International Journal of Human-Computer Interaction*, vol. 20, pp. 271–284, Jul 2006.
- [3] J. Corso, G. Ye, D. Burscbka, and G. Hager, "A practical paradigm and platform for video-based human-computer interaction," *IEEE Computer*, vol. 41, pp. 48–55, May 2008.
- [4] G. Ye, J. Corso, and G. Hager, "Visual modeling of dynamic gestures using 3D appearance and motion features," in *Real-Time Vision for Human-Computer Interaction*, ch. 7, pp. 103–120, Springer-Verlag, Aug 2005.
- [5] G. Yahav, G. Iddan, and D. Mandelbom, "3D imaging camera for gaming application," *ICCE '07: Proceedings of the 25th International Conference on Consumer Electronics*, pp. 1–2, Jan 2007.
- [6] J. Leitner, M. Haller, K. Yun, W. Woo, M. Sugimoto, and M. Inami, "IncreTable, a mixed reality tabletop game experience," *ACE '08: Proceedings of the 5th International Conference on Advances in Computer Entertainment Technology*, pp. 9–16, Dec 2008.

- [7] V. Buchmann, T. Nilsen, and M. Billinghurst, "Interaction with partially transparent hands and objects," AUIIC '05: Proceedings of the 6th Australasian User Interface Conference, vol. 40, pp. 17–20, Jan 2005.
- [8] M. Kanbara, T. Okuma, H. Takemura, and N. Yokoya, "A stereoscopic video see-through augmented reality system based on real-time vision-based registration," VR '00: Proceedings of the 2nd IEEE Virtual Reality Conference, pp. 255–262, Mar 2000.
- [9] J. Fischer, B. Huhle, and A. Schilling, "Using time-of-flight range data for occlusion handling in augmented reality," EGVE '07: Proceedings of the 13th Eurographics Symposium on Virtual Environments, pp. 109–116, Jul 2007.
- [10] V. Jurgens, A. Cockburn, and M. Billinghurst, "Depth cues for augmented reality stakeout," CHINZ '06: Proceedings of the 7th ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction, pp. 117–124, Jul 2006.