

실내 구면 파노라마에서의 볼륨 검출을 통한 AR 게임

공간 구성

김기식[○], 박종승
인천대학교 컴퓨터공학과
{kisik, jong}@inu.ac.kr

Creating AR Game Space via Volume Detection from Indoor Spherical Panorama

Ki-Sik Kim[○], Jong-Seung Park
Dept. of Computer Science & Engineering, Incheon National University

요 약

본 논문에서는 구면 파노라마를 기반으로 직육면체 형태의 실내 공간에서 볼륨 검출을 통해 AR 게임 공간을 구성하는 방법을 제안한다. 기존의 공간 인식 방법은 철저하게 관측 가능한 공간에 대해서만 평면 검출이 가능하며, 밀도와 노이즈의 변화에 민감하다. 또한 기존의 AR 게임 콘텐츠는 평면 카메라에 의존적이기 때문에 공간 인식을 위한 데이터 공급량에 한계가 있다. 제안 시스템은 기존의 평면 검출 방식에서 벗어나 Oriented Bounding Box를 탐색하고 교정하는 방법을 통해 AR 게임 공간을 구성한다. 또한, 제안 시스템은 구면 파노라마를 기반으로 하여 AR 게임 공간을 구축해 공간 인식의 효율을 극대화한다. 제안 방법은 실험을 통해 기존의 방법보다 수행 속도가 월등히 빠르고, 밀도와 노이즈의 변화 등 환경 제약 요소에 강건함을 보인다.

1. 서 론

실내 Augmented Reality(AR) 게임 환경에서 콘텐츠를 구현하기 위해서는 AR 시스템이 실제 공간을 인식할 유의미한 정보를 검출해야 한다. 이 정보는 대표적으로 점군, 오브젝트, 평면이 있다. 이 정보들 중 평면 인식은 바닥면과 벽면, 테이블, 장애물 등 공간을 가장 직관적으로 인식하는 방법이다. 평면 인식은 실내 AR 공간을 구성하는 가장 핵심적인 기술이다.

현재 AR 게임 공간 구축은 평면 검출에 의존하고 있다. 기존의 평면 검출 방법에는 주로 Hough Transform[1], RANSAC[2], Region Growing[3]이 복합적으로 활용된다. 기존의 평면 검출 방법은 입력 데이터인 점군의 영역을 분할하고, 분할 영역

에 대한 원시 평면을 획득하여 주변으로 평면의 영역을 확장하는 Bottom-Up 방법을 기반으로 한다[4][5]. 평면 검출 기술은 해당 알고리즘을 기반으로 노이즈에 강건하고 정확한 평면을 검출하기 위해 많은 발전을 거듭했다. 그러나 기존의 평면 검출 방법은 근본적인 한계점이 존재한다. 기존의 방법은 철저하게 관측 가능한 공간에 대해서만 평면 검출이 가능하기 때문에 점군이 없는 공간의 평면 검출은 불가능하다. 그리고 밀도가 낮을 경우 원시 평면이 검출되지 않거나 검출률이 급격하게 저하된다. 검출을 위해 점군 샘플의 법선 벡터를 획득하며, 영역을 분할하는 등의 많은 전처리 과정으로 실시간성이 떨어진다.

AR 콘텐츠를 활용하는 대부분의 실내 공간은 방이다. 대부분의 방은 직육면체 모양으로, 가장

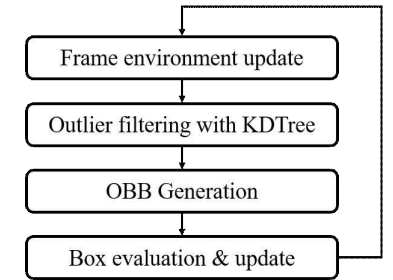
크고 중요한 평면은 바닥과 천장, 벽면이다. 이에 적합한 알고리즘은 대부분의 실내 공간에서 활용이 가능하며, 직육면체 공간에서 매우 효율적이다.

AR에서 가장 중요한 것은 공간의 '인식'과 시각적 '표현'이다. 인식된 공간이 다양하고 넓을수록 표현의 자유도는 증가하기 때문에, 인식의 효율성은 AR에서 중요하다. 기존의 AR은 평면 카메라에 의존적이다. 평면 카메라는 시야각이 한정되어 있기 때문에 AR 공간 인식을 위한 데이터 공급량에 한계가 있다. 구면 카메라와 같이 시야각의 제한이 없는 카메라에 기반할 경우 인식을 위한 데이터의 공급이 매우 많아질 뿐만 아니라 사용자의 시야 밖을 활용한 콘텐츠 생성 역시 가능하다.

본 논문에서는 직육면체 형태의 실내 공간에서 점군 데이터를 기반으로 볼륨 검출을 통해 빠르고 정확하게 AR 게임 공간을 구성하는 방법을 제안한다. 제안 방법은 공간 인식의 효율성을 극대화하고 AR 공간 속 표현의 자유도를 높이기 위해 구면 파노라마 비디오를 활용한다. 제안 방법은 바닥, 천장, 벽면의 주요 평면 정보를 검출하기 위해 기존의 알고리즘에서 벗어나 Oriented Bounding Box(OBB) 검출[6] 방법을 활용한다. 제안 방법은 분할된 최소 영역에서 확장하는 Bottom-Up 방식이 아닌, 전체의 영역을 한 번에 감싸 교정하는 Top-Down 방식이다. 따라서 점군의 밀도에 영향을 받지 않고, 별도의 전처리 과정이 없다. 제안 방법은 검출된 볼륨을 정량적으로 평가하고, 실시간으로 더욱 좋은 품질의 볼륨 데이터로 갱신하며 AR 공간의 품질을 향상시킨다.

2. 실내 공간의 볼륨 검출

[그림 1]은 제안하는 볼륨 검출 방법의 흐름도이다. 제안하는 볼륨 검출 시스템은 3개의 메인 파트로 구성된다. 먼저, 구면 파노라마 기반 Vision SLAM을 통해 점군 데이터를 획득하고, KDTree를 구성해 이상점을 제거한다. 다음으로, 점군 데이터를 감싸는 최적의 OBB를 생성한다. 마지막으로, 생성된 OBB의 볼륨의 품질을 수치화하여 정량적으로 평가하고, 좋은 품질의 OBB로 갱신한다.



[그림 1] 제안하는 볼륨 검출 흐름도

2.1 이상점이 제거된 점군 획득

제안하는 시스템은 실시간으로 점군 데이터를 획득하기 위해서 OpenVSLAM[7]을 활용하였다. 해당 SLAM은 구면 파노라마 비디오를 입력으로 하여 영상 속 특징점에 대한 3차원 점군과 실시간 카메라의 위치 및 자세 추정 정보를 생성한다. 해당 SLAM은 우수한 정확성을 보이지만, 영상 분석을 통한 추정이기 때문에 점군 데이터는 오차를 포함한 비선형적 데이터이다.

점군 속 샘플의 미세한 오차는 AR 공간 인식에서 큰 영향을 주지 않지만, 오차가 큰 샘플은 지대한 영향을 줄 수 있다. 제안하는 시스템에서는 치명적인 이상점을 제거하기 위해 3차원 KDTree를 활용하였다. KDTree는 BST를 기반으로 하여 다중 차원에서 한 점을 중심으로 각 축들을 번갈아가며 주변의 점과의 거리에 따라 노드를 구성한다.

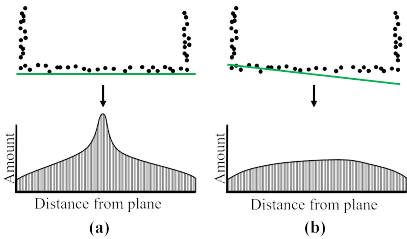
$$e = E(X) + \sigma(X). \quad (1)$$

본 시스템에서는 모든 점의 평균에 가장 가까운 점을 기준으로 KDTree를 구성하였다. 그리고 모든 점의 가장 가까운 이웃점과의 거리 집합 X 에 대해 평균값 $E(X)$ 와 표준편차 $\sigma(X)$ 를 획득하고, X_i 가 두 값을 더한 e 보다 클 경우 i 번째 점을 이상점으로 분류하였다.

2.2 OBB 검출과 품질 평가

OBB는 점군의 영역 전체를 감싸는 최적의 박스이다. 이상적인 환경에서 직육면체 공간의 점군에 대한 OBB는 직육면체 공간의 표면을 정확하게 감싼다. 본 시스템은 Naujoks 등[8]이 제안한 OBB 검출 알고리즘을 활용한다. 해당 알고리즘은 점군을 둘러싸는 최소한의 점들인 Convex Hull을 검출하여 점들을 감싸는 박스를 생성하고, Convex Hull의 극점을 활용한 박스 교정을 통해 좋은 품질의 박스를 생성한다.

제안 시스템은 실시간으로 더욱 좋은 품질의 볼륨 데이터로 갱신하며 AR 공간의 품질을 향상시킨다. 이를 위해서는 OBB의 품질을 수치화하여 정량적으로 평가하는 과정이 필요하다. 제안 시스템은 박스의 6면에 대해 평면상에 존재하는 점이 많을수록, 그리고 점들이 적절히 분포되어있을수록 정확하게 검출한 박스라 가정하였다.



[그림 2] 평면 품질 평가 예시: (a) 좋은 평면의 예시, (b) 좋지 않은 평면의 예시

$$s = \sum_{i=1}^6 \frac{n_i}{\sigma(D_i)} \quad (2)$$

제안 시스템은 박스의 6개의 평면에 대해 식 (2)과 같은 방법으로 품질 점수 s 를 산출했다. 만약 6개의 평면 중 i 번째 평면에 대한 점수를 구한다고 할 때, 평면상의 점의 개수 n_i 를 평면상의 점들의 거리에 대한 집합 D_i 에 대한 표준편차 $\sigma(D_i)$ 로 나누어준다면 해당 평면의 품질 점수를

구할 수 있다. [그림 2]의 예시처럼, 점수가 낮은 경우 평면상에 점이 충분하지 않거나, 점들의 거리가 일정하지 않아 표준편차가 큰 경우이다. 반대의 경우에는 점수가 높다.

본 시스템에서는 매 프레임마다 OBB 볼륨에 대한 점수를 산출하여 이전에 존재했던 가장 높은 품질의 OBB보다 점수가 높다면 OBB를 갱신한다. 반대의 경우 기존의 OBB 상태를 유지한다.

3. 볼륨 데이터 기반 AR 게임 환경 구성

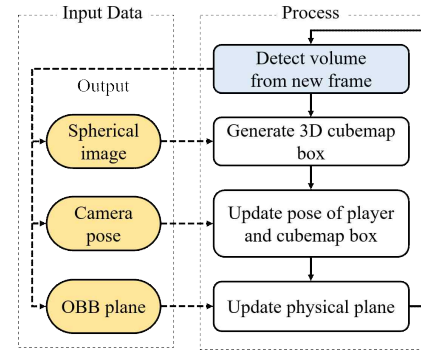
AR 게임 환경을 구성하기 위해서는 2장의 과정을 통해 인식한 공간의 정보와 실제 공간에 대한 영상을 일치시키는 과정이 필요하다. 이를 위해서는 가상 세계 안에서 비디오를 촬영하는 카메라의 위치 및 자세를 알아야 한다. 그리고 비디오와 함께 해당 위치 및 자세에서 좌표계가 변환된 가상 세계의 콘텐츠를 동시에 출력해 완전하게 가상 세계와 비디오의 시점이 일치해야 한다. 본 장에서는 가상 세계를 구축하고 실제 세계와 일치시키는 방법과, AR 콘텐츠의 기반이 되는 3차원 오브젝트의 스폰 지점을 계산하는 방법을 설명한다. 본 논문에서는 Unreal Engine 4 환경에서 AR 게임 환경을 구축하였다.

3.1 가상 세계 구축과 실제 세계와의 일치화

일반적인 평면 영상을 기반으로 하는 AR 환경에서는 평면 영상 위에 가상 세계를 덮어서 출력해 사용자에게 보여준다. 제안하는 시스템은 구면 파노라마를 기반으로 하므로 차이가 있다. 인간의 눈은 핀홀 투영을 통해 평면 영상 정보를 만든다. 따라서 왜곡량이 많은 구면 파노라마 정보를 인간의 시각에서 보기 위해 왜곡이 없는 360° 영상으로 변환이 필요하다.

[그림 3]은 가상 세계 구축과 일치화에 대한 흐름도이다. 흐름도 속 파란색 프로세스는 2장의 볼륨 검출 과정을 의미한다. 노란색 데이터 블록은 볼륨 검출 과정에서 획득할 수 있는 데이터이다. 볼륨 검출 과정 이후에는 3개의 파트로 나누어 수

행하여 가상 세계를 구축하고 실제 세계와 일치화한다.



[그림 3] 가상 세계 구축과 일치화 흐름도

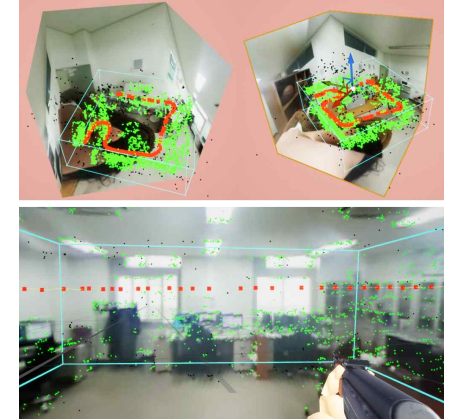
먼저, 구면 파노라마의 왜곡을 제거하기 위해 3차원 큐브맵 이미지로 변환한다. 큐브맵 이미지는 거대한 스카이박스처럼 가상 세계 안에 거대한 정육면체 형태로 배치한다. 정육면체의 안쪽 면에 큐브맵 이미지의 6면이 투영되어 가상 세계의 전체를 감싸는 형태이다.

다음으로, 가상 세계 안에서의 카메라의 위치 및 자세를 반영한다. OpenVSLAM은 실시간으로 영상 속 특징점에 대한 3차원 점군과 카메라의 위치 및 자세 추정 정보를 생성한다. 따라서 카메라의 위치 및 자세 정보는 OpenVSLAM을 통해 획득한다. AR 콘텐츠는 1인칭이기 때문에 플레이어의 위치 및 자세는 광학 장치의 위치 및 자세이다. 실제 세계의 영상과 가상 세계가 일치되어야 하므로 가상 세계에 배치된 큐브맵 정육면체의 위치 및 자세 역시 카메라의 위치 및 자세로 변환한다.

마지막으로, 2장의 방법을 통해 획득한 OBB의 6개의 평면을 가상 세계에 배치한다. 해당 평면은 가상 세계상에서 점군을 분석해 획득한 평면이기 때문에 별도의 위치 및 자세 변환을 하지 않는다.

[그림 4]의 위쪽 그림은 3.1장의 내용을 통해 구축된 가상 세계의 모습이다. 큐브맵 육면체는 훨씬 거대하지만 예시를 위해 [그림 4]에서만 사이즈를 축소하였다. 아래쪽 그림은 사용자의 1인칭 시점에

서 실제 세계와 가상 세계가 일치된 모습이다. 아래쪽 그림에는 실제 세계의 영상, 특징점 점군, 카메라 이동경로, OBB의 6개 평면 정보가 동시에 출력되어 있다. 이는 개발 측면에서의 출력이기 때문에, 사용자에게 제공되는 AR 콘텐츠에서는 모든 정보를 감추고 3D 모델 등 콘텐츠 요소만 출력한다.



[그림 4] 구현된 AR 게임 환경: (위)구축된 가상 세계와 (아래)일치화된 플레이어 시점의 모습

3.2 가상 세계에서의 오브젝트 스폰

일반적인 3차원 오브젝트의 스폰은 AR 콘텐츠의 시작점이다. 획득한 평면을 활용하기 위해서는 각각 평면의 좌표에서의 오브젝트를 생성해야 한다. 먼저, 오브젝트의 스폰 지점을 평면의 중앙 지점이라 가정할 때 오브젝트의 위치 및 자세는 식 (3), (4)과 같다.

$$R_n = \begin{bmatrix} r_R \\ r_P \\ r_Y \end{bmatrix} = \begin{bmatrix} 0 \\ \arctan(z_n / \sqrt{x_n^2 + y_n^2}) \\ \arctan(y_n / x_n) \end{bmatrix} \quad (3)$$

$$P_o = [R_o | t_o] = [Mat(R_n) | p_c] \quad (4)$$

평면의 법선 벡터 $\vec{v}_n = [x_n, y_n, z_n]^T$ 의 자세(회전량) R_n 은 식 (3)과 같은 방법을 통해 획득할 수 있다. 오브젝트의 초기 위치 P_o 은 R_n 의 회전행렬 $Mat(R_n)$ 와 평면의 중앙 지점 위치 \mathbf{p}_c 을 통해 식 (4)과 같이 표현된다.

오브젝트가 평면에 안착하여 이동하기 위해서는 평면의 좌표계에서의 계산이 필요하다. 평면의 좌표계에서 캐릭터를 \vec{v}_t 방향으로 이동시키고자 한다면, 월드 좌표계에서는 $\vec{v}_{wt} = Mat(R_n)\vec{v}_t$ 방향으로 이동시키면 된다.

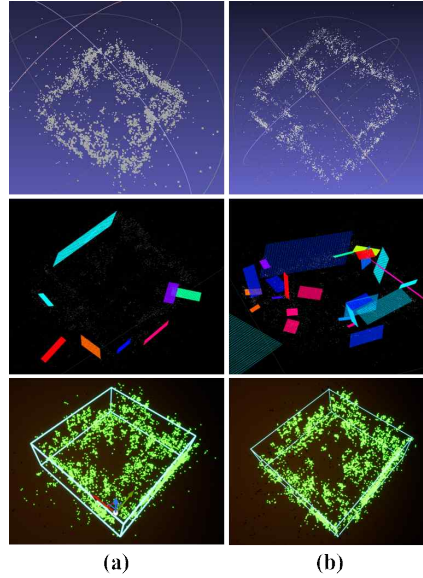


[그림 5] 볼륨 데이터를 통해 스폰된 3D 오브젝트

4. 실험 결과

제안하는 볼륨 검출 알고리즘의 성능을 평가하기 위해 실험을 수행하였다. 실험은 기존의 다중 평면 검출 알고리즘 중 최근에 제안되고 좋은 성능을 보인 Araujo 등[9]이 제안한 시스템과 비교하였다. 실험을 위해 OpenVSLAM을 통해 획득한 실측 점군 데이터를 통해 두 시스템의 평면 검출 결과를 비교하였다. 본 논문에서 제안하는 시스템은 실시간으로 점점 나온 평면을 검출할 수 있는 알고리즘이지만, 동일한 조건을 가정하기 위해 완성된 점군 데이터에 대해 실험하였다.

[그림 6]과 [표 1]은 실험에 대한 결과이다. 실험은 SLAM을 통해 획득한 (a), (b) 두개의 점군 세트를 통해 수행했다.



[그림 6] 평면 검출 결과: 위에서부터 점군 원본 데이터, 기존의 평면 검출 방법, 제안 방법

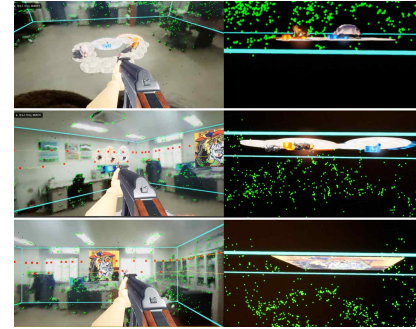
[표 1] 두 방법의 수행 시간

실험 세트	수행 시간(ms)	
	기존 방법	제안 방법
(a)	89,421	29
(b)	92,124	30

OpenVSLAM으로 획득한 실측 데이터는 점군의 밀도가 낮고, 노이즈가 포함되어 있다. 점군의 밀도와 노이즈에 많은 영향을 받는 기존의 평면 검출 방법의 경우 두 점군 세트 모두 좋은 검출 결과를 보이지 않았을 뿐만 아니라, 실내 공간의 6개의 면 중 한 개의 면만 검출되었다. 반면, 제안 방법의 경우 6개의 면에 대해 모두 좋은 평면 검

출 결과를 보였다. 수행 시간에서도 큰 차이를 보였다. 전처리 과정이 많고 반복 수행이 많은 기존의 방법의 경우 평균 90772ms의 수행 시간으로, 약 90초 이상의 시간이 필요했다. 제안 방법의 경우 평균 29.5ms의 수행 시간이 필요했다. 기존의 방법은 볼륨 검출을 통해 밀도가 낮고 노이즈가 큰 점군을 입력으로 매우 안정적인 결과를 보였다. 또한, 짧은 수행 시간을 통해 실시간성을 유지하였다.

가상 세계의 오브젝트는 현실 세계와 잘 동화되었다. 오브젝트는 추정된 평면 위에 안착하여 올바르게 생성되었다. 카메라가 이동하는 등 현실 세계의 시점이 변화해도 오브젝트가 있어야 하는 지점에서 흔들림 없이 잘 고정되어 있었다.



[그림 7] 가상세계 오브젝트 스폰 결과

5. 결 론

본 논문에서는 직육면체 형태의 실내 공간에서 볼륨 검출을 통해 주요 평면을 검출하는 방법을 제안하였다. 제안 방법은 OBB 검출 방법을 활용하여 빠르고 정확하게 평면을 검출할 수 있으며, 구면 파노라마 기반 OpenVSLAM을 통해 공간 인식의 효율성을 극대화하였다. 제안 방법은 구면 파노라마를 기반으로 하여 AR 게임 환경을 구성하였다. 실험 결과 기존의 다중 평면 검출 방법보다 밀도와 노이즈의 변화에 강건하며, 수행 속도가 월등히 빠름을 확인할 수 있었다.

제안 방법은 실내 AR 콘텐츠가 사용되는 공간

의 특성을 반영하여 논문에서 가정하는 모든 직육면체 형태의 공간에서 활용할 수 있다. 향후에 제안 방법을 확장하여 직육면체 형태가 아닌 다른 형태의 공간에 대한 모양을 추정할 수 있다면 활용 방안이 확대될 것이다. 또한, 기존의 다중 평면 기술과의 조합을 통해 바닥, 천장, 벽면 등의 6면은 제안 방법을 통해 획득하고, 기존의 방법을 통해 부분적인 평면 검출을 수행한다면 많은 평면을 가진 더욱 좋은 품질의 AR 게임 공간을 구축할 수 있을 것이다.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2019R1F1A1060828).

참고문헌

- [1] A. S. Hassanein, S. Mohammad, M. Sameer, M. E. Ragab, "A Survey on Hough Transform, Theory, techniques and Applications", arXiv:1502.02160, 2015.
- [2] R. Raguram, J. M. Frahm, M. Pollefeys, "A Comparative Analysis of RANSAC Techniques Leading to Adaptive Real-Time Random Sample Consensus", In Proc. European Conf. computer vision, pp. 500-513, 2008.
- [3] S. A. Hojjatoleslami, J. Kittler, "Region growing: a new approach," IEEE Trans. Image Processing, vol. 7, pp. 1079-1084, 1998.
- [4] M. Hulik, M. Spanel, P. Smrz, Z. Mateme, "Continuous plane detection in point-cloud data based on 3D Hough Transform", J. Visual Commun. Image Represent., 2014.
- [5] Lin Li, Fan Yang, Haihong Zhu, Dalin Li, You Li, Lei Tang, "An Improved RANSAC for 3D Point Cloud Plane Segmentation Based on Normal Distribution Transformation Cells", Remote Sensing, 2017.
- [6] J. O'Rourke, "Finding Minimal Enclosing Boxes," Int. J. Parallel Programming, vol. 14, no. 3, 1985.

- [7] S. Sumikura, M. Shibuya, K. Sakurada, "OpenVSLAM: A Versatile Visual SLAM Framework," In Proc. 27th ACM Int. Conf. Mult., pp. 2292-2295, 2019.
- [8] B. Naujoks, H. J. Wuensche, "An Orientation Corrected Bounding Box Fit Based on the Convex Hull under Real Time Constraints," In Proc. IEEE Int. Vehicles Symp., pp. 415-420, 2018.
- [9] Abner M. C. Araujo, Manuel M. Oliveira, "A Robust Statistics Approach for Plane Detection in Unorganized Point Clouds," Pattern Recognition, 2019.