

실내 환경에서 직육면체 검출을 통한 AR 게임 공간 구성

김기식, 박종승
인천대학교 컴퓨터공학부
{kisik, jong}@inu.ac.kr

Constructing AR Game Space through Cuboid Detection in Indoor Environment

Ki-Sik Kim, Jong-Seung Park

Dept. of Computer Science and Engineering, Incheon National University

요 약

본 논문에서는 직육면체 형태의 실내 환경에서 볼륨 검출을 통해 AR 게임 공간을 구성하는 방법을 제안한다. 기존의 공간 인식 방법은 관측 가능한 제한된 공간에 대해서만 평면 검출이 가능하며, 밀도와 노이즈의 변화에 민감하다. 제안 방법은 기존의 평면 검출 방식에서 벗어나 구면 파노라마로부터 OBB를 탐색하는 방법을 통해 AR 게임 공간을 구성한다. 제안 방법은 실험을 통해 기존의 방법보다 수행 속도가 빠르고, 또한 밀도와 노이즈의 변화 등의 환경 제약 요소에 강건함을 보였다.

ABSTRACT

In this paper, we propose a method of constructing AR game spaces through cuboid detection in indoor environment. Conventional spatial recognition methods can detect planes only in limited spaces that can be well observed. They are also vulnerable in density and noise. The proposed method overcomes the limitations of the conventional method by constructing AR game spaces by a method of detecting OBBs from spherical videos. Experimental results showed that the proposed method is faster than the conventional method and it is also robust against environmental constraints such as changes in density and noisy.

Keywords : Augmented Reality(증강 현실), Point Cloud(점군), Game Space(게임 공간), Cuboid (직육면체)

Received: Jul. 15. 2021 Revised: Aug. 26. 2021
Accepted: Sep. 03. 2021
Corresponding Author: Jong-Seung Park (Incheon National University)
E-mail: jong@inu.ac.kr

ISSN: 1598-4540 / eISSN: 2287-8211

© The Korea Game Society. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

컴퓨터 비전에서 공간 인식 문제는 현실 공간을 분석하여 유의미한 기하 정보를 도출하는 것을 목적으로 한다. 기하 정보는 점군, 오브젝트, 평면 등 다양한 형태로 표현된다. 인식된 공간의 정보는 더욱 고도화된 다양한 응용 결과를 생성하는 기반이 된다. 실내 Augmented Reality(AR) 게임 환경에서 게임 콘텐츠를 구현하기 위해서는 공간의 기하 정보가 필요하다. 평면 인식은 바닥면과 벽면, 테이블, 장애물 등 공간을 가장 직관적으로 인식하는 방법이다. 현재의 AR 게임 공간 구축 방법은 평면 인식에 의존하고 있으며, 이는 실내 AR 공간을 구성하는 가장 핵심적인 기술이다.

실내 AR 콘텐츠를 활용하는 대부분의 실내 공간은 직육면체 형태의 공간이다. 실내 공간에서 가장 크고 중요한 평면은 가상 콘텐츠를 생성하기에 충분히 크고 평평한 바닥면과 천장면, 벽면이다. 이러한 공간 표현 방식은 대부분의 실내 공간에서 활용이 가능하며, 특히 직육면체 공간에서 매우 효율적이다.

기존의 평면 검출 방법은 점군을 입력 데이터로 하여 점군의 영역을 분할하고, 분할 영역에 대한 원시 평면을 획득하여 주변으로 평면의 영역을 확장하는 bottom-up 방식을 기반으로 한다. 평면 검출 기술은 이러한 알고리즘을 기반으로 정확하고 노이즈에 강건한 평면을 검출하기 위해 많은 발전을 거듭했다. 그러나 기존의 평면 검출 방법은 근본적인 한계점이 존재한다. 기존의 방법은 철저하게 관측 가능한 공간에 대해서만 평면 검출이 가능하므로 점군이 없는 공간의 평면 검출은 불가능하다. 그리고 밀도가 낮을 경우 원시 평면이 검출되지 않거나 검출률이 급격하게 저하된다. 또한, 많은 전처리 과정으로 실시간성이 떨어진다.

AR에서 가장 중요한 세부 기술은 공간의 인식과 시각적 표현이다. 인식된 공간이 다양하고 넓을수록 표현의 자유도는 증가하기 때문에, 인식의 효율성은 AR에서 중요하다. 기존의 AR은 평면 카메라

에 의존적이다. 평면 카메라는 시야각이 한정되어 있기 때문에 AR 공간 인식을 위한 데이터 공급량에 한계가 있다. 구면 카메라와 같이 시야각의 제한이 없는 카메라를 기반으로 하는 경우 인식을 위한 데이터의 공급이 매우 많아질 뿐만 아니라 사용자의 시야 밖을 활용한 콘텐츠 생성 역시 가능하다.

본 논문에서는 실내 공간에서 점군 데이터를 기반으로 직육면체 공간 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 바닥면, 천장면, 벽면의 주요 평면 정보를 검출하기 위해 기존의 알고리즘에서 벗어나 Oriented Bounding Box(OBB) 검출 방법을 활용한다[1]. 제안 방법은 분할된 최소 영역에서 확장하는 bottom-up 방식이 아닌, 전체의 영역을 한 번에 감싸 교정하는 top-down 방식의 접근이다. 따라서 점군의 밀도에 영향을 받지 않고, 별도의 전처리 과정이 없다. 제안 방법은 검출된 기하 공간을 정량적으로 평가하고, 실시간으로 더욱 좋은 품질의 공간 데이터로 갱신하여 AR 공간의 품질을 높인다.

2. 관련 연구

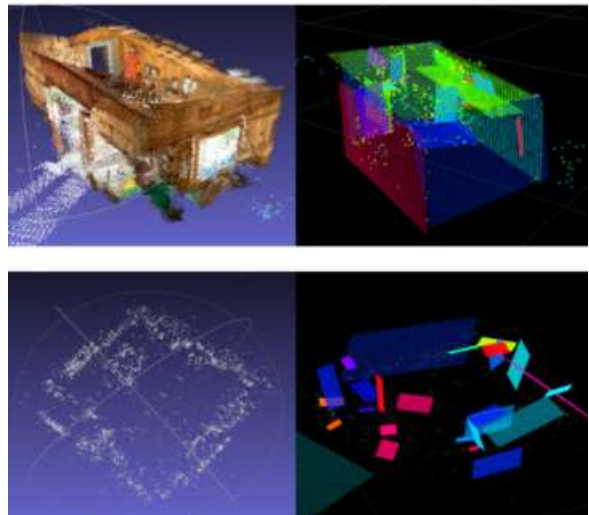
과거부터 공간을 인식하기 위한 평면 검출은 활발히 연구되었다. 기존의 평면 검출 방법에는 주로 Random Sample Consensus (RANSAC)[2], Hough Transform[3], Region Growing[4] 등의 방법이 복합적으로 활용되고 있다. 초기에는 평면 검출을 위해 점군 실측 데이터 치명적인 이상점에도 강건한 RANSAC 기반의 검출 방법[5]이 제안되었고, 이후 RANSAC의 비일관성이 가지는 한계점을 극복하기 위해 Hough Transform 기반의 검출 방법[6]이 제안되었다. 해당 방법들은 모든 점군을 반복적으로 검색하며 탐색하기 때문에 같은 샘플을 중복하여 검색해야 했다. 또한, 점군의 평

면성이 일관되지 않을 경우 평면의 영역 결정에 대한 오차가 많았다. 이를 보완하기 위해 영역을 분할하고 평면성이 우수한 원시 평면을 기준으로 영역을 확장해나가는 Region Growing 기반의 검출 방법[7]이 제안되었다. 이후 다른 많은 평면 검출 알고리즘이 등장했으나, 너무 많은 조정 인자가 필요하거나 출력물의 일관성이 지나치게 떨어지는 등의 이유로 인해 RANSAC, Hough Transform, Region Growing을 중심으로 발전해왔다.

최근에 제안된 평면 검출 알고리즘은 주로 점군 전처리와 함께 점군의 영역을 분할하고, 분할 영역의 원시 평면을 획득하여 평면의 영역을 확장하고 교정하는 알고리즘이다. 예시로, Araujo 등[8]이 제안한 방법은 리프노드가 n 개 이하가 될 때까지 점군 샘플의 영역을 분할한 Octree를 구성하고, 각 리프노드의 점들이 속해 있는 부분적인 영역의 점들에 대한 Median Absolute Deviation(MAD) Estimator를 활용해 기존의 Region Growing에서 활용했던 PCA 기반 방법보다 노이즈에 강건한 Planar Patch를 획득한다. 이후 유사한 평면성을 가지는 Planar Patch들을 합치며 평면의 영역을 확장한다. Hulik 등[9]이 제안한 방법의 경우 점군 샘플을 2-Level 스케일의 영역으로 분할하여 Hough Transform을 기반으로 평면을 검출 후 평면을 교정, 삭제한다. Li 등[10]이 제안한 방법은 점군을 NDT[11] 복셀 영역 분할하여 각 영역의 평면성을 판단하여 검출한 후 평면을 교정, 합체한다.

이러한 방법은 최소영역에서 확장하며 교정하는 bottom-up 방식으로, 근본적인 한계점이 존재한다. 먼저, 점군의 밀도에 많은 영향을 받는다. 원시 평면을 생성하고, 영역을 확장하는 과정에서 밀도에 변화가 많을 경우 평면 생성이 되지 않거나 정확한 확장을 할 수 없다. 따라서 밀도가 높은 데이터에서는 좋은 결과를 보이지만, 밀도가 낮아질 경우 검출률이 급격하게 하락한다. 그리고 철저하게 관측 가능한 공간만 평면 검출이 가능하다. 만약 점군이 관측되지 않은 영역이 있다면, 하나의 실제

평면이 두 개 이상으로 나뉘어 표현될 수 있다. [Fig. 1]은 Araujo 등의 방법에 대해 입력 점군의 밀도에 따른 결과의 차이를 보여준다. 높고 균일한 밀도의 점군의 경우 벽면과 바닥을 비롯해 시각적으로 분석이 가능한 반면, 낮은 밀도를 가지는 실측 데이터의 경우 바닥, 천장, 벽면을 비롯한 많은 평면이 검출되지 않거나 두 개 이상으로 나누어져 검출되었다. 또한, 기존의 방법은 영역 분할과 점군의 법선 벡터를 획득하는 등의 많은 전처리 과정을 가지기 때문에 연산량이 많고 실시간성이 떨어진다. Araujo의 방법은 독창적인 Statistical Estimator를 활용한 Planar Patch 검출을 통해 기존의 평면 검출 방법보다 노이즈에 강건하지만, 최소 영역에서 확장하며 교정하는 bottom-up 방식의 근본적인 한계점을 여전히 가지고 있다.



[Fig. 1] The results of the Araujo's plane detection method according to the density of the input point cloud: (upper) dense points, (lower) sparse points.

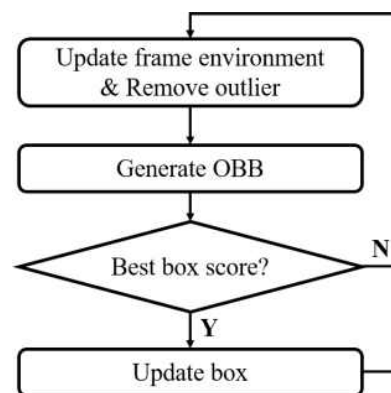
OBB 검출은 점군을 감싸는 최소한의 박스를 검출하는 기술이다. OBB 검출 기술을 통해 직육면체형 공간의 6면을 검출한다면 기존의 평면 검출이 가지는 한계점을 극복할 수 있다. OBB 검출은 점군을 감싸는 최소한의 울타리 점인 Convex Hull[12]을 획득하여 볼륨 검출이 수행된다. 이 때, 원본 점군과 Convex Hull의 OBB 검출 결과가 동

일하다. 이는 OBB 검출 결과가 점군의 밀도에 거의 영향을 받지 않음을 보여준다. 또한, OBB 검출 과정에서 사용하는 L-Shape[13]과 같은 기술을 통해 관측된 면의 점군을 기반으로 마주 보는 관측되지 않은 면을 검출할 수 있도록 한다. 즉, 실내 직육면체 공간 3면에 대한 점군이 관측된다면 나머지 면을 검출하여 모든 벽면과 바닥, 천장을 검출할 수 있다.

기존의 OBB 검출 기술은 영상 속 오브젝트의 불륨 검출을 위한 기술로 활용되었다. 이는 학습을 기반으로 초기 불륨을 생성하고 교정하는 검출 방법[14][15]이 대다수이며, 영역을 잘게 분할하고 최소영역에서 확장하며 탐색하는 bottom-up 방식의 불륨 검출 방법[16]이 대부분이다. 학습을 기반으로 하는 경우 많은 학습 데이터가 필요하며, 점군 이상의 시각적인 데이터를 요구한다. 시각적인 데이터를 위해 한 장의 영상 프레임 내에 오브젝트의 모습이 전부 보여야 정확한 불륨이 검출되기 때문에 한 장으로 답을 수 없는 실내 공간의 경우 검출이 보장되지 않는다. Bottom-up 방식의 불륨 검출은 소형 오브젝트에 적합하다. 모든 영역을 포괄하는 실내 공간에 대한 직육면체를 탐색하기에는 불륨을 확장하는 과정에서 빈 공간 등의 환경적인 제약 요소로 인해 확장이 멈출 수 있다. 실내 공간의 직육면체를 탐색하기 위해서는 전체 영역 불륨을 검출한 후 교정하는 top-down 방식으로 탐색하는 방법이 적합하다.

3. 직육면체 기반 AR 게임 공간 구성

AR 게임 환경 구성은 실제 세계를 기반으로 한 가상의 환경을 구축하고, 이를 실제 세계와 일치시키는 과정이 필요하다. 이를 위해서는 우선 현실 공간의 이해를 기반으로 한 가상공간의 구축이 필요하다. 제안하는 시스템은 현실 공간의 기하 정보를 검출하여 공간의 바닥면, 벽면, 천장면 등의 주요 정보를 획득한다.



[Fig. 2] The pipeline of our proposed cuboid detection process.

[Fig. 2]는 제안하는 실내 직육면체 불륨 검출 시스템의 흐름도이다. 제안하는 시스템은 3개의 주요 과정으로 구성되어 있다. 먼저, 구면 파노라마 기반의 Vision SLAM을 통해 실시간으로 점군 데이터를 획득하고, KDTree를 구성해 이웃점과 거리가 먼 이상점을 제거한다. 다음으로, 점군 데이터를 감싸는 최적의 OBB를 검출한다. 마지막으로, 생성된 OBB 불륨의 품질을 정량적으로 평가하고, 실시간으로 좋은 품질의 OBB로 생산한다. 각 세부 과정은 다음 절에서 자세하게 설명한다.

3.1 이상점이 제거된 점군 획득

제안하는 시스템은 OpenVSLAM을 활용하여 실시간으로 점군 데이터를 획득한다[17]. 해당 SLAM은 구면 파노라마 비디오를 입력으로 하여 실시간으로 영상 속의 특징점에 대한 3차원 점군과 카메라의 위치 및 자세 추정 정보를 생성한다. OpenVSLAM은 우수한 정확성을 보이지만, 출력 데이터인 점군과 카메라의 위치 및 자세 정보는 영상 분석을 통한 추정 데이터이기 때문에 오차를 포함한 비선형적인 데이터이다. 따라서 점군 데이터를 활용하기 위해서는 오차를 고려해야 한다.

향후 OBB를 검출할 때 점군 속 샘플의 미세한 오차는 큰 영향을 주지 않지만, 오차가 큰 샘플은 심각한 영향을 줄 수 있다. 따라서 본 시스템은 치명적인 샘플을 제거하기 위해 이웃점과의 거리가

지나치게 먼 이상점을 제거한다. 본 시스템은 이를 위해 KDTree를 구성한다. KDTree는 BST를 기반으로 하여 다중 차원에서 한 점을 중심으로 각 축들을 번갈아가며 주변의 점과의 거리에 따라 노드를 구성한 자료구조 트리이다. KDTree를 통해 한 점에서 가장 가까운 이웃점을 빠르게 검색할 수 있다.

우선, 본 시스템은 모든 점의 평균에 가장 가까운 점을 기준으로 KDTree를 구성한다. 그 다음, 모든 점에 대해 가장 가까운 이웃점과의 거리를 저장한 집합 $X = [x_1, \dots, x_n]^T$ 을 생성한다. 그리고 X 에 대해 평균값 $E(X)$ 와 표준편차 $\sigma(X)$ 을 계산한다. $\sigma(X)$ 은 각 점들에 대한 이웃점의 평균 거리를 중심으로 얼마나 퍼져 있는지를 의미하는 산포도의 일종이다. 그러므로 한 점에 대한 이웃점과 거리가 $E(X) + \sigma(X)$ 의 합 e 보다 멀다는 것은(식 (1) 참조) 다른 점들에 비해 지나치게 거리가 멀다는 것을 의미한다. 본 시스템에서는 X 에 대해 x_i 가 e 보다 큰 i 번째 점을 이상점으로 분류하였다.

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

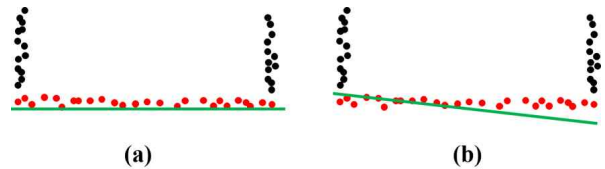
본 시스템의 OBB를 활용한 기하 공간 검출 방법은 실제 공간에 대한 Convex Hull을 획득할 수 있는 소량의 노이즈가 적은 점이 있는 것이 중요하다. 따라서 본 시스템에서는 많은 점들을 이상점으로 분류하여 밀도가 낮아지더라도 최대한 노이즈가 적은 점들만을 획득한다. 그러나 본 시스템은 이상점으로 분류된 샘플을 삭제하지 않는다. 이상점은 향후에 추가될 점군으로 인해 이웃과의 거리가 멀지 않은 활용 가능한 샘플로써 동작할 수 있다. 따라서 이상점은 현재 프레임에서만 이상점으로 분류하여 보관한다.

3.2 실내 환경의 직육면체 공간 검출

OBB는 점군의 영역 전체를 감싸는 최소의 박스

다. 노이즈가 없는 환경에서 직육면체 공간의 점군에 대한 OBB는 직육면체 공간의 표면을 정확하게 검출해 공간을 감싼다. 본 시스템은 Naujoks 등 [18]이 제안한 OBB 검출 알고리즘을 활용한다. 해당 알고리즘은 점군을 둘러싸는 가장 최적의 샘플 집합인 Convex Hull을 검출하고, Convex Hull의 극점을 활용한 박스 교정을 통해 높은 근사치의 박스를 생성한다. 해당 알고리즘은 방향 교정에서 단순히 가장 부피가 작은 박스를 지향하는 것이 아닌, 평면상에 많은 점을 포함할 수 있는 방향으로 교정한다. 따라서 직육면체 점군과 같이 모서리 지점이 분명한 점군에서는 많은 점을 포함한 면들에 밀착하여 박스를 형성하기 때문에 신뢰도가 높은 공간 검출이 가능하다.

본 시스템은 실시간으로 이전의 결과와 비교를 통해 점진적으로 우수한 실내 기하 공간 정보를 갱신한다. 이를 위해서는 OBB의 품질을 수치화하여 정량적으로 평가하는 과정이 필요하다. 제안 시스템은 박스의 6개의 평면에 대해 평면상에 존재하는 샘플의 수가 많을수록, 그리고 각 면으로부터 근접한 샘플들까지의 거리가 균일할수록 올바른 높은 품질 점수를 부여하였다.



[Fig. 3] Two difference situations to the uniformity of the distances from the plane(green line) to adjacent samples(red points) among the point cloud(black and red points).

[Fig. 3]은 평면으로부터 근접한 샘플들로부터 검출된 평면까지의 거리가 비교적 일정한 상황(a)과 일정하지 않은 상황(b)의 예시이다. 만약 (a)와 같이 평면으로부터 근접한 샘플의 거리가 일정하다면, 샘플의 거리 집합에 대한 표준편차가 낮다. 반면 평면의 방향이 올바르지 않아서 샘플의 거리가 일정하지 않은 (b)의 상황에서는 비교적 표준편차

가 높다. 그러므로 평면으로부터 근접한 샘플의 거리에 대한 표준편차를 통해서 평면의 방향과 거리가 비교적 정확한지 평가할 수 있다.

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

제안 시스템은 박스의 6개 평면에 대해 식 (2)와 같은 방법으로 품질 점수 s 을 산출하였다. 6개의 면에 대한 평면 주변 샘플들의 거리에 대한 집합 $[D_1, \dots, D_6]$ 가 존재할 때, i 번째 평면상의 점의 개수 n_i 을 D_i 의 표준편차로 나누어준 값을 해당 평면의 점수로 산출하였고, 6개의 평면에 대한 점수의 합을 OBB의 점수로 산출하였다.

$$B_{best} = \begin{cases} B_{cur} & \text{if } s_{cur} \geq s_{best} \\ B_{best} & \text{otherwise} \end{cases} \quad (3)$$

본 시스템에서는 매 프레임마다 OBB에 대한 점수를 산출하여 이전에 존재했던 가장 높은 품질의 OBB보다 점수가 높다면 OBB를 갱신한다. 반대의 경우 기존의 OBB 상태를 유지한다. 식 (3)과 같이 현재 프레임에서의 박스 B_{cur} 에 대한 품질 점수 s_{cur} 가 이전에 가장 좋은 품질의 박스 B_{best} 에 대한 품질 점수 s_{best} 보다 크거나 같다면, B_{best} 을 B_{cur} 로 업데이트한다. 만약 작다면, 현재의 박스는 B_{best} 보다 정확하지 않은 것으로 판단하고 소멸시킨다. 매 프레임마다 B_{best} 은 업데이트될 수도 있고, 아닐 수도 있다. 실질적으로 모든 프레임에 대해 가상의 공간을 구성하는 박스는 B_{best} 이다. AR 게임 공간에서 플레이어에게 제공되는 직육면체 공간 역시 B_{best} 을 기반으로 구축된다.

제안하는 시스템은 실시간 환경에서 이전에 관측된 점군을 포함한 상태로 지속적으로 점군을 확장한다. 새로 갱신하는 OBB는 이전의 OBB를 획득하기 위해 활용했던 모든 점군 정보를 포함하여

활용한다. 그러므로 새로 갱신하는 OBB는 이전의 OBB에 대한 모든 의미 있는 정보를 포함하고 있다. 따라서 본 시스템은 OBB를 갱신하는 과정에서 이전의 OBB 정보를 별도로 고려하지 않는다.

이 과정을 통해 정해진 OBB 데이터는 직육면체 형태의 직관적인 기하 모델이다. OBB 데이터는 8개의 꼭짓점으로 이루어져 있으며, 한 평면 상에 존재하는 4개의 꼭짓점을 분리하여 총 6개의 평면을 생성할 수 있다. 각 평면은 직육면체 형태의 실내 공간을 둘러싸는 바닥, 천장, 그리고 4개의 벽면에 대해 대응하는 평면의 정보를 담고 있다.

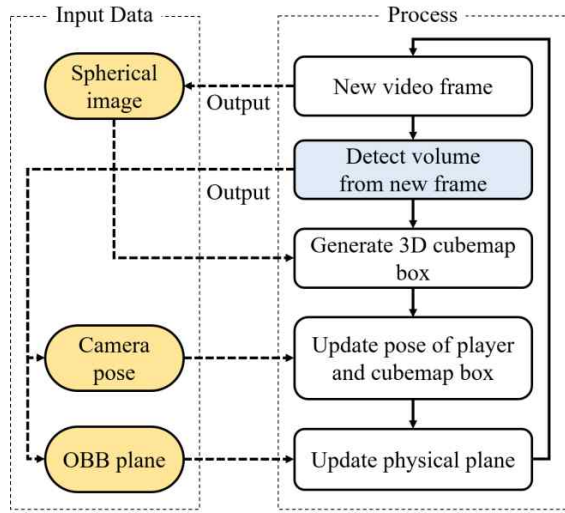
3.3 가상 세계 일치화와 AR 공간 구성

기하 공간 검출 과정을 통해 인식한 공간의 정보는 Vision SLAM을 통해 획득한 점군을 기반으로 구축하였기 때문에 좌표계와 스케일이 현실 세계와 다르다. 그러므로 AR 게임 환경을 구성하기 위해서는 인식한 공간의 정보와 현실 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 논문에서는 게임 엔진(Unreal Engine 4) 환경에서 가상 세계 정보와 실제 세계 정보의 일치화를 통해 AR 게임 환경을 구축하였다.

일반적인 핀홀 카메라로 촬영한 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.

[Fig. 4]는 가상 세계 구축과 일치화에 대한 흐름도이다. 본 시스템은 새로운 비디오 프레임이 입력되면 먼저 기하 공간 검출 과정을 수행하고, 이

후 가상 세계를 업데이트한다. 흐름도에서 파란색 프로세스는 3.1장과 3.2장의 기하 공간 검출 과정을 의미한다. 노란색 데이터 블록은 비디오 프레임 입력 과정과 기하 공간 검출 과정에서 획득할 수 있는 데이터이다. 기하 공간 검출 과정 이후에는 3개의 과정으로 나누어 수행하여 가상 세계를 구축하고 실제 세계와 일치화한다.



[Fig. 4] The pipeline of our augmented space construction and synchronization method.

기하 공간 검출 과정 이후에는 먼저 구면 파노라마의 왜곡을 제거하기 위해 3차원 큐브맵 이미지로 변환한다[19]. 큐브맵 이미지는 거대한 스카이박스처럼 가상 세계 안에 매우 큰 정육면체 형태로 배치한다. 해당 정육면체의 안쪽 면에는 큐브맵 이미지의 6면이 투영되어 가상 세계 전체를 감싸는 형태이다.

다음으로, 가상 세계 안에서의 카메라의 위치 및 자세를 플레이어와 큐브맵 정육면체에 반영한다. OpenVSLAM은 실시간으로 영상 속 특징점에 대한 3차원 점군과 카메라의 위치 및 자세 추정 정보를 생성한다. 따라서 가상 세계 안에서의 카메라의 위치 및 자세 정보는 OpenVSLAM을 통해 획득한 카메라 위치 및 자세 추정 정보이다. AR 콘텐츠는 1인칭 시점이기 때문에 플레이어의 위치 및 자세는 카메라의 위치 및 자세와 동일하다. 실

제 세계의 영상과 가상 세계가 일치되어야 하므로 가상 세계에 배치된 큐브맵 정육면체의 위치 및 자세 역시 카메라의 위치 및 자세로 변환한다.

마지막으로, 공간 검출 과정을 통해 획득한 OBB의 6개의 평면을 가상 세계에 배치한다. 해당 평면은 가상 세계 상에서 점군을 분석해 획득한 평면이기 때문에 별도의 위치 및 자세 변환을 하지 않는다.

3.4 게임 환경에서 가상 오브젝트 스폰

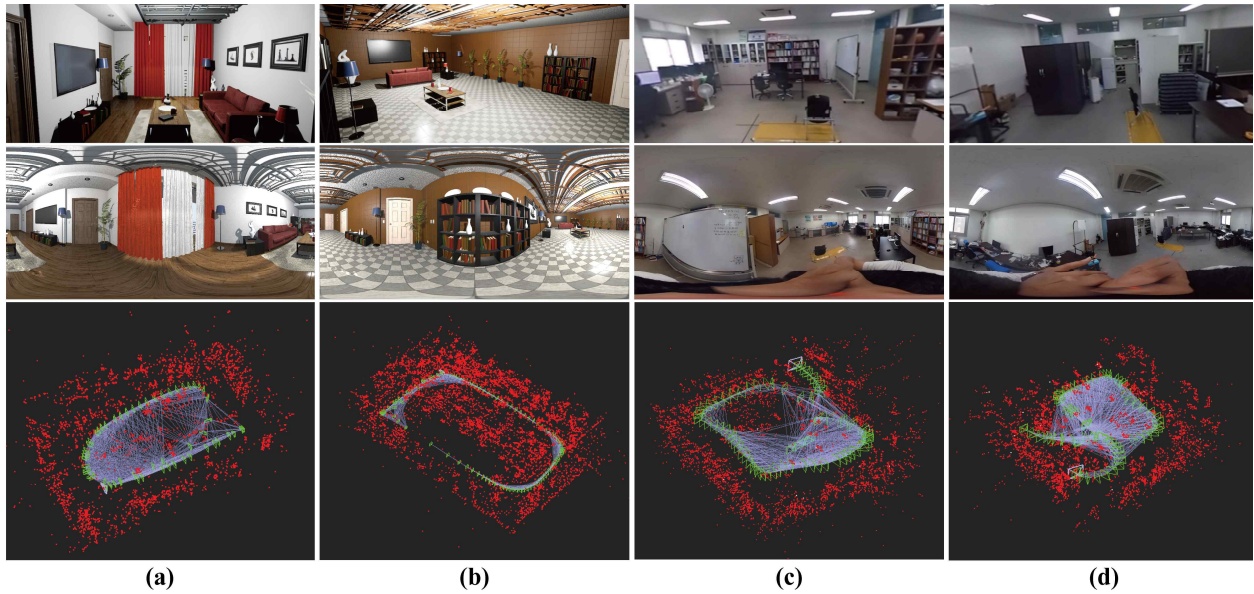
3차원 가상 오브젝트의 스폰(spawn)은 AR 콘텐츠의 가장 중요하면서도 기본적인 시작점이다. 공간 검출을 통해 획득한 평면을 활용하기 위해서는 각각의 평면에 대한 좌표계에서의 오브젝트를 생성해야 한다. 먼저, 오브젝트의 스폰 지점을 평면의 중앙 지점이라 가정할 때 오브젝트의 위치 및 자세는 식 (4)~(5)와 같다.

$$\mathbf{r}_n = \begin{bmatrix} r_R \\ r_P \\ r_Y \end{bmatrix} = \begin{bmatrix} 0 \\ \arctan(z_n / \sqrt{x_n^2 + y_n^2}) \\ \arctan(y_n / x_n) \end{bmatrix} \quad (4)$$

$$P_o = [R_o | \mathbf{t}_o] = [Mat(\mathbf{r}_n) | \mathbf{p}_c] \quad (5)$$

오브젝트가 생성된 후 평면의 법선 벡터와 같은 자세로 회전시키고, 평면의 중앙 지점으로 이동시키면 해당 평면의 중앙에 밀착되어 평면상에 존재하게 된다. 평면의 법선 벡터 $\vec{v}_n = [x_n, y_n, z_n]^T$ 의 자세(회전량) \mathbf{r}_n 은 식 (4)와 같은 방법을 통해 획득할 수 있다. 오브젝트의 초기 위치 및 자세 P_o 은 식 (5)와 같이 \mathbf{r}_n 의 회전행렬 $Mat(\mathbf{r}_n)$ 와 평면의 중앙 지점 위치 \mathbf{p}_c 을 통해 표현된다.

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



[Fig. 5] The datasets for quality measurement: (upper) snapshot of test environment, (middle) spherical panorama of the space, (lower) spherical video-based point cloud generated via OpenVSLAM. The two datasets (a),(b) are for virtual space, and (c),(d) are for real space.

으로 이동시킨다.

4. 실험 결과

4.1 공간 인식의 성능 평가

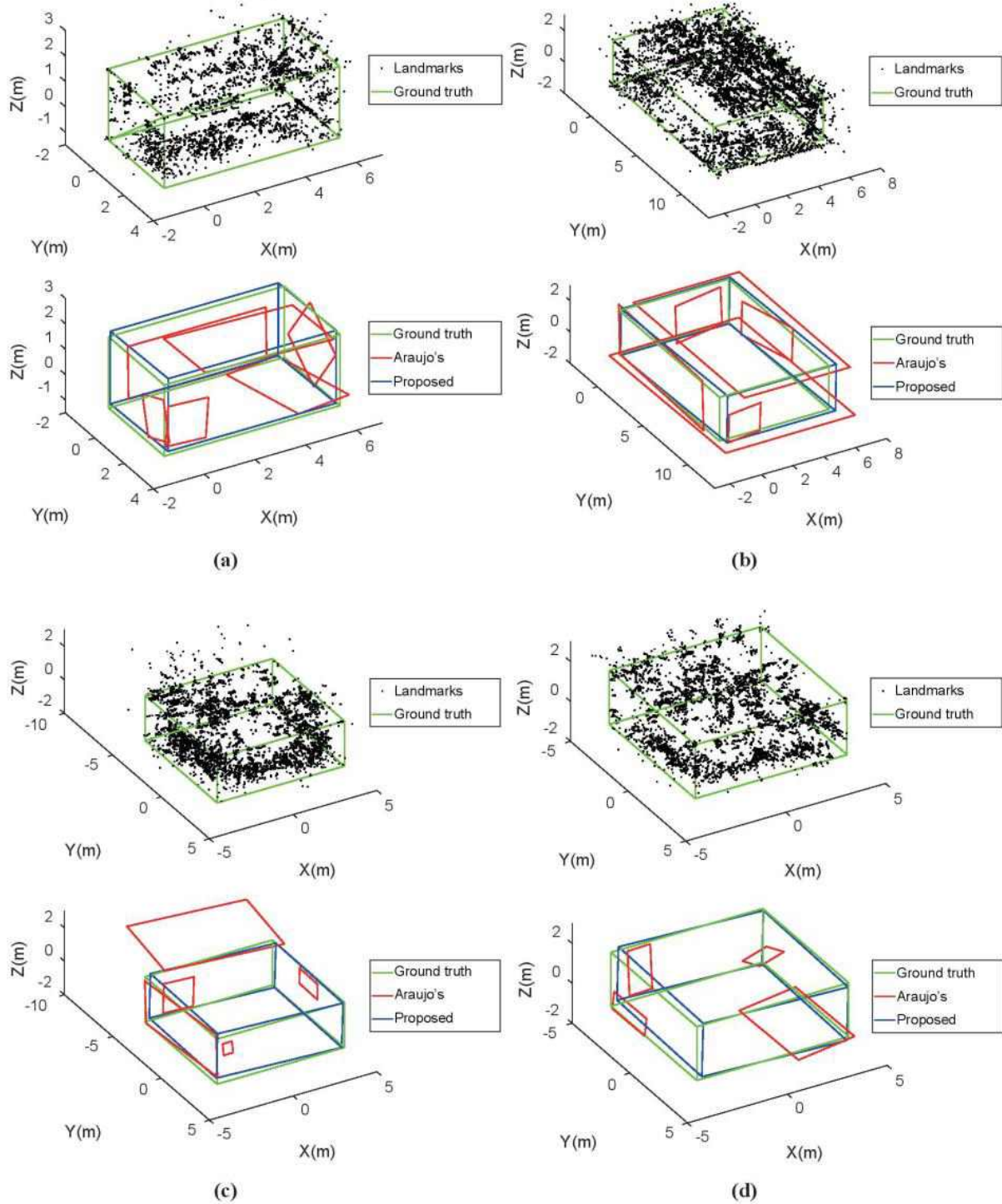
공간 인식의 성능을 평가하기 위해 기존의 다중 평면 검출 알고리즘 중 최근에 제안되고 좋은 성능을 보인 Araujo 등이 제안한 시스템과 비교를 통해 실험을 수행하였다. 실험을 위한 입력 데이터는 점군이며, 이는 구면 파노라마 비디오를 OpenVSLAM을 통해 분석한 출력 데이터이다. 실험은 3.40GHz Intel Core 7-6700 CPU와 32.0GB RAM의 사양을 가지는 Desktop PC를 통해 수행했으며, Windows 운영체제 환경에서 수행하였다. 실험 속 공간 인식을 위한 연산은 GPU를 사용하지 않았으며, CPU를 통해 수행하였다. 실험 속 수행 시간 측정은 영상을 불러오는 과정을 제외하고, 점군을 입력으로 하여 평면 검출을 마친 시점까지 시간을 측정하였다.

실험은 [Fig. 5]와 같이 4개의 데이터를 분석하

여 수행하였다. 그림에서 (a),(b)는 가상 세계에 대한 데이터이고, (c),(d)는 실제 세계에 대한 데이터이다. 실제 세계를 기반으로 구면 파노라마 비디오를 촬영할 경우 비디오 스탠드나 촬영자가 등장하여 점군의 신뢰성이 소폭 하락한다. 이를 위해 실험을 가상 세계와 실제 세계에서 모두 수행하였다. 가상 세계의 경우 완벽하게 촬영자가 등장하지 않기 때문에 이상적인 데이터를 기반으로 실험을 수행할 수 있다. 실제 세계의 경우 실제 사용 환경과 동일하기 때문에 환경 제약 요소가 적절히 반영된 환경에서 실질적인 성능을 분석할 수 있다. 가상 세계에 대한 비디오는 Unreal Engine 4 상에서 실제 세계와 거의 유사한 극 실사 환경을 구성하여 구면 파노라마 비디오를 녹화하였다.

제안 시스템은 비디오 프레임을 입력으로 하여 매 프레임마다 확장되는 점군을 기반으로 매 프레임 공간을 검출하여 점점 좋은 품질의 공간 데이터로 갱신한다. 그러나 실험의 공정성을 위해 두 시스템 모두 완전히 관측이 완료된 실내 점군 데이터를 대상으로 1회만 검출하여 비교하였다.

실험은 Araujo의 시스템과 제안 시스템의 결과



[Fig. 6] Comparison of the detection results from the Araujo's system and our proposed system: (upper) point cloud and Ground truth, (lower) ground truth, Araujo's system and our proposed system.

물을 Ground truth와 비교하여 수행하였다. 비교 결과가 [Fig. 6]에 있다. Ground truth는 가상공간의 바닥, 천장, 4개의 벽면에 대한 평면이다. 가상세계의 Ground truth는 가상 세계 속의 평면의 위치, 크기, 회전량을 알아내고, OpenVSLAM상의 카메라 경로와 가상 세계 속 카메라의 경로의 이동거리 비율을 통해 정확히 스케일을 일치시켜 획득했다. 실제 세계의 Ground truth는 촬영되는 공간의 너비, 폭, 높이를 정확히 측정하고, 공간의 중앙에 촬영 장치를 일정한 거리만큼 움직이는 슬라이드를 설치하여 OpenVSLAM상의 카메라 경로와 실제 세계의 카메라 경로의 이동거리 비율을 통해 좌표계와 스케일을 정확히 일치시켜 획득했다.

[Table 1] Comparison of the Araujo's method and the proposed method.

	Type	Method	N	E(cm)	T(s)
(a)	Virtual	Araujo's	6	43.47	89.421
		Proposed	6	4.80	0.078
(c)	Virtual	Araujo's	6	26.39	98.332
		Proposed	6	4.41	0.084
(e)	Real	Araujo's	5	58.62	91.488
		Proposed	6	4.66	0.080
(d)	Real	Araujo's	4	64.47	90.192
		Proposed	6	5.12	0.079

평면 검출 성능 비교를 위해 [Table 1]과 같이 두 시스템에서 검출된 평면의 개수(N), 검출된 평면에 대해 Ground truth와의 평균 오차(E), 수행 시간(T)을 구하여 비교하였다. E 은 Ground truth 평면의 네 꼭짓점과의 거리 오차를 1미터 당 오차(단위 cm)로 산출했다. N 은 Ground truth 평면의 회전값과 코사인 유사도가 0.9 이상이고, 평면상에

존재하거나 평면 가까이에 존재하는 평면이 있을 경우 검출된 것으로 가정하였다. Araujo의 방법의 경우 한 Ground truth 평면을 두 개 이상의 평면으로 검출할 수도 있다. 이럴 경우, 두 평면의 영역을 합치고 회전값의 평균을 반영하였다.

두 시스템은 5개의 데이터에 대해 거의 모든 일부 이상의 평면을 검출하였다. 그러나 [Fig. 6]에서 볼 수 있듯이, Araujo의 시스템은 대부분의 평면에 대해 일부 영역만 검출한 반면, 제안 시스템은 매우 정확한 회전상태와 크기로 모든 평면을 검출하였다. 제안 시스템은 1미터 당 평균 4.74cm의 오차를 보인 반면, Araujo의 시스템은 48.23cm의 오차를 보여 약 10배가량의 정확도 차이를 보였다.

제안 시스템이 모든 데이터에 대해 거의 일관된 E 을 보인 반면, Araujo의 시스템은 E 가 매우 불규칙하다. 이는 Araujo의 시스템이 점군의 밀도의 변화에 매우 불안정하기 때문이다. 가상공간의 경우 [Fig. 6] 속 각 데이터의 위쪽 그래프(점군 + Ground truth 그래프)에서 볼 수 있듯이, 데이터 (a)에 대한 가상공간은 무늬가 단조로운 반면 데이터 (b)에 대한 가상공간은 패턴 무늬의 바닥과 벽면, 천장을 가져 점군의 밀도가 비교적 균일하다. Araujo의 시스템은 이러한 점군의 밀도에 많은 영향을 받는 결과를 보인다. 또한 Araujo의 시스템은 밀도가 낮고 밀도 변화가 심한 데이터일수록 평면이 확장되지 못하고 일부의 영역만 검출된 것을 볼 수 있다.

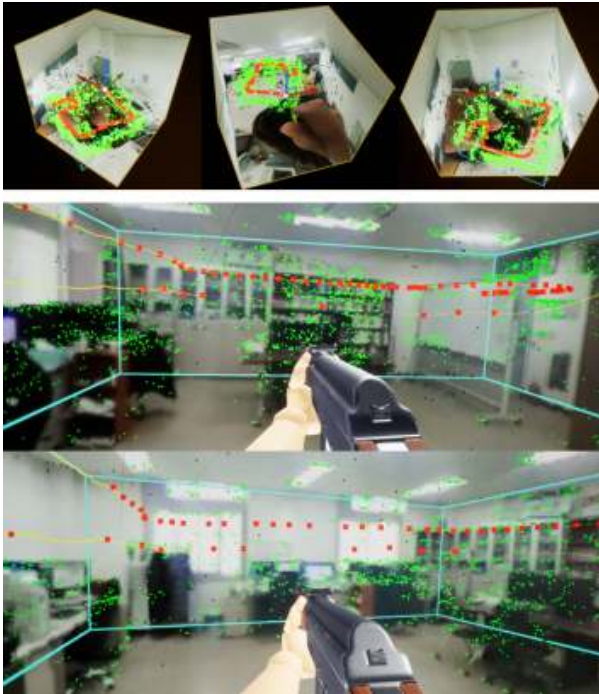
Araujo의 시스템은 실제 세계에 대한 데이터 (c),(d)에서 더욱 높은 E 을 보였다. 이는 Araujo의 시스템이 실제 세계에서 발생하는 환경 제약 요소에 많은 영향을 받는다는 것을 의미한다. 반면, 제안 시스템은 실제 세계에 대해서도 비교적 균일한 E 을 보였다. 이는 환경 제약 요소가 많은 실제 사용 환경에서도 우수한 성능을 보인다는 것을 의미한다.

두 시스템은 수행 시간에서도 많은 차이를 보였다. Araujo의 시스템은 평균 92.3초의 수행 시간을 가졌던 반면, 제안 시스템은 평균 0.08초의 수행 시

간을 보여 제안 시스템이 1000배 이상 빠른 수행 속도를 보였다. 제안 시스템은 연속적인 스트리밍 환경에서도 실시간성을 확보할 수 있다.

4.2 구현된 AR 게임 환경 평가

제안 방법의 일치화 과정과 오브젝트 스폰 과정을 통해 사용자의 시점에서 정확하게 일치화된 가상 세계를 확인할 수 있다. 또한, 오브젝트의 스폰을 통해 인식된 공간의 활용성을 검토할 수 있다. [Fig. 7]의 상단 그림은 일치화 과정을 통해 구축된 가상 세계의 모습이다. 가상 월드 속에는 점군, 카메라 경로, 그리고 구면 파노라마 프레임의 큐브맵 정육면체가 생성되어 있다. 큐브맵 정육면체는 훨씬 거대하지만, 확인을 위해 그림에서만 크기를 축소하였다. 하단 그림은 사용자의 1인칭 시점에서 실제 세계와 가상 세계가 일치된 모습이다.



[Fig. 7] The implemented AR game environment: (upper) virtual world, (lower) synchronized player's point of view.



[Fig. 8] The spawned 3D objects in the constructed AR game space.

[Fig. 7]의 하단 그림에는 실제 세계의 영상을 비롯해 Vision SLAM을 통해 획득한 점군, 카메라 이동 경로, 공간 검출을 통해 획득한 OBB의 6개 평면 정보가 모두 출력되어 있다. 그림에서 볼 수 있듯이, OBB를 구성하는 선분은 실내 직육면체 공간의 모서리를 정확하게 찾아내어 일치되어 있다.

구현된 AR 게임 응용에서는 공간 검출을 통해 획득한 평면상에 오브젝트를 생성하여 실제 세계의 바닥면, 천장면, 벽면에 밀착하여 실제 세계와 일치화되어 동작되도록 하였다. [Fig. 8]에 구현된 AR 게임 응용의 장면이 있다. 제안 시스템에서 생성한 평면들은 직육면체 공간에서 획득할 수 있는 가장 거대한 평면들이기 때문에 많은 영역을 활용하는 콘텐츠를 생성할 수 있었다.

5. 결 론

본 논문에서는 직육면체 형태의 실내 공간에서 공간 기하 정보 검출을 통해 AR 게임 공간을 구성하는 방법을 제안하였다. 제안 시스템은 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. 제안 방법은 전체의 영역을 한 번에 감싸는 top-down 방식을 활용해 점군의 밀도에 영향을 받지 않고, 전처리 과정이 필요 없다는 점에서 우수성이 있다. 실험 결과 기존의 다중 평면 검출 방법과 비교해 밀도와 노이즈의 변화에 매우 강건하며, 수행 속도가 월등히 빨라 실시간성을 확보함을 확인할 수 있었다.

제안 방법은 실내 AR 콘텐츠가 사용되는 공간의 특성을 반영하여 논문에서 가정하는 모든 직육면체 형태의 공간에서 활용할 수 있다. 그러나 사용 환경을 직육면체형 공간에 제한하기 때문에, 소수의 다른 형태 공간이나 실외 공간에 대해서는 활용할 수 없다는 한계점이 있다. 향후에 제안 방법을 확장하여 직육면체 형태가 아닌 다른 형태의 공간에 대한 모양을 추정할 수 있도록 하면 활용 가능한 환경이 확장될 것이다. 또한, 바닥면, 천장면, 벽면 등의 6면과 더불어, 기존의 다중 평면 인식 방법을 통해 부분적인 평면 검출을 수행한다면 더욱 융통성있는 AR 게임 공간을 구축할 수 있을 것이다.

ACKNOWLEDGMENTS

This work was supported by Incheon National University Research Grant in 2021.

REFERENCES

- [1] J. O'Rourke, "Finding Minimal Enclosing Boxes," *Int. J. Parallel Programming*, Vol. 14, No. 3, 1985.
- [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] A. S. Hassanein, S. Mohammad, M. Sameer, M. E. Ragab, "A Survey on Hough Transform, Theory, Techniques and Applications", *International Journal of Computer Science*, Vol. 12, Issue 1, pp. 1-18., 2015.
- [4] S. A. Hojjatoleslami, J. Kittler, "Region Growing: A New Approach," *IEEE Trans. Image Processing*, Vol. 7, Issue 7, pp. 1079-1084, 1998.
- [5] R. Schnabel, R. Wahl, R. Klein, "Efficient RANSAC for Point-Cloud Shape Detection", *Computer Graphics Forum*, Vol. 26, No. 2, pp. 214-226, 2007.
- [6] U. Bauer, K. Polthier, "Detection of Planar Regions in Volume Data for Topology Optimization", *Lecture Notes in Computer Science* 4975, pp. 119-126, 2008.
- [7] J. Poppinga, N. Vaskevicius, A. Birk, K. Pathak, "Fast Plane Detection and Polygonalization in noisy 3D Range Images", In *IROS '08*, pp. 3378-3383, 2008.
- [8] Abner M. C. Araujo, Manuel M. Oliveira, "A Robust Statistics Approach for Plane Detection in Unorganized Point Clouds," *Pattern Recognition*, Vol. 100, pp. 1-12, 2020.
- [9] 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.*, Vol. 25, Issue 1, pp. 86-97, 2014.
- [10] 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*, Vol. 9, No. 5, 2017.
- [11] P. Biber, W. Strasser, "The Normal

- Distributions Transform: A New Approach to Laser Scan Matching”, in IEEE/RJS Intern. Conf. on Intelligent Robots and Systems, Vol. 3, pp. 2743-2748, 2003.
- [12] R. V. Chadnov, A. V. Skvortsov, “Convex hull algorithms review”, In Proc. 8th Russian-Korean Int. Symp. on Science and Technology, Vol. 2, pp. 112 - 115, 2004.
- [13] N. Tayem, H. M. Kwon, “L-Shape 2-Dimensional Arrival Angle Estimation with Propagator Method”, IEEE Trans. Antennas Propag., Vol. 53, No. 5, pp. 1622-1630, 2005.
- [14] D. Dwibedi, T. Malisiewicz, V. Badrinarayanan, A. Rabinovich, “Deep Cuboid Detection: Beyond 2D Bounding Boxes”, arXiv:1611.10010, 2006.
- [15] S. Fidler, S. Dickinson, R. Urtasun, “3D Object Detection and Viewpoint Estimation with a Deformable 3D Cuboid Model”, In NIPS, 2012.
- [16] Z. S. Hashemifar, K. W. Lee, N. Napp, and K. Dantu, “Consistent Cuboid Detection for Semantic Mapping”, In Proc. IEEE 11th ICSC, 2017.
- [17] S. Sumikura, M. Shibuya, K. Sakurada, “OpenVSLAM: A Versatile Visual SLAM Framework,” In Proc. 27th ACM Int. Conf. Mult., pp. 2292-2295, 2019.
- [18] 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.
- [19] T. El-Ganainy, M. Hefeeda, “Streaming virtual reality content”, arXiv:1612.08350, 2016.



김기식 (Kim, Ki-Sik)

약력 : 2015-2020 인천대학교 임베디드시스템공학과(공학사)
2020-현재 인천대학교 컴퓨터공학과(석사과정)

관심분야 : 게임개발, 게임엔진, AR/VR, 게임인공지능



박종승 (Park, Jong-Seung)

약력 : 1988-1992년 경북대학교 전자계산학과(이학사)
1992-1994년 POSTECH 컴퓨터공학과(공학석사)
1994-1999년 POSTECH 컴퓨터공학과(공학박사)
2004년-현재 인천대학교 컴퓨터공학부 교수

관심분야 : 게임제작, 게임프로그래밍, 증강현실

