

가상 환경의 배경 정보를 위한 2D 영상 기반의 3D 데이터 생성 방법

이성배, *김규현
경희대학교

rhee@khu.ac.kr, *kyuheonkim@khu.ac.kr

2D Image based 3D Data Generation Method for Background
Information of Virtual RealitySeongbae Rhee, *Kyuheon Kim
Kyunghee University

요약

가상 현실(VR: Virtual Reality) 기술은 대표적인 몰입형 미디어 기술로써, 컴퓨터 등을 통해 인공적으로 만들어낸 실제와 유사한 특정 환경, 상황 혹은 그 기술 자체를 의미한다. VR 기술은 비교적 간단한 장비를 착용한 것만으로 가상 세계에 구현된 모든 공간을 체험할 수 있기 때문에 사용자가 신체적 불편하더라도 손쉽게 유명 관광 명소를 여행할 수 있다. 또한, 실제 작전 지역을 가상 세계에 반영함으로써, 가상 세계에서 안전한 군사 훈련이 가능하다. 이와 같은 활용을 가능하게 하기 위해서는 가상 세계의 배경 정보가 실제 세계의 모습과 흡사한 실사 그래픽으로 구성되어야 한다. 그러나 실사 그래픽을 제작하는 것은 제작 난이도가 높고 제작비용이 비싸다는 제한 사항으로 인하여 실사 그래픽을 바탕으로 한 VR 콘텐츠의 수는 부족하다. 이에 본 논문에서는 일반 카메라를 통해서 촬영한 단일 영상 또는 다시점 영상을 통해서 Point Cloud 데이터를 생성하고, 이를 가상 세계의 배경 정보로 활용하기 위한 방법을 제안하고자 한다.

1. 서론

가상 현실(VR: Virtual Reality) 기술은 컴퓨터 등을 통해 인공적으로 만들어낸 실제와 유사한 특정 환경, 상황 혹은 그 기술 자체를 의미한다[1]. VR은 대표적인 몰입형(Immersive) 미디어 기술로써, VR 사용자가 콘텐츠에 몰입하는 것을 가능하게 하며, 가상 세계에서 실제 세계와 유사한 현장감을 느낄 수 있다는 장점이 있다.

VR이 사용자에게 현장감을 제공하기 위해서는 기본적으로 가상 세계와 사용자 상호 간에 데이터 정보의 교환이 이루어져야 하며, 이를 위하여 다양한 디바이스가 활용되고 있다. 이때, 동작, 음성, 조작 등의 사용자 정보를 가상 세계에 전달하기 위해 카메라 센서, 음성 센서, 트래커(Tracker), 조이스틱, 키보드, 마우스 등이 활용되고 있으며, 가상 세계의 데이터를 사용자에게 전달하기 위하여 HMD(Head Mounted Display), 스피커, 진동 센서 등이 활용될 수 있다[2].

상기와 같은 디바이스를 착용하는 것만으로 사용자는 가상 세계에 구현된 모든 공간을 갈 수 있고 체험할 수 있기 때문에 다양한 분야에서 VR의 활용이 기대되고 있다. 일례로, VR을 통해서 관광 서비스를 하는 경우, 실제로 여행을 하는 것보다 적은 비용으로 관광 명소에 방문할 수 있고, 현실에서 신체적 불편함으로 여행이 제한되는 사용자까지도 디바이스만을 착용함으로써 관광이 가능해진다. 또 다른 예시로는, 실제 작전 지역을 가상 세계에 반영함으로써, 가상 세계에서 안전한 군사 훈련을 진행할 수 있다[3].

일반적으로, 가상 세계에 존재하는 배경 데이터는 그림같이 표현되는 캐주얼 그래픽, 실제 모습과 유사한 유사(반실사) 그래픽, 실제 세계

와 구분하기 어려운 실사 그래픽으로 분류되는데, 앞선 예시에서 언급한 서비스들이 가능하기 위해서는 실사 그래픽으로 제작된 콘텐츠가 필요하다[4]. 그러나 실사 그래픽으로 콘텐츠를 제작하는 것은 제작 난이도가 높고, 제작비용이 비싸다는 제한 사항이 있기 때문에 실사 그래픽을 기반으로 한 콘텐츠의 수는 부족하다.

이에 본 논문에서는 실사 그래픽 기반의 배경 정보를 제작하는 것이 제작 난이도가 높고, 제작비용이 비싸다는 제한 사항을 극복하기 위하여 일반 카메라를 통해서 촬영된 영상을 바탕으로 포인트 클라우드(Point cloud)[5] 데이터를 생성하여 가상 환경에 맵핑함으로써, 보다 쉬운 실사 그래픽 제작 방법을 제안하고자 한다. 이와 같은 방법을 지원하기 위하여 본 논문의 2장에서는 배경 기술을 간단하게 분석하고, 3장에서는 일반 영상을 통해서 포인트 클라우드 데이터를 생성하여 가상 세계에 맵핑하는 방법을 설명한다. 이어지는 4장에서는 테스트 영상을 통해서 제안 기술의 결과를 설명하며, 5장에서는 결론을 내린다.

2. 배경 기술 분석

앞서 설명한 바와 같이 가상 현실 기술에서 활용하는 배경 정보는 그림 1에서 나타나는 바와 같이 캐주얼 그래픽, 유사(반실사) 그래픽, 실사 그래픽으로 구분할 수 있다. 이때, 사용자의 완전한 몰입을 위해서라면 가상 세계의 배경 정보는 실사 그래픽을 바탕으로 제작되어야 한다.

실사 그래픽은 3D 모델링, 맵핑(Mapping), 지형 작업, 색채 및 휘도 설정 작업 등의 과정으로 생성되며[6], 이때 배경 정보가 실제 환경과 구분되지 않게 제작되기 위해서는 3D 모델링과 색채 및 조도 설정 작업



(a) 캐주얼 그래픽



(b) 반실사 그래픽



(c) 실사 그래픽

그림 1. 그래픽 종류 예시

이 중요하다. 그러나 3D 모델링은 대리석을 조각하는 것과 같이 시각 블록에서 원하는 모양으로 성형하는 과정을 의미하며, 색채 및 조도 설정은 모델링된 형태에 실제와 같은 색을 입히고 빛이 반사되는 것과 그늘이 지는 것까지를 설정하는 것을 의미한다. 이와 같은 작업들은 수작업으로 진행되기 때문에 개발자의 역량에 따라서 결과의 품질이 크게 좌우되며, 해당 작업에는 섬세한 표현이 필요하기 때문에 작업에 소요되는 시간이 길다는 제한 사항이 존재한다.

이와 같이 실사 배경 데이터를 구현함에 있어서 3D 모델링과 색채 및 조도 설정의 제한 사항을 극복하기 위하여, 포인트 클라우드(Point cloud) 기술을 활용한 가상 환경의 배경 정보를 생성하는 방법을 이어지는 3장에서 제안하고자 한다.

3. 제안 기술명

본 논문에서는 포인트 클라우드 데이터를 활용하여 가상 환경의 배경 정보를 생성하는 기술에 대하여 설명하고 있다. 구체적으로, 실사 그래픽을 생성하는 과정 중에서 3D 모델링과 색채 및 조도 설정의 어려움을 해소하기 위하여 포인트 클라우드 데이터를 활용하여 가상 환경의 배경 정보를 생성하는 기술을 제안하고 있다.

일반적으로, 포인트 클라우드 데이터를 생성하기 위해서는 라이다 센서나 다수의 스테레오 카메라가 필요하기 때문에, 해당 장비를 갖추기 힘든 일반인이 포인트 클라우드 데이터를 생성하는 것은 제한될 수 있다. 이에 본 논문에서는 그림 2에서 나타나는 바와 같이 스테레오 카메라가 아닌 일반 카메라로 배경을 촬영하고, MegaDepth[7], MiDaS[8] 등과 같은 딥러닝 기반의 Depth estimation 기술을 활용하여 영상 한 장으로부터 depth 정보를 추출하여 포인트 클라우드 데이터를 생성하는 방법을 제안한다.

그림 2에서 나타나는 바와 같이 한 장의 영상으로 포인트 클라우드를 생성하는 과정과 다수의 시점에서 촬영된 영상으로 포인트 클라우드를 생성하는 과정으로 구분되며, 이때 다수의 시점에서 촬영된 영상으로 포인트 클라우드 데이터를 생성하는 경우에는 Registration 기술을 통해서 하나의 공간으로 정합하는 과정이 추가된다. 배경 정보를 담은 포인트 클라우드 데이터는 가상 세계에서 원하는 크기로 맵핑되기 위하여 scaling을 조절한다. 이때 크기가 조절되는 과정에서 점과 점 사이에 빈 공간이 생성되기 때문에 Upsampling과 Interpolation 과정이 필요하다. 앞선 과정이 진행된 이후, 포인트 클라우드 데이터는 가상 세계의 정해진 공간에 mapping되고 사용자의 니즈에 따라서 사물의 모양이 수정

되거나 채색과 휘도에 대한 추가적인 설정을 진행하여 3D 배경 정보의 생성은 마무리된다.

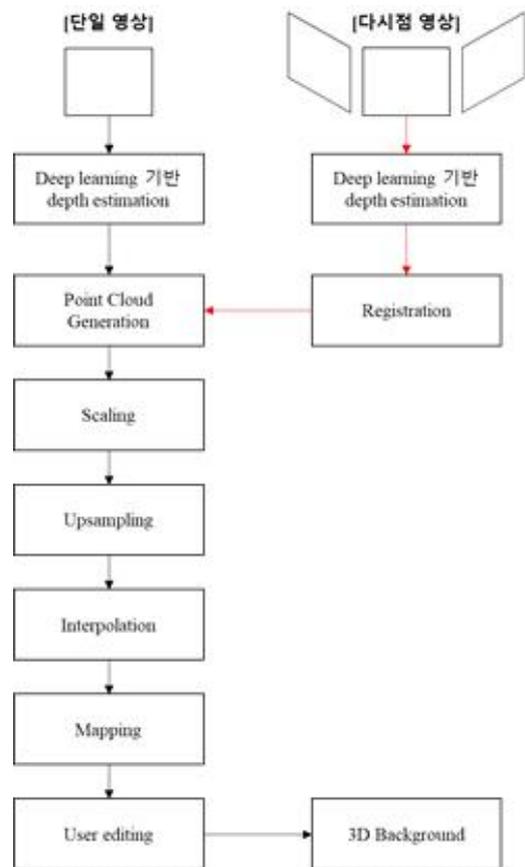


그림 2. 일반 영상을 통한 3D 배경 정보 생성 구조도

이와 같이 일반 카메라를 통해서 획득한 2D 영상을 활용하여 포인트 클라우드 데이터를 생성하고 3D 모델링의 기초 틀로 활용하는 것은 그래픽 제작자가 실사 그래픽을 제작하는 과정에서 스케치 역할을 수행하여 제작 난이도를 줄일 수 있을 것으로 기대된다. 이에, 이어지는 4장에서는 스테레오 카메라를 통해서 촬영된 영상이 아닌, 일반 카메라를 통해서 촬영된 일반 2D 영상 한 장을 통해서 딥러닝 기반 Depth estimation 기술을 적용하여 포인트 클라우드 데이터로 변환하는 과정을 테스트 영상을 통하여 진행하고자 한다.

4. 실험 결과

본 논문에서는 일반 카메라를 통해 획득한 2D 영상 한 장을 딥러닝 기반 Depth estimation 기술에 적용하여 depth 정보를 추출하고, 이를 포인트 클라우드 데이터를 생성하여 VR의 배경 데이터로 활용하는 방법을 제안하고 있다. 본 논문의 실험 환경은 Window 10, i7-10700의 CPU, RTX-2080의 GPU, Python 3.7, Google Colaboratory을 통해 설계되었다.

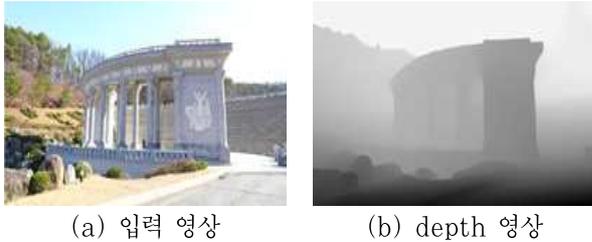


그림 3. 딥러닝 기반 depth estimation 결과 예시

본 논문에서 활용한 2D 영상은 그림 3의 (a)이며, 실험에서 활용한 딥러닝 기반 Depth estimation 모델은 MiDaS[8]이다. 딥러닝 기반 Depth estimation 모델인 MiDaS의 추출 값은 실수 값으로 추출되며, Open3D 라이브러리에 필요로 하는 depth 맵과 거리 값이 반대로 정의되어 있기 때문에 그림 3의 (a)를 입력으로 추출된 depth 맵에 대하여 식 1과 식 2의 변환 과정을 통해 그림 3의 (b)를 생성하였다.

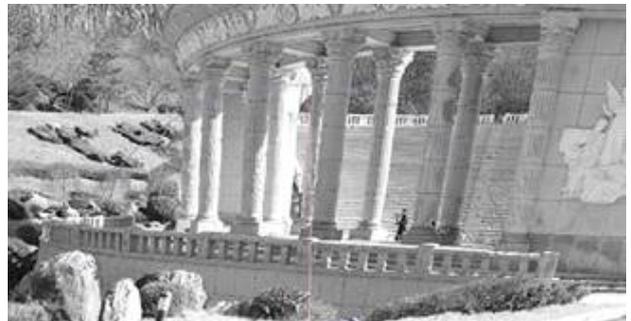
$$D_{norm} = \frac{D(x,y) - \min(D)}{\max(D) - \min(D)} \times 255 \quad (1)$$

$$D_{ref} = |255 - D_{norm}| \quad (2)$$

상기 식 1에서는 딥러닝 기반 Depth estimation에서 추출된 depth 정보를 0부터 255까지의 값으로 양자화를 하는 과정에 해당하며, 식 2는 Open3D 라이브러리에서 활용하는 depth 맵의 형태를 맞추기 위하여 중앙값을 기준으로 대칭 변환한 값에 해당한다.



(a) 정면 모습



(b) 30도 측면 모습

그림 4. 생성된 포인트 클라우드 결과 예시

입력 영상인 그림 3의 (a)와 딥러닝 기반 Depth estimation 네트워크를 통해 추출한 depth 맵인 그림 3의 (b)를 활용하여 Open3D 라이브러리를 통해서 포인트 클라우드 데이터를 생성한 결과는 그림 4와 같다. 그림 4의 (a)는 정면에서 보는 포인트 클라우드 배경의 모습이고, 그림 4의 (b)는 측면 30도 방향에서 보는 포인트 클라우드 배경의 모습이다. 해당 결과는 한 장의 영상만을 사용하여 생성한 포인트 클라우드 데이터로, 더 많은 입력 영상을 활용하는 경우 더 넓은 범위의 포인트 클라우드 배경 데이터를 생성할 수 있을 것으로 기대된다.

5. 결론

VR은 대표적인 몰입형(Immersive) 미디어 기술로써, VR 사용자는 콘텐츠에서 현장감을 느낄 수 있기에 관광 서비스, 군사 서비스 등 다양한 분야에서의 활용이 기대되고 있다. 이와 같은 서비스가 가능하기 위해서는 가상 세계의 배경 정보가 실사 그래픽으로 구성되어야 하지만, 실사 그래픽을 제작하는 난이도가 높고, 제작비용이 비싸다는 제한 사항으로 인해 실사 그래픽 기반의 콘텐츠의 수는 상당히 부족한 상황이다.

이에 본 논문에서는 실사 그래픽을 제작의 제한 사항을 극복하기 위하여 일반 영상을 딥러닝 기반 Depth estimation 기술에 적용함으로써, depth 정보를 생성하고 이를 포인트 클라우드 데이터로 변환하는 방법에 대하여 설명하였다. 또한, 실험에서 테스트 영상으로 포인트 클라우드 데이터를 생성함으로써, 제안하는 기술이 가상 세계의 배경 정보로 활용 가능함을 확인하였다.

이와 같이 일반 카메라를 통해서 획득한 2D 영상을 활용하여 포인트 클라우드 데이터를 생성하고, 3D 모델링의 스케치 역할로 활용하는 것은 제작 난이도를 줄일 수 있을 것으로 기대된다. 또한, 포인트 클라우드 데이터를 Mesh 단위로 처리함으로써, 캐주얼 그래픽 또는 반실사 그래픽 제작에도 활용될 수 있을 것으로 기대된다.

추가로, 다시점에서 촬영된 영상을 본 논문에서 제안한 절차로 더 넓은 공간의 포인트 클라우드 데이터를 생성하는 것과 맵스 정보의 양자화를 고도화하여 세밀한 표현까지 구현하는 등의 다양한 추가 연구가 가능할 것으로 기대된다.

참고문헌

- [1] Anthes, Christoph, et al. "State of the art of virtual reality technology." 2016 IEEE Aerospace Conference. IEEE, 2016.
- [2] BOAS, Y. A. G. V. Overview of virtual reality technologies. In: Interactive Multimedia Conference. 2013.
- [3] Lee, Eun-Seok, et al. "An Efficient Acceleration Method for Terrain Rendering in Virtual Reality Environment." 한국차세대컴퓨팅학회 논문지, Vol.17, No.1, pp. 83-92, 2021.
- [4] Lee, Yo-Seop. "2D-3D 변환 기술의 동향 및 전망." The Magazine of the IEIE Vol.38, No.2, pp.37-43, 2011
- [5] Kim, Junsik, et al. "3D motion estimation and compensation method for video-based point cloud compression." IEEE Access 8, pp.83538-83547, 2020.
- [6] Han, Jong-Sung, and Geun-Ho Lee. "VR tourism content using the HMD device." The Journal of the Korea Contents Association, Vol.15, No.3, pp.40-47, 2015.
- [7] Li, Zhengqi, and Noah Snavely. "Megadepth: Learning single-view depth prediction from internet photos." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [8] Ranftl, René, et al. "Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer." arXiv preprint arXiv:1907.01341, 2019.