

일반논문 (Regular Paper)

방송공학회논문지 제29권 제6호, 2024년 11월 (JBE Vol.29, No.6, November 2024)

<https://doi.org/10.5909/JBE.2024.29.6.1033>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

G-PCC 압축 포인트 클라우드의 복셀 기반 포인트 클라우드 렌더링

강형우^{a)}, 장선영^{a)}, 오병태^{b)†}

Voxel-based Point Clouds Rendering of G-PCC Compressed Point Clouds

Hyungwoo Kang^{a)}, Seonyoung Jang^{a)}, and Byung Tae Oh^{b)†}

요약

본 논문에서는 Geometry-based Point Cloud Compression (G-PCC)로 압축된 포인트 클라우드의 효율적인 렌더링 적용을 위하여 복셀 기반 처리를 통한 포인트 클라우드 렌더링 방식을 제안한다. G-PCC의 기하 정보 압축 특성을 활용하여 포인트 클라우드를 복셀로 변환함으로써, 희소한 포인트 클라우드에서도 홀 (Hole)이 발생하지 않고 기존의 그리드 기반 CNN이 적용할 수 있도록 네트워크를 구성하였다. 또한, 포인트 클라우드의 지역적, 전역적 동시에 추출하고 개별 포인트들의 특성까지 추출하여, 3차원 모델의 세부적 표현을 가능하게 하여 압축된 포인트 클라우드 스트리밍 환경에서도 네트워크의 표현 능력을 향상하도록 한다. 실험을 통해, 복셀 기반 처리가 포인트 클라우드의 지역적 및 전역적 특성을 효과적으로 분석할 수 있음을 확인하였다. 이러한 특성들은 네트워크가 새로운 데이터에 대해 더 정확하고 안정적인 성능을 발휘하는 데 기여하였으며, 이를 통해 네트워크는 다양한 종류의 포인트 클라우드 데이터에서 높은 일반화 능력과 희소한 데이터에서도 우수한 성능을 달성할 수 있었다.

Abstract

In this paper, we propose a method to improve point cloud rendering performance through voxel-based processing when using point clouds compressed by G-PCC as input. By transforming the point cloud into voxels, leveraging the geometric compression characteristics of G-PCC, the network is structured to apply existing grid-based CNNs without causing holes, even in sparse point clouds. Additionally, it enhances the network's representational capability in compressed point cloud streaming environments by analyzing both local and global features of the point clouds, as well as individual point features, enabling detailed representation of 3D models. Through experiments, it was confirmed that the voxel-based processing effectively analyzes both local and global features of point clouds. These features contributed to the network's ability to deliver more accurate and stable performance on new data. As a result, the network demonstrated strong generalization ability across various types of point cloud data and excellent performance even on sparse data.

Keyword : Point cloud rendering, G-PCC, Voxel-based, Deep learning, Deep network

Copyright © 2024 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. 서론

메타버스 산업은 AI 기술 발전과 비대면 수요 증가로 급 성장하고 있으며, 이에 사용자 경험을 극대화하기 위한 관련 기술의 중요성도 함께 커지고 있다. Neural Radiance Fields (NeRF) 기술^[1]의 등장으로, 다양한 시점에서 촬영된 2차원 영상들을 기반 3차원 공간상의 라디언스 (Radiance) 값을 추정하여 임의시점에서의 영상을 사실적으로 합성하는 것이 가능해졌다. 이 기술을 통해 간단한 장비와 적은 메모리로도 3차원 공간을 렌더링할 수 있게 되었으나, 카메라의 정확한 위치 및 방향이 필요하며, 이를 기반으로 한 레이 트레이싱을 통해 라디언스 값을 계산해야 하므로 계산 복잡도가 높다는 한계점이 존재한다.

최근, 센서 기술의 발전으로 LiDAR, ToF 등의 거리측정 센서로 얻을 수 있는 3차원 데이터가 점차 주목받고 있으며, 특히 가상현실과 같은 실시간 상호작용을 요구하는 응용 분야가 증가하면서, 보다 효율적이고 실시간 처리가 가능한 렌더링 기술의 필요성이 대두되고 있다. 다양한 3차원 데이터 중에서도 포인트 클라우드 데이터는 명시적인 공간 정보를 제공하기 때문에 학습시간을 단축하고 실시간 추론이 가능하다는 점에서 많은 관련 연구가 진행되고 있다.

현재 포인트 클라우드 렌더링 연구는 주로 고정된 장면에 집중하고 있으며, 일부 연구는 움직이는 장면을 다루지만, 알려진 장면에서 학습된 내용에만 기반하여 새로운 장면에서는 제대로 동작하지 않는 한계가 있다. 새로운 장면을 처리할 수 있는 일반화 모델 연구도 존재하지만, 계산

복잡도가 높아 실시간 추론이 어렵다. 또한, 대부분의 연구는 원본 포인트 클라우드를 입력 데이터로 사용하는 것을 가정하고 있어, 이는 실시간으로 전송되는 포인트 클라우드를 처리해야 하는 비디오 스트리밍 환경이나 대규모 데이터를 다뤄야 하는 응용 분야에 적용하기 어렵게 만드는 주요한 제약이 되고 있다.

본 논문에서는 이러한 문제를 해결하기 위해, 동일 인물의 다양한 동작을 학습하여 전역적 특징을 파악하고, 이를 바탕으로 개별 동작을 예측하는 부분적 일반화 모델을 제안한다. 이와 함께 원본 포인트 클라우드 대신 압축 과정을 통해 옥트리 (Octree) 구조로 표현된 포인트 클라우드를 입력 데이터로 사용하는 딥러닝 기반 렌더링 방식을 도입한다. 구체적으로, Geometry-based Point Cloud Compression (G-PCC)로 압축된 포인트 클라우드를 활용하여 압축 과정에서 손실된 정보를 복원하기 위해 포인트 주변 영역을 복셀 단위로 처리하고, 각 복셀의 특징을 추출해 렌더링 성능을 향상시키는 것을 목표로 한다. 이를 통해 계산 비용과 메모리 사용을 줄이면서도 높은 품질의 렌더링을 실현하고, 다양한 실험을 통해 그 성능을 검증하고자 한다.

본 논문은 다음과 같이 구성된다. 먼저 2장에서 포인트 클라우드 렌더링의 기존 기술에 대한 대략적인 소개 및 기술적인 분류와 포인트 클라우드 압축 방식을 소개한다. 3장에서는 기존 기술을 바탕으로 새롭게 제안된 기술을 소개하며, 4장에서는 제안 기술에 대한 검증을 위한 실험 결과 및 분석을 진행한다. 마지막으로 5장에서 제안 연구에 대한 결론을 제시한다.

a) 한국항공대학교 항공전자정보공학부(Electronics and Information Engineering, Korea Aerospace University)

b) 한국항공대학교 컴퓨터공학부(Computer Engineering, Korea Aerospace University)

‡ Corresponding Author : 오병태(Byung Tae Oh)
E-mail: byungoh@kau.ac.kr
Tel: +82-2-300-0409

ORCID: <https://orcid.org/0000-0003-1437-2422>

※ This work was supported by the IITP(Institute of Information & Communications Technology Planning & Evaluation)-ICAN(ICT Challenge and Advanced Network of HRD)(IITP-2024-RS-2024-004 37857, 50%) grant funded by the Korea government(Ministry of Science and ICT) and by the GRR program of Gyeonggi Province [2023-B02, Development of Intelligent Interactive Media and Space Convergence Application System, 50%].

· Manuscript September 13, 2024; Revised October 21, 2024; Accepted October 22, 2024.

II. 관련 기술

1. 포인트 클라우드 렌더링

포인트 클라우드 렌더링 기술은 그림 1과 같이 레스터화 모델, 레이 트레이싱 모델, 스플래팅 모델로 분류할 수 있다. 이 세 가지 모델은 각기 다른 방식으로 3차원 포인트 클라우드를 2차원 평면에 렌더링하는 방식이다.

레스터화 모델은 3차원 포인트 클라우드를 2차원 평면에 투영하는 방식으로, 간단하게 구현되어 실시간 렌더링에

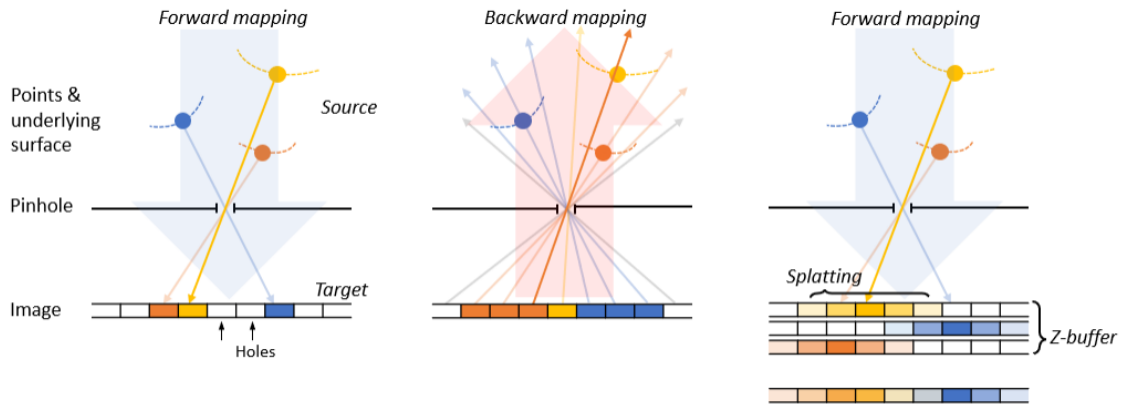


그림 1. 포인트 클라우드 렌더링 분류: 레스터화 모델 (왼쪽), 레이 트레이싱 모델 (중앙), 스플래팅 모델 (오른쪽)
 Fig. 1. Classification of point clouds rendering: Rasterization model (left), Ray tracing model (center), Splatting model (right)

적합하지만 깊이 정보 처리에 한계가 있다. 이를 보완하기 위해 학습 가능한 특징 벡터 삽입^[2], 구면 조화 함수 적용^[3], 표면 예측^[4] 및 주파수 정보 추출^[5] 등의 연구가 존재한다. 희소한 포인트 클라우드가 입력으로 들어오는 경우 2차원 평면에 홀이 발생하지만, 이미지-레벨 처리를 통해 홀을 제거할 수 있어서 비교적 안정적인 성능을 제공한다.

레이 트레이싱 모델은 NeRF^[1]를 포인트 클라우드에 적용하여, 각 픽셀에서 출발하는 광선을 추적해 포인트 클라우드와의 교차점을 계산하는 방식이다. 이는 짧은 학습시간과 빠른 추론을 가능하게 하지만, 실시간 렌더링에는 여전히 어려움이 있다. 광선과 포인트 클라우드의 교차점을 통해 표면 예측을 하는 다양한 연구가 존재한다. 구체적으로 광선과 포인트의 교차점 주변 포인트들을 활용하거나^[6], 법선 벡터를 사용해 더 정확한 표면을 예측하는 방법^[7], 희소한 포인트 클라우드를 조밀한 볼륨으로 변환하여 광선 추적 효율성 개선하는 방법^[8] 등이 연구되고 있다. 희소한 포인트 클라우드가 입력으로 들어오는 경우 포인트 간 간격이 넓어지면 광선이 포인트와 교차하는 빈도가 줄어들어 정확한 표면 예측이 어려워지고 렌더링 품질이 저하된다.

스플래팅 모델은 포인트 클라우드를 3차원 가우시안으로 변환해 2차원 평면에 투영하는 방식^[9]으로, 포인트 간 연속성을 보장하여 높은 품질의 영상을 생성하면서 물리 기반 알고리즘, 컴퓨터 그래픽스를 통해 매우 빠른 추론 시간을 갖는다. 관련 연구로는 딥러닝을 활용하여 가우시안 파라미터를 예측하는 방식^[10], 2차원 가우시안으로 변환하

여 메모리를 줄이는 방식^[11]이 존재한다. 스트리밍 환경에서 포인트 클라우드가 시간 t 의 입력 없이 제공될 경우, 새로운 포인트 클라우드에 대한 가우시안 파라미터를 초기화하기가 어렵고, 파라미터를 최적화하여 계산하는 데 시간이 오래 걸린다.

본 연구에서는 레스터화 모델을 기반으로 포인트 클라우드 렌더링 네트워크를 설계하였다. 입력된 희소한 포인트 클라우드에서 각 포인트를 복셀로 처리하고, 포인트 클라우드가 해당 복셀의 영역을 대표하도록 하여 희소한 포인트 클라우드에서도 안정적으로 동작하도록 하였다. 또한, 각 복셀의 특징을 계산하여 압축 과정 손실된 정보를 복원하여 렌더링 성능을 향상하고자 한다.

2. 포인트 클라우드 압축

포인트 클라우드 압축 방식은 G-PCC와 Video-based Point Cloud Compression (V-PCC)의 2가지 다른 접근방법으로 개발되고 있다^[12-13].

G-PCC는, 그림 2와 같이 기하 기반의 포인트 클라우드 압축 기술로, 포인트 클라우드의 기하학적 정보를 중심으로 데이터를 압축하는 방식이다. 점들의 위치나 형상을 유클리드로 효율적으로 표현하여 데이터 크기를 줄인다. 포인트 클라우드의 구조적 특성을 이용해 고정밀 데이터를 압축하는 데 유리하다. V-PCC는, 그림 3과 같이 비디오 기반 포인트 클라우드 압축 기술로, 포인트 클라우드를 여러 개

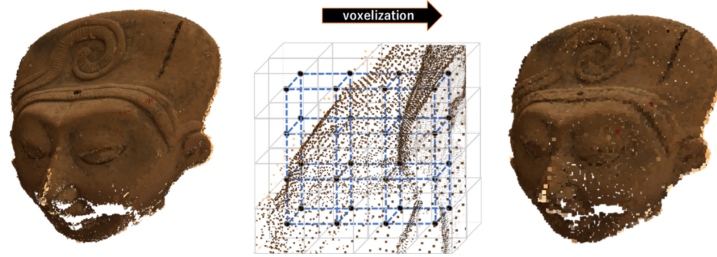


그림 2. G-PCC 부호화 과정에서의 복셀화^[14]
 Fig. 2. Voxelization in the G-PCC encoding process^[14]

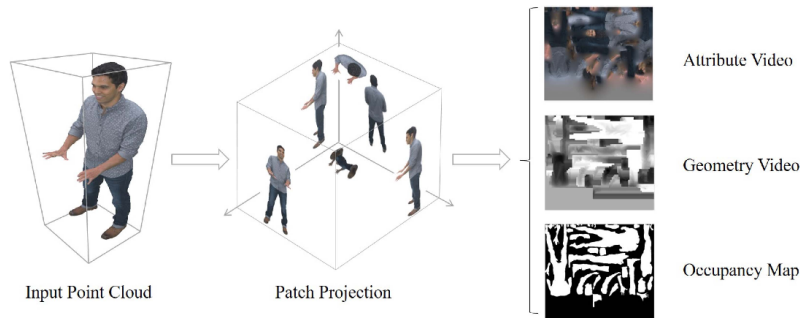


그림 3. V-PCC 부호화 과정에서의 패치 프로젝션^[15]
 Fig. 3. Patch projection in the V-PCC encoding process^[15]

의 2D 영상 프레임으로 변환하여 압축하는 방식이다. 기존 비디오 압축 코덱을 활용해 데이터를 압축하여 동적인 포인트 클라우드의 효율적 압축에 적합하다.

III. 제안 기법

1. 전체적인 네트워크 구조

제안 방식에서는 그림 4와 같이 포인트 클라우드 전역적 구조 정보와 개별 포인트 특성을 각각 예측하여 렌더링 성능을 개선하는 방식을 제안한다. 시스템의 첫 번째 과정으로서 개별 포인트 클라우드 데이터를 복셀화하여 복셀 데이터를 생성한다. 이러한 복셀화 과정을 통해서 희소한 포인트 클라우드에서 조밀한 3차원 데이터를 획득하여 생성 영상에서 홀 (hole)의 발생을 막을 수 있으며, 복셀로 변환이 된 포인트 클라우드 데이터에 정형화된 CNN 기술을 적용할 수 있다는 장점을 가지게 된다. 그림 5는 G-PCC로 압축된 포인트 클라우드가 옥트리 구조를 통해 기하 정보

를 표현하는 과정을 보여준다. 이 구조는 공간을 계층적으로 분할하여 복셀화 시 중복을 방지하고, 희소한 포인트 클라우드를 밀집된 복셀로 효율적으로 변환할 수 있게 한다. 그림 5에서 왼쪽은 개별 포인트를, 오른쪽은 변환된 복셀을 나타내며, 복셀화 과정을 통해 포인트 클라우드가 밀집된 형태로 변화하는 모습을 시각적으로 설명한다.

이후 단계로서 Voxel Network와 Point Network 두 가지 경로를 통해 포인트 클라우드의 특성 정보를 처리한다. Voxel Network에서는 전체 데이터의 구조적, 공간적 정보를 처리하는 부분으로, 3차원 컨볼루션 레이어를 사용하여 복셀 데이터를 처리하며, 데이터의 전역적 및 지역적 특성을 추출한다. Point Network에서는 깊이 영상을 통해 개별 포인트의 상세 정보를 처리하는 부분으로, MLP를 통해 라디언스 정보 추출한다. 마지막 Image Network에서는 Voxel Network와 Point Network에서 얻은 정보를 결합하여 최종 영상을 생성한다. 이 단계에서는 2차원 컨볼루션 레이어를 사용하여 생성된 영상에서 발생할 수 있는 홀이나 아티팩트를 제거하고, 더욱 상세한 영상을 제공한다. 이때, 깊이 정보를 통한 라디언스 정보 예측, 2차원 컨볼루션

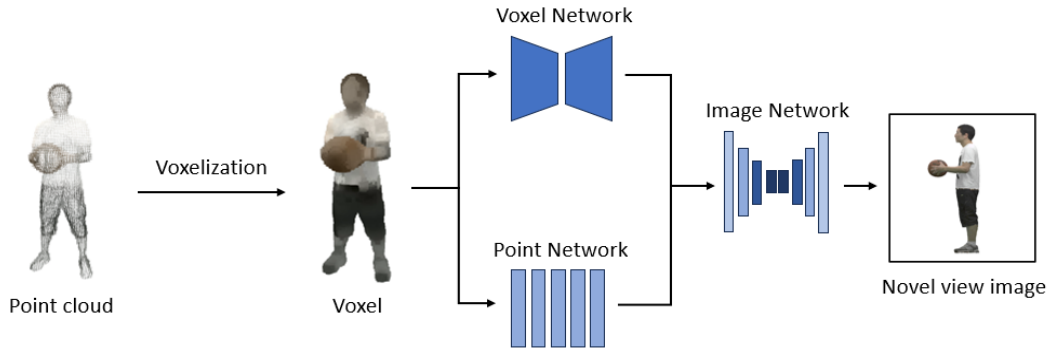


그림 4. 전체 네트워크 구조
 Fig. 4. Overall network architecture of the proposed system

레이어를 통한 아티팩트 제거 방식은 BPCR^[4]과 유사하지만, 복셀로 변환하여 전체 구조적 정보를 예측하는 과정에서 제안 방식과 큰 차이가 있다. 이에 대한 상세한 설명은 아래 세부 절에서 설명한다.

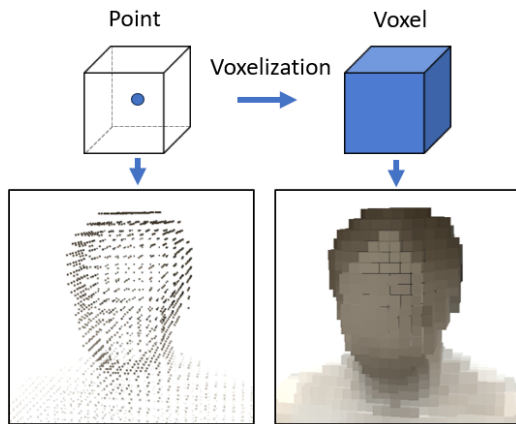


그림 5. 복셀화: 희소한 포인트 클라우드에서 조밀한 복셀로 변환
 Fig. 5. Voxelization: Transforming sparse point clouds into dense voxels

2. 복셀 기반 구조적 정보 예측

복셀 기반의 구조적 정보와 시각적 특성을 예측하기 위해 제안 방식에서는 색 정보를 포함한 복셀로 생성한다. 제안 방식에서는 VoxelNet^[16]의 구조를 그대로 차용하였다. VoxelNet은 기존 3D CNN과 다르게 PointNet 레이어와 3D 컨볼루션 레이어로 구성되어 희소한 데이터 처리에 효율적으로 빠르게 전체 복셀에서 특징을 추출하는데 효율적이다. 기존 VoxelNet의 입력은 하나의 복셀 안에 여러 포인

트가 존재하며, 이들 포인트의 집합을 통해 복셀의 특성을 계산하는 방식이다. 제안된 방식에서는 옥트리로 표현된 포인트 클라우드를 입력으로 사용하여 각 포인트가 개별 복셀을 대표하며, 해당 포인트의 특성을 추출하기 위해 VoxelNet을 활용하였다. 전체 3차원 특징 복셀은 카메라 파라미터를 통해 2차원 특징 영상으로 투영한다.

수식 (1)에 따라, 복셀의 특징 F_v 는 포인트 좌표 \mathbf{x} , 색 정보 \mathbf{c} 를 기반으로 Voxel network Θ_v 를 통해 얻어진다.

$$F_v = \Theta_v(\mathbf{x}, \mathbf{c}) \quad (1)$$

3. 포인트 기반 라디언스 정보 예측

포인트 기반 라디언스 정보 예측 단계에서는, 색 정보를 포함하지 않은 복셀이 깊이 영상을 생성하며, NeRF와 유사한 방식으로 그림 6과 같이 라디언스 매핑을 진행한다. 수식 (2)에 따라, 깊이 영상의 깊이 값 z , 카메라 위치 \mathbf{o} , 카메라 방향 벡터 \mathbf{d} 를 통해 카메라와 포인트 클라우드의 표면 교차점 \mathbf{x} 을 계산하는 데 사용한다.

$$\mathbf{x} = \mathbf{o} + z\mathbf{d} \quad (2)$$

포인트의 개별 좌표 \mathbf{x} 와 카메라 방향 벡터 \mathbf{d} 는 positional encoding을 수행하는 함수 $\gamma(\cdot)$ 를 통해 차원이 확장된 후, MLP의 입력으로 들어간다. 여러 층으로 구성된 MLP를 통과하면서 포인트의 라디언스 특성이 점차적으로 학습된다. 최종적으로 출력된 특징값은 각 포인트의 라디언스 특성을

나타내며, 이 값은 2차원 포인트 특징 맵으로 변환된다. 이러한 과정을 통해 3차원 공간 내의 각 지점에 대해 정확한 라디언스 값을 추정할 수 있다.

수식 (3)에 따라, 포인트의 라디언스 특징 F_p 는 포인트 좌표 x , 카메라 방향 벡터 d 를 기반으로 Point network Θ_p 를 통해 얻어진다.

$$F_p = \Theta_p(\gamma(x), \gamma(c)) \quad (3)$$

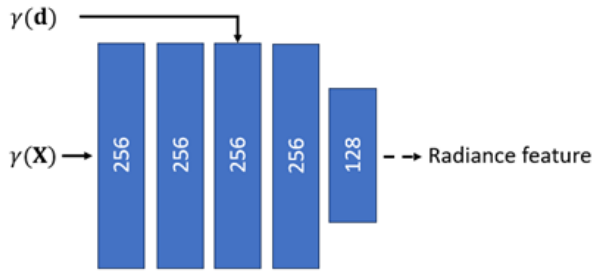


그림 6. Point Network 구조
Fig. 6. Point Network structure

4. 네트워크 구조 및 학습

생성된 두 개의 2차원 특징 맵은 U-Net을 통해 결합되며, 이 과정에서 복셀화 과정에서 발생할 수 있는 홀이나 아티팩트를 제거하기 위해 Image Network에서 정제화 과정이 추가된다. 최종적으로, 이러한 정제 과정을 거친 후 고품질의 RGB 영상이 생성된다. 최종적으로 사용한 손실함수 L_T 는 수식 (4)과 같이 구해진다.

$$L_T = (1 - \lambda)L_S + \lambda L_D \quad (4)$$

여기서 L_S 와 L_D 는 각각 상세 정보를 위한 L2 손실함수와 구조 정보를 위한 Perceptual 손실 함수^[17]를 의미하며, 실험에서 $\lambda = 0.01$ 으로 설정하였다.

IV. 실험 결과

1. 데이터셋

제안 기술의 성능측정을 위해 MPEG PCC Dynamic Human Sequence Dataset이 실험 데이터로 사용되었다. 해당 데이터셋은 약 3천만 개의 포인트로 구성된 시퀀스로, 각 시퀀스는 20초 동안 30fps로 기록되었다. 복셀 크기는 2048×2048×2048로 설정되었으며, 데이터셋의 포인트 클라우드는 G-PCC로 압축된 포인트 클라우드를 사용하

표 1. QP에 따른 출력 포인트 클라우드와 원본 포인트 클라우드와의 포인트 수와 D1-PSNR, D2-PSNR 비교

Table 1. Comparison of output point cloud using various QPs and original point cloud in terms of the number of points, and D1-PSNR, D2-PSNR

	QP	# pts. (M)	D1-PSNR	D2-PSNR
R01	51	0.01	49	59
R02	46	0.05	55	66
R03	40	0.2	69	83
R04	34	0.8	75	88
R05	28	1.7	88	98
R06	22	2.3	93	101



그림 7. 다양한 QP가 적용된 압축된 포인트 클라우드와 원본 포인트 클라우드와의 비교: R01, R02, R03, R06 및 원본 포인트 클라우드 (왼쪽부터)

Fig. 7. Comparison of compressed output point cloud using various QPs and the original point cloud: R01, R02, R03, R06, and the original point cloud (from left to right)

였다. 구체적으로 기하 정보는 옥트리로 손실 압축, 속성 정보는 RATH로 손실 압축을 하였고 실험은 R03 설정 하에서 진행되었다. 그림 7은 Quantization Parameter (QP)에 따른 G-PCC로 압축된 포인트 클라우드와 원본 포인트 클라우드의 비교를 보여준다. 표 1은 QP에 따른 포인트 수, D1-PSNR, D2-PSNR의 차이를 비교한 결과를 나타낸다. 이를 통해 압축 품질이 QP 값에 따라 어떻게 변화하는지 확인할 수 있다.

2. 네트워크 학습 설정

일반화 성능을 확인하기 위해 두 가지 방식으로 네트워크를 학습시켰다. 첫 번째 방식은 기존 일반화 모델의 학습 방식과 동일하게 서로 다른 오브젝트와 다양한 동작을 사용하여 학습을 진행하였다. 예를 들어, *Dancer*, *Model*, *Thai dancer* 오브젝트의 동작을 학습하여 *basketball player* 오브젝트를 테스트 하였다. 두 번째 방식은 동일한 오브젝트의 다른 동작만을 사용하여 학습하는 방식이다. 이 방식은 데이터 확보가 더 용이하며 실용적이다. 또한, 기존의

일반화 모델은 새로운 장면에 대응하기 위해 복잡하고 큰 모델이 필요하지만, 동일 오브젝트의 다양한 동작을 학습하는 방식은 모델이 처리해야 할 변수가 줄어들어, 비교적 간단하고 경량화된 모델로 구현할 수 있다.

실험에 사용된 데이터세트는 각각 600개의 포인트 클라우드로 구성되었으며, 이 중 501번부터 600번까지의 100개 포인트 클라우드를 사용하여 학습을 진행하였다. 테스트는 1번부터 200번까지의 200개 포인트 클라우드를 대상으로 수행하였다. 네트워크 학습에는 Adam 옵티마이저를 사용하며, 배치 크기는 1이다. VoxelNet, MLP와 U-Net의 초기 학습률은 각각 $1e-4$, $5e-4$ 와 $1.5e-4$ 이며, 학습률은 매 스텝마다 0.9999씩 곱해서 감소한다.

3. 실험 결과

그림 8과 표 2에서는 다양한 포인트 클라우드 렌더링 기술과 비교한 결과를 나타낸다. TriVol의 경우, 다른 오브젝트와 다른 동작으로 학습했을 때에도 PSNR 29.14, SSIM 0.956, LPIPS 0.054를 기록하며 우수한 일반화 성능을 보



그림 8. 다양한 포인트 클라우드 렌더링 기술의 시각적 비교: BPCR, Point-NeRF, TriVol, Proposed method, Ground truth (왼쪽부터), 다른 오브젝트의 다른 동작으로 학습 결과 (첫 줄), 동일 오브젝트의 다른 동작으로 학습 결과 (두 번째 줄)

Fig. 8. Visual comparison of various point cloud rendering techniques: BPCR, Point-NeRF, TriVol, Proposed method, Ground Truth (from left to right), results from training with different objects and different motions (first row), results from training with the same object and different motions (second row)

표 2. 다양한 포인트 클라우드 렌더링 기술의 성능 비교 (다른 오브젝트의 다른 동작 학습 결과 / 동일 오브젝트의 다른 동작 학습 결과)
 Table 2. Performance comparisons of various point cloud rendering methods (results from training with different objects and different motions/results from training with the same object and different motions)

	PSNR	SSIM	LPIPS ^[18]	FPS
BPCR ^[4]	19.94 / 23.26	0.803 / 0.852	0.202 / 0.159	68
Point-NeRF ^[6]	18.61 / 27.84	0.794 / 0.920	0.226 / 0.094	0.05
TriVol ^[8]	29.14 / 29.98	0.956 / 0.962	0.054 / 0.048	0.4
Proposed method	22.87 / 30.23	0.898 / 0.954	0.100 / 0.050	44

여주었다. 동일 오브젝트의 다른 동작으로 학습할 때도 PSNR 29.14 SSIM 0.962, LPIPS 0.048로 높은 수치를 기록하였으나, FPS는 0.4로 매우 낮았다. 반면, 제안 방식은 동일 오브젝트의 다른 동작으로 학습할 때 가장 높은 PSNR

을 보이면서 SSIM과 LPIPS에서도 우수한 결과를 보였다. 특히 FPS 44를 기록하며, 실시간 렌더링에 매우 적합한 성능을 확인할 수 있다.

그림 9는 제안 방식의 학습 방식에 따른 시각적 결과를



그림 9. 제안 방식 정성적 결과: 다른 오브젝트의 다른 동작으로 학습 결과 (첫번째 줄), 동일 오브젝트의 다른 동작으로 학습 결과 (두 번째 줄), 원본 데이터 (세 번째 줄)

Fig. 9. Qualitative results of the proposed method: results from training with different objects and different motions (first row), results from training with the same object and different motions (second row), and the ground truth (third row)

표 3. 제안 방식 정량적 결과 (다른 오브젝트의 다른 동작 학습 결과 / 동일 오브젝트의 다른 동작 학습 결과)

Table 3. Quantitative results of the proposed method (results from training with different objects and different motions / results from training with the same object and different motions)

	# pts (M)	PSNR	SSIM	LPIPS ^[18]	FPS
Basketball player	0.2	22.40 / 30.58	0.899 / 0.958	0.103 / 0.045	39
Dancer	0.18	23.70 / 30.21	0.905 / 0.957	0.098 / 0.047	41
Model	0.07	21.98 / 29.89	0.887 / 0.945	0.100 / 0.062	52
Thai dancer	0.17	23.41 / 30.24	0.902 / 0.954	0.098 / 0.046	43
Avg.	0.16	22.87 / 30.23	0.898 / 0.954	0.100 / 0.050	44

보여준다. 표 3에서 확인할 수 있듯이, 다른 오브젝트의 다른 동작으로 학습할 경우 성능이 다소 낮아지지만, 동일 오브젝트의 다른 동작으로 학습할 경우 PSNR 30.23, SSIM 0.954, LPIPS 0.050을 기록하며 높은 성능을 달성하였다.

V. 결론

본 논문에서는 G-PCC로 압축된 포인트 클라우드를 입력 데이터로 사용하여 렌더링 성능을 향상시키는 딥러닝 기반 렌더링 모델을 제안하였다. 제안 방식은 포인트 클라우드를 복셀화하여 복셀 기반의 처리를 통해 전역 특성을 추출하는 동시에, MLP를 통해 개별 포인트의 특성을 추출하여, 3차원 모델의 정밀한 표현을 가능하게 하고, 장면 간의 일반화 성능을 향상시켰다. 실험 결과, 압축된 희소한 포인트 클라우드 데이터에서도 제안 방식은 G-PCC의 압축 특성을 효과적으로 활용하여 기존의 포인트 클라우드 렌더링 방식보다 뛰어난 성능을 달성하였다. 특히, 복셀화 과정을 통해 3차원 모델의 전체 구조를 효율적으로 파악하는 것이 렌더링 성능 개선에 중요한 역할을 한다는 점이 확인되었다.

참고 문헌 (References)

- [1] Mildenhall, Ben, et al. "Nerf: Representing scenes as neural radiance fields for view synthesis." *Communications of the ACM* 65.1 (2021): 99-106.
doi: <https://doi.org/10.1145/3503250>
- [2] Aliev, Kara-Ali, et al. "Neural point-based graphics." *Computer Vision - ECCV 2020: 16th European Conference, Glasgow, UK, August 23 - 28, 2020, Proceedings, Part XXII* 16. Springer International Publishing, 2020.
- [3] Rakhimov, Ruslan, et al. "Npbg++: Accelerating neural point-based graphics." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.
doi: <https://doi.org/10.1109/CVPR52688.2022.01550>
- [4] Huang, Xiaoyang, et al. "Boosting point clouds rendering via radiance mapping." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 37. No. 1. 2023.
doi: <https://doi.org/10.1609/aaai.v37i1.25175>
- [5] Zhang, Yi, et al. "Frequency-modulated point cloud rendering with easy editing." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
doi: <https://doi.org/10.1109/CVPR52729.2023.00020>
- [6] Xu, Qiangeng, et al. "Point-nerf: Point-based neural radiance fields." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
doi: <https://doi.org/10.48550/arXiv.2201.08845>
- [7] Chang, Jen-Hao Rick, et al. "Pointersect: Neural rendering with cloud-ray intersection." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
doi: <https://doi.org/10.1109/CVPR52729.2023.00808>
- [8] Hu, Tao, et al. "Trivol: Point cloud rendering via triple volumes." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023.
doi: <https://doi.org/10.1109/CVPR52729.2023.01986>
- [9] Kerbl, Bernhard, et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering." *ACM Trans. Graph.* 42.4 (2023): 139-1.
doi: <https://doi.org/10.1145/3592433>
- [10] Zheng, Shunyuan, et al. "Gps-gaussian: Generalizable pixel-wise 3d gaussian splatting for real-time human novel view synthesis." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.
doi: <https://doi.org/10.48550/arXiv.2312.02155>
- [11] Huang, Binbin, et al. "2d gaussian splatting for geometrically accurate radiance fields." *ACM SIGGRAPH 2024 Conference Papers*. 2024.
doi: <https://doi.org/10.1145/3641519.3657428>
- [12] Byeon, Ju-Hyeong, Han-Sol Choe, and Dong-Gyu Sim. "MPEG G-PCC International Standard Technology." *Broadcasting and Media Magazine* 26.2 (2021): 31-45.
- [13] Jang, Ui-Seon. "Introduction to MPEG Video-based Point Cloud Compression Standard." *Broadcasting and Media Magazine* 26.2

- (2021): 18-30.
- [14] Graziosi, Danillo, et al. "An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC)." *APSIPA Transactions on Signal and Information Processing* 9 (2020): e13.
doi: <https://doi.org/10.1017/ATSIP.2020.12>
- [15] Choi, Yihyun, and Eun-Seok Ryu. "The Comparison of Compression and Conversion Performance for Real-time Volumetric Video Streaming by Using Dynamic Point Cloud and Mesh Coding." (2024).
doi: <https://doi.org/10.21203/rs.3.rs-4488649/v1>
- [16] Zhou, Yin, and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
doi: <https://doi.org/10.48550/arXiv.1711.06396>
- [17] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei. "Perceptual losses for real-time style transfer and super-resolution." *Computer Vision - ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*. Springer International Publishing, 2016.
doi: <https://doi.org/10.48550/arXiv.1603.08155>
- [18] Zhang, Richard, et al. "The unreasonable effectiveness of deep features as a perceptual metric." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
doi: <https://doi.org/10.1109/CVPR.2018.00068>

저 자 소 개



강 형 우

- 2023년 2월 : 한국항공대학교 항공전자정보공학부 (공학사)
- 2023년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부 (석사과정)
- ORCID : <https://orcid.org/0009-0008-6468-9069>
- 주관심분야 : 포인트 클라우드, 렌더링



장 선 영

- 2023년 2월 : 한국항공대학교 항공전자정보공학부 (공학사)
- 2023년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학부 (석사과정)
- ORCID : <https://orcid.org/0009-0000-5850-1447>
- 주관심분야 : 포인트 클라우드, 딥러닝



오 병 태

- 2003년 8월 : 연세대학교 전자공학과 (학사)
- 2007년 2월 : University of Southern California (석사)
- 2009년 8월 : University of Southern California (박사)
- 2009년 11월 ~ 2013년 2월 : 삼성종합기술원 연구원
- 2013년 3월 ~ 현재 : 한국항공대학교 교수
- ORCID : <http://orcid.org/0000-0003-1437-2422>
- 주관심분야 : 영상처리, 영상 포렌식, 비전, 딥러닝