

특집논문 (Special Paper)

방송공학회논문지 제30권 제6호, 2025년 11월 (JBE Vol.30, No.6, November 2025)

<https://doi.org/10.5909/JBE.2025.30.6.876>

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

## 원통형 투영 영상 특징 기반 효율적 포인트 클라우드 렌더링

강형우<sup>a)</sup>, 오병태<sup>b)†</sup>

### Efficient Point Cloud Rendering using Cylindrical Projection Feature Maps

Hyungwoo Kang<sup>a)</sup> and Byung Tae Oh<sup>b)†</sup>

#### 요약

기존의 포인트 클라우드 렌더링 방식은 완전한 포인트 클라우드가 입력으로 제공된다는 가정에 기반한다. 그러나 실제 응용, 특히 스트리밍 환경에서는 이러한 렌더링 방식이 한계를 보인다. 실제로 포인트 클라우드의 획득과 전송에는 압축이 수반되며, 그 결과 열화된 포인트 클라우드를 효과적으로 렌더링해야 한다. 이러한 문제점을 해결하기 위해 본 논문에서는 스트리밍 환경에 최적화된 원통 투영 기반의 새로운 포인트 클라우드 렌더링 기법을 제안한다. 제안 기법은 장면의 중심축을 기준으로 전체 포인트 클라우드를 방사 방향으로 2차원 원통면에 투영함으로써 왜곡을 최소화하고, 격자 기반 2차원 처리를 가능하게 한다. 제안 기법은 희소하거나 불완전한 포인트 클라우드에서도 안정적이고 일관된 성능을 보인다. 실험 결과, 제안 기법은 추가 미세 조정 없이도 다양한 장면과 객체를 효과적으로 렌더링하며, 기존 방법 대비 스트리밍 환경에서 우수한 성능과 일반화 능력을 입증한다.

#### Abstract

Conventional point cloud rendering methods proceed under the assumption that the input point cloud is complete. However, in real-world setting, quality degradation often occurs during acquisition and transmission. To address this issue, a rendering network that operates reliably even on degraded point clouds is needed. This paper proposes a novel point cloud rendering method that employs a cylindrical projection tailored for streaming environments. The proposed method projects the entire point cloud onto a cylindrical plane with respect to its central axis, reducing spatial distortion and enabling efficient grid-based 2D processing. In particular, the method delivers strong performance even for sparse or incomplete point clouds. Extensive experiments show that the proposed approach effectively renders diverse scenes and objects without additional fine-tuning and achieves superior performance and generalization in streaming environments compared with conventional methods.

Keyword : Point cloud rendering, Cylindrical projection, Image-based, Deep learning, Deep network

## I. 서 론

과거 임의 시점 영상 합성의 표준으로 자리 잡았던 Neural Radiance Fields (NeRF)<sup>[1]</sup>는 다각도 2차원 영상들을 바탕으로 3차원 공간 내 방사값을 추정하여 사실적인 영상을 합성할 수 있었지만, 신경망 기반 암시적 표현과 볼륨 렌더링으로 인한 느린 추론 시간으로 실시간 스트리밍 적용에 한계가 있었다. 반면 3D Gaussian Splatting (3DGS)<sup>[2]</sup>은 장면을 명시적 가우시안 합성으로 표현해 GPU 병렬화에 최적화된 초고속 렌더링을 제공하면서도 NeRF를 능가하는 사실적 표현력을 보여주어, 현재 임의 시점 영상 합성 분야의 표준으로 자리매김했다.

그러나 현재 3DGS 연구를 비롯한 명시적 데이터를 사용하는 포인트 클라우드 렌더링 연구에는 몇 가지 한계가 존재한다. 첫 번째, 대부분의 연구는 정적인 장면에 초점을 맞추고 있다. 일부 동적인 장면을 다루는 연구가 존재 하지만, 대체로 프레임 단위 최적화에 의존하기 때문에 새로운 장면 적용 시 성능 저하가 발생한다. 몇몇 일반화 가능 모델도 제안되었으나, 높은 계산 복잡도와 다중 시점 영상 의존성으로 인해 실시간 응용과 실제 환경에서의 실용성이 떨어진다. 더 나아가, 기존 연구는 장면별 미세 조정을 요구하는 경우가 많아 실시간성을 저해한다. 둘째, 많은 연구가 비압축되고 완전한 포인트 클라우드를 입력으로 가정한다. 그러나 실제 환경에서 포인트 클라우드의 획득과 전송 과정에서 압축이 불가피하고 이로 인한 입력

데이터의 품질 저하는 렌더링 품질 저하로 이어진다. 이러한 한계들은 실제 스트리밍 환경에서 더욱 두드러지는데, 스트리밍 환경에서는 포인트 클라우드가 프레임 단위로 실시간 입력되고, 즉시 단일 패스 추론으로 렌더링하며, 장면별 미세 조정이 허용되지 않기 때문에 시간축을 명시적으로 활용한 4차원 처리가 불가능하다. 따라서 각 프레임을 독립적인 3차원 입력으로 다루면서도 안정적인 성능을 보장할 수 있는 강건한 일반화 모델이 필수적이다.

본 논문에서는 실시간 처리가 가능한 일반화된 포인트 클라우드 렌더링 시스템으로서, 열화된 입력에서도 강건하게 동작하는 것을 목표로 한다. 이를 위해 먼저 3차원 포인트 클라우드를 원통형 투영을 통해 2차원 평면으로 변환하는 새로운 렌더링 방식을 적용하였다. 이 과정을 통해 비정형 포인트 클라우드를 정형화된 격자로 변환하여 지역적, 전역적 특징을 효율적으로 추출할 수 있으며, 회소하거나 불완전한 입력 포인트 클라우드에서도 안정적으로 동작할 수 있다. 나아가 동일 인물의 다양한 동작을 사전에 학습하는 부분적 일반화 학습 방식을 도입하여 전역적 형상 특징을 안정적으로 파악하고, 각 입력 프레임의 개별 동작을 독립적으로 예측할 수 있도록 설계하였다. 이러한 접근은 장면별 미세 조정 없이도 빠른 추론과 최소한의 일반화 능력을 동시에 달성하며, 완전 일반화 모델 대비 계산 부담을 줄여 스트리밍 환경에서 요구되는 실시간성과 강건성을 확보한다. 마지막으로 다양한 실험을 통해 제안 기법의 성능을 검증한다.

본 논문은 다음과 같이 구성된다. 2장에서는 포인트 클라우드 렌더링 방식의 분류에 해당하는 기존 기술들을 개괄적으로 소개한다. 3장에서는 이러한 기존 기술들을 바탕으로 본 논문에서 새롭게 제안하는 기술을 상세히 설명한다. 4장에서는 제안 기술의 성능을 검증하기 위한 실험 결과와 그에 대한 분석을 제시한다. 마지막으로 5장에서는 제안 연구에 대한 결론을 제시한다.

## II. 관련 기술

본 장에서는 포인트 클라우드의 렌더링 방식에 따른 분류와 포인트 클라우드 렌더링의 동적 장면 모델과 일반화 모델에 대한 개략적 설명과 장단점 분석, 그리고 최신 기술 동향을 살펴본다.

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

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

‡ Corresponding Author : 오병태(Byung Tae Oh)  
E-mail: byungoh@kau.ac.kr  
Tel: +82-2-300-0408

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

※ 이 논문의 결과 중 일부는 한국방송·미디어공학회 2025년 하계 학술대회에서 발표한 바 있음

\* This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean Government (MSIT) (RS-2023-00229330), and by a grant(RS-2025-02633421, Drone SAR System Development for Monitoring Landslide Structure and Collapse Analysis) funded by Ministry of Interior and Safety(MOIS, Korea)

· Manuscript September 9, 2025; Revised October 14, 2025; Accepted October 14, 2025.

## 1. 포인트 클라우드 렌더링 분류

포인트 클라우드 렌더링은 매핑 방법에 따라 순방향 매핑 (forward mapping)과 역방향 매핑 (backward mapping)으로 구분할 수 있다. 순방향 매핑은 원본 도메인에서 타겟 도메인으로 직접 매핑하는 방식이고, 역방향 매핑은 타겟 도메인에서 출발해 원본 도메인 좌표를 역으로 찾아가 샘플링하는 방식이다.

순방향 매핑 기반 포인트 클라우드 렌더링은 3차원 포인트 클라우드를 2차원 평면에 투영하는 방식으로, 간단한 구현이 가능하지만 포인트 클라우드 자체가 텍스쳐 정보가 부족하다는 한계가 있다. 이를 해결하기 위해 학습 가능한 특징 벡터 삽입<sup>[3]</sup>, 시점 의존적 표현을 위한 구면 조화 함수 (spherical harmonics)<sup>[4]</sup>, 또는 고주파 성분 보강을 위한 주파수 인코딩 기법<sup>[5]</sup> 등이 제안되었다. 그러나 여전히 입력 포인트 클라우드가 희소하거나 불완전할 경우 2차원 평면에 구멍이 자주 발생하는 근본적인 한계가 있으며, 이로 인해 필수적으로 추가적인 영상 수준 후처리가 요구된다.

이러한 한계를 극복하기 위해 최근에는 포인트 클라우드의 각 포인트를 3차원 가우시안으로 변환한 뒤 이를 평면으로 투영하는 방식이 제안되었다<sup>[2]</sup>. 이러한 방식은 2차원 평면에서 구멍 발생을 막을 수 있어 추가적인 영상 후처리 없이 빠른 추론을 가능하게 한다. 한 픽셀에 여러 가우시안이 매핑되어 알파 블렌딩 (alpha blending)을 통해 보다 사실적인 표현을 가능하게 한다. 그러나 이러한 3차원 가우시안의 위치, 크기, 방향성 등 파라미터를 안정적으로 예측하기 어렵다는 한계가 있으며, 이를 개선하기 위해 최근 연구에서는 계층적 혼합 모델을 통한 점진적 추정<sup>[6]</sup>, 학습 기반 초기화와 정규화를 통한 파라미터 보정<sup>[7]</sup> 등의 기법이 제안되고 있다.

역방향 매핑 기반 포인트 클라우드 렌더링은 픽셀에서 출발해 입력 포인트 클라우드와의 교차점을 역으로 찾아가는 방식으로, 포인트 클라우드에서 물체의 표면을 추정하여 교차점을 계산하기 때문에, 2차원 평면에 빈 공간이 발생하지 않고 텍스쳐 정보를 복원하여 비교적 안정적인 표현을 제공한다. 그러나 입력 포인트 클라우드가 희소하거나 불완전할 경우 정확한 표면 추정이 어려워 렌더링 성능

이 저하되는 한계가 있다. 이를 해결하기 위해 다양한 표면 보정 기법들이 제안되었으며, 대표적으로 주변 포인트를 활용<sup>[8]</sup>, 법선 벡터 통합<sup>[9]</sup>, 희소 포인트 클라우드를 조밀한 특징 볼륨 변환<sup>[10]</sup> 등이 있다. 하지만 역방향 매핑은 구조적으로 높은 계산 복잡도를 수반하기 때문에 실시간 렌더링이 어렵고 다양한 입력 분포에 대한 일반화에도 제약이 따른다.

## 2. 포인트 클라우드 렌더링에서의 동적 모델과 일반화 모델

포인트 클라우드 렌더링에서 동적 모델은 크게 기준 공간 기반 변형 네트워크 접근과 시간축을 포함한 4차원 포인트 클라우드 접근으로 나눌 수 있다.

기준 공간 기반 접근은 동적 장면을 정적 기준 공간에 정렬한 뒤 변형 필드를 학습하여 타겟 시점을 예측하는 방식이다. 시간 일관성을 유지할 수 있으나 계산량이 많고 큰 변형에서 불안정하다. 이를 해결하기 위해 루트 파인딩 안정화 (root-finding stabilization)<sup>[11]</sup>, 스키닝 기반 정규화 (skinning based normalization)<sup>[12]</sup>, 시공간 제약<sup>[13]</sup> 등이 제안되었다.

4차원 포인트 클라우드 접근은 입력에 시간축을 포함해 직접 처리하는 방식으로, 대규모 모션과 시계열 일관성을 자연스럽게 다룰 수 있다. 그러나 메모리 요구가 크고 깜빡임 억제를 위한 정규화가 필요하다. 이를 해결하기 위해 시간 민감도 기반 프루닝<sup>[14]</sup>, 저랭크 표현, 계층적 가우시안 구조<sup>[7]</sup> 등이 제안되었다.

두 방식 모두 명시적 시간축을 활용하고, 장면별 미세조정이 필요해 스트리밍 환경 적용에는 한계가 있다.

일반화 모델의 최근 동향은 3차원 데이터의 지역적, 전역적 특징을 효율적으로 추출하는 것에 초점을 둔다. 대표적으로 포인트 클라우드 자체에서 직접 3차원 특징을 학습하는 방식<sup>[10]</sup>과 다중 시점 영상을 이용해 3차원 코스트 볼륨 (cost volume)을 구성하는 Multi-View Stereo (MVS) 기반 방식<sup>[15]</sup>이 있다. 3차원 특징 추출 기반 방식은 입력의 구조적 정보를 직접 활용할 수 있으나, 계산 복잡도가 높고 희소하거나 열화된 입력에서 성능 저하가 크다. MVS 기반 접근은 시점별 영상으로 보완할 수 있다는 장점이 있으나, 입력 조건에 따라 성능이 크게 좌우되어 실용성이 낮다.

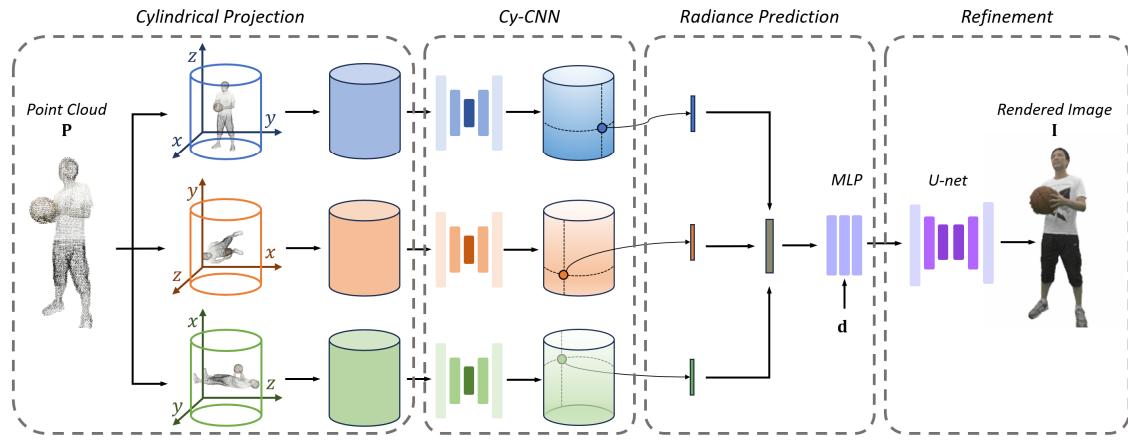


그림 1. 제안 기법의 네트워크 구조

Fig. 1. Network architecture of the proposed method

### III. 제안 기술

본 논문에서는 3차원 포인트 클라우드의 특징을 효과적으로 추출하는 새로운 렌더링 모델을 제안한다. 그림 1은 제안 방식의 전체 시스템 개략도를 나타낸다. 즉, 제안 시스템은 3차원 포인트 클라우드를  $x, y, z$  축 기준으로 원통형으로 투영하여 2차원 평면으로 변환하여 2차원 이미지 기반 처리를 진행하는 단계, 그리고 3개의 2차원 영상 기반 추출된 특징을 통해 지역적, 전역적 특징을 효율적으로 추출하는 단계, 그리고 포인트 기반 방사값 예측 단계, 영상 수준 정제 단계로 이루어져 있다. 각 단계별 상세한 설명은 아래의 세부 절에서 설명한다.

#### 1. 원통형 투영

첫 번째 단계는 포인트 클라우드를 원통면에 투영한다. 포인트 클라우드는  $x, y, z$  축을 기준으로 반지름  $\rho$ 의 원통형 표면에 투영된다. 이 과정은 직교 좌표계에서 각 점  $\mathbf{p}_i = (x_i, y_i, z_i)$ 를 원통 좌표계  $\mathbf{p}_{i,z}^{CY} = (\rho_{i,z}, \theta_{i,z}, h_{i,z})$ 로 매핑하여 원통형 포인트 클라우드를 생성한다. 직교 좌표계에서 원통형 좌표계로의 매핑은 식 (1)-(3)에 의해 정의된다:

$$\rho_{i,z} = \rho_0, \quad (1)$$

$$\theta_{i,z} = \arctan\left(\frac{y_i}{x_i}\right), \quad (2)$$

$$h_{i,z} = z_i, \quad (3)$$

여기서  $\rho_0$ 은 고정된 반지름 값으로, 모든 점이  $z$ 축으로부터 동일한 거리에 위치하도록 하며,  $\theta_{i,z}$ 는  $xy$  평면에서 점  $\mathbf{p}_i$ 의 방위각이고,  $h_{i,z}$ 는  $z$ 축을 따라 높이를 그대로 유지한다. 이 단계는 포인트 클라우드의 높은 해상도를 유지하면서 효율적 계산을 가능하게 한다.

다음 단계는 원통 좌표의  $\theta$ 와  $h$  값을 사용해 점들을 2차원 평면에 투영한다.  $\theta$  값은 2차원의 영상의 수평 방향에 대응하며,  $h$  값은 수직 방향에 대응한다. 2차원 영상의 좌우 경계가 연속적으로 연결되어 원통형 구조를 2차원 평면에 매핑한다. 따라서 2차원 영상은 경계에서의 정보 손실을 방지하는 연속적인 연결을 갖는다. 2차원 평면 투영에서는 식 (4)-(5)에 의해 평면상의 좌표  $(u_{i,z}, v_{i,z})$ 를 정의한다:

$$u_{i,z} = \frac{(\theta_{i,z} - \theta_{\min})}{(\theta_{\max} - \theta_{\min})} W, \quad (4)$$

$$v_{i,z} = \frac{(h_{i,z} - z_{\min})}{(z_{\max} - z_{\min})} H', \quad (5)$$

여기서  $u_{i,z}$ 는  $\theta_{i,z}$ 를  $(\theta_{\min}, \theta_{\max})$  사이로 정규화되어 2차원 평면의 폭  $W'$ 에 매핑된다.  $v_{i,z}$ 는  $h_{i,z}$ 을  $z_{\min}$ 에서  $z_{\max}$  사이로 정규화되어 높이  $H'$ 에 매핑된다.

포인트 클라우드에 대해 원통형 투영을 수행할 때, 투영 축에 따라 정보 손실의 양이 달라진다. 이러한 손실을 보완하기 위해 본 논문에서는  $x, y, z$ 축이라는 세 가지 주요 축을 따라 원통형 투영을 수행했다. 이를 통해 3차원 포인트 클라우드의 다중 2차원 평면을 생성하여 공간 정보를 포괄적으로 분석하고 효율적으로 처리할 수 있도록 했다.

이 과정은 3차원 포인트 클라우드를 2차원 평면으로 변환하여 전체 구조를 유지하면서 계산 복잡성을 줄이고 처리 효율성을 높인다. 원통형 투영 방법은 포인트 클라우드의 원래 기하학적 세부 정보를 유지하면서 고품질 영상을 생성하고 빠른 영상 생성을 가능하게 한다.

## 2. 원통형 컨볼루션 뉴럴 네트워크

본 절의 목표는 2차원 원통면으로부터 특징을 추출하는 것이다. 이를 위해, 원통형 슬라이딩 윈도우 (Cylindrical Sliding Window, CSW) 기술<sup>[16]</sup>을 효과적으로 활용할 수

있도록 설계된 원통형 컨볼루션 뉴럴 네트워크 (Cy-CNN) 아키텍처를 제안한다.

그림 2는 원통형 뉴럴 컨볼루션 네트워크를 나타낸다. CSW는 원통형으로 투영된 평면 이미지의 왼쪽과 오른쪽 경계를 연결하여 연속적으로 처리한다. 이러한 경계 연결은 CNN 커널이 경계를 넘어 연속적으로 연산할 수 있도록 하여 원통형 좌표 시스템의 구조적 특성을 효과적으로 반영한다.

기존의 2D CNN 기반 투영 방식, 특히 구형 투영 기반 방식은 극점 부근에서의 왜곡과 불연속적인 경계로 인해 연속적 특징 학습이 어렵다. 반면 제안하는 Cy-CNN은 원통형 좌표계의 주기성을 CNN 내로 직접 통합함으로써, 별도의 패딩이나 좌표 보정 없이도 일관성을 유지하며, 연속적, 효율적 특징 추출이 가능하다.

## 3. 포인트 기반 방사값 예측

본 단계는 앞 절에서 얻은 원통형 특징을 포인트 클라우드에 삽입한 뒤, 목표 시점의 렌더링에 필요한 포인트별 방사값을 빠르게 산출하는 과정이다. 먼저 각 포인트에 Cy-CNN이 추출한 특징을 부착하고, 카메라 방향 벡터를

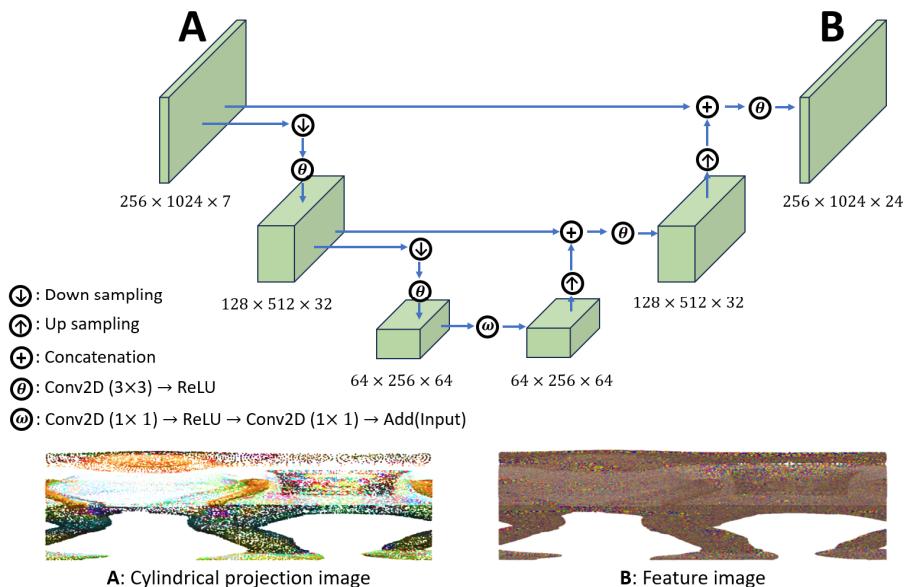


그림 2. 원통형 뉴럴 컨볼루션 네트워크 세부 구조  
Fig. 2. Detailed architecture of the cylindrical convolution neural network

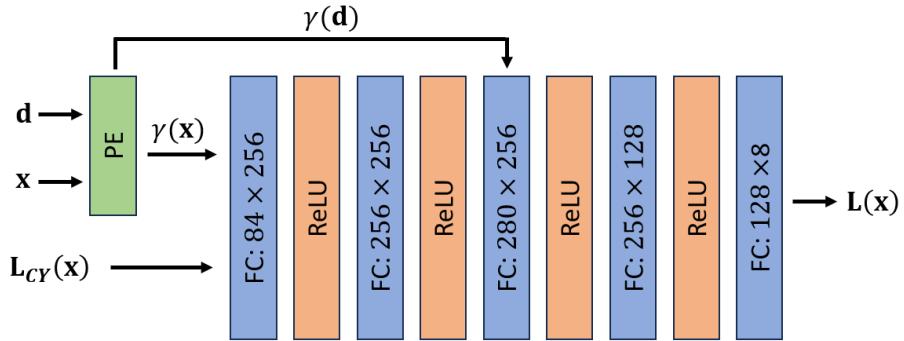


그림 3. 포인트 기반 라디언스 정보 예측기 세부 구조

Fig. 3. Detailed architecture of the point-based radiance predictor

통해 2차원 평면으로 투영한다. 이 과정에서 피셀-포인트 대응이 즉시 정해지며 별도 레이 마칭 (ray marching) 없이 교차 지점을 얻는다. 그림 3은 포인트 기반 방사값 예측기 세부 구조를 나타낸다. MLP는 포인트 좌표  $\mathbf{x}$ , 카메라 방향 벡터  $\mathbf{d}$ , Cy-CNN 특징  $L_{CY}(\mathbf{x})$ 으로부터 포인트별 방사값 정보  $\mathbf{L}(\mathbf{x})$ 를 직접 예측한다. 앞단의 Cy-CNN이 전체 포인트 클라우드에서 추출한 지역적, 전역적 정보를 각 포인트 단위 특징으로 구체화되어, 시점 의존적 표현이 가능해지며, 장면별 미세 조정 없이도 프레임 독립적으로 안정적인 초기 렌더링을 얻을 수 있다.

#### 4. 영상 수준 정제 네트워크

최종 잡재 특징은 2차원 영상으로 재구성되며, 비어 있는 피셀의 값은 0으로 할당된다. MLP를 통해 출력된 특징 영상은 이후 2D U-Net으로 구현된 정제 단계로 전달된다. 정제 네트워크는 3차원 공간 정보와 독립적으로 작동하며, 누락된 영역을 채우고 왜곡을 제거한다. 학습 과정에서는 L2 손실과 지각 손실(perceptual loss)의 가중합을 사용한지도 학습을 수행한다.

## IV. 실험 및 결과

### 1. 실험 세부 사항

제안 모델은 사전 학습 없이 end-to-end로 학습된다. AdamW

( $\beta_1=0.9$ ,  $\beta_2=0.999$ ) 옵티마이저를 사용하였으며, 배치 크기는 1로 설정하였다. 손실함수는  $L = L_{RGB} + \lambda L_{perceptual}$ 로 정의되며,  $\lambda$ 는  $5 \times 10^{-3}$ 이다. 렌더링 해상도는  $512 \times 512$ , 원통 투영 맵 크기는  $256 \times 1024$ 로 설정하였다. GT 이미지는 메시 데이터를 2D로 투영하여 생성하였다.

### 2. 데이터셋

본 논문에서는 원통형 투영을 통한 포인트 클라우드 렌더링 성능을 정적 장면과 스트리밍 환경에서 평가한다. 정적 장면 렌더링 모델 평가를 위해 데이터셋은 NeRF Synthetic Dataset<sup>[1]</sup>를 사용하였다. 각 장면에 대해서 100개의 시점에 대해서 학습을 하고, 200개의 임의 시점에 대해서 렌더링 성능을 평가하였다.

스트리밍 환경 렌더링 모델 평가를 위해 데이터셋은 동적 장면으로 구성된 MPEG PCC Human Sequence Dataset<sup>[17]</sup>을 하였다. 해당 데이터셋은 600개의 포인트 클라우드 프레임으로 구성되어 있다. 모델의 부분적 일반화 성능을 평가하기 위해 동일 시퀀스에서 학습 및 테스트 데이터셋을 분리하였다. 학습 데이터로는 프레임 501번부터 600번까지 총 100개의 프레임을 사용하였고, 테스트 데이터로는 다양한 동작을 포함하는 프레임 1번부터 200번 까지 총 200개의 프레임을 사용하여 평가를 진행하였다. 스트리밍 환경을 구현하기 위해 포인트 클라우드를 기하 기반 포인트 클라우드 압축 (G-PCC)을 사용하여 압축하였다. 압축은 R01에서 R06까지 다양한 양자화 파라미터를 사용

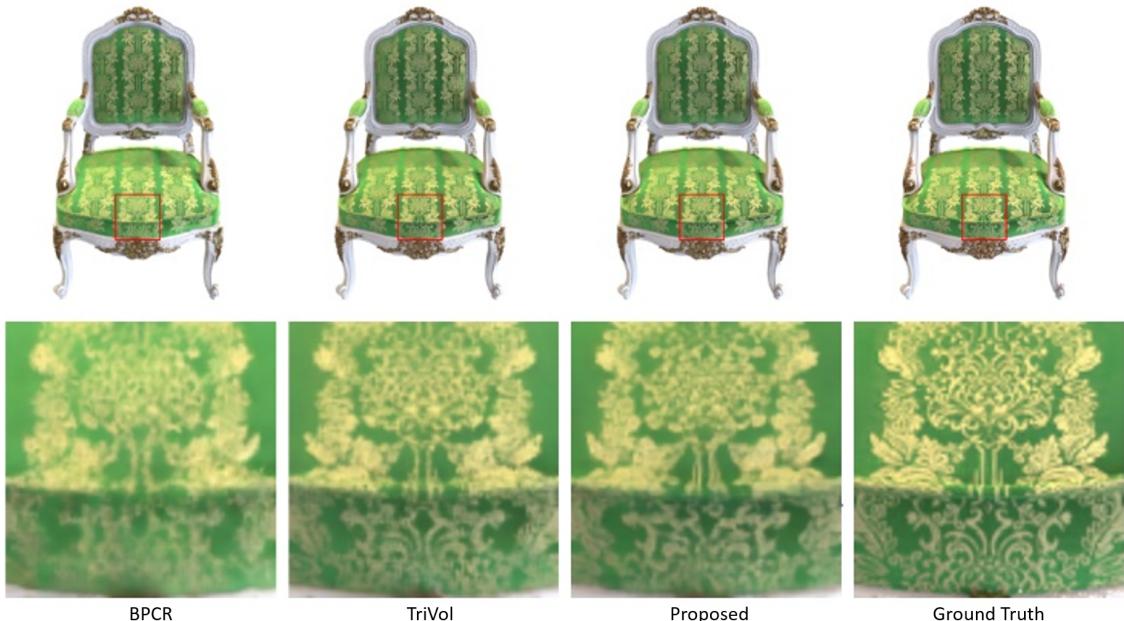


그림 4. NeRF Synthetic 데이터세트에서 시각적 결과 비교  
Fig. 4. Visual comparison on the NeRF Synthetic dataset

하여 모델의 포인트 클라우드 품질의 강인성을 평가하였다. 이때, R01부터 R06 양자화 파라미터는 G-PCC 표준 실험을 참고하였다<sup>[18]</sup>.

표 1. NeRF Synthetic 데이터세트에서 정량적 결과 비교  
Table 1. Quantitative comparison on the NeRF Synthetic dataset

NeRF Synthetic Dataset <sup>[1]</sup>			
Method	PSNR	SSIM	LPIPS
BPCR <sup>[20]</sup>	30.84	0.946	0.056
FreqPCR <sup>[5]</sup>	31.14	0.950	0.049
Point2Pix <sup>[21]</sup>	31.09	0.948	0.081
TriVol <sup>[10]</sup>	31.34	0.951	0.047
Proposed	31.86	0.959	0.038

### 3. 정적 장면 실험 결과

그림 4와 표 1은 정적 장면 렌더링 평가를 위한 NeRF Synthetic Dataset에서의 정성, 정량 결과이다. 제안 기법은 정성, 정량 양 측면에서 일관되게 우수한 성능을 보여준다. 정성 비교에서 제안 기법은 텍스처 디테일과 경계 선명도를 잘 보존하며, 정량 지표에서도 모든 평가 지표에서 최고 성능을 기록했다. 이는 원통형 투영을 통한 포인트 클라우드 처리가 기존

기법 대비 텍스처 정보를 잘 포착하고 복원함을 입증한다.

### 4. 스트리밍 환경 실험 결과

그림 5와 표 2는 G-PCC로 압축한 다양한 양자화 구간에서의 성능을 보여준다. 제안 기법은 전 구간에서 일관되게 우수한 렌더링 성능을 보이며, 특히 포인트 수가 급격히 감소하는 고압축 조건에서 격차가 가장 크게 벌어진다. 정성 비교에서도 경계 보존, 텍스처 연속성, 홀 인페인팅 등에서 아티팩트가 현저히 감소해 세부 복원이 안정적으로 유지됨을 확인할 수 있다.

이 결과는 제안된 Cy-CNN이 희소한 입력 데이터에서도 강인성을 보이며, 포인트 클라우드에서 다양한 희소성 수준을 효과적으로 처리할 수 있음을 보여준다. 이러한 성능은 Cy-CNN이 희소한 데이터에서도 높은 성능을 유지할 수 있음을 보여주며, 기존 방법들보다 효과적임을 보여준다.

더불어 포인트 클라우드를 원통형 투영 기반 2차원 격자로 구조화하고 2차원 처리하는 설계로 연산 부담을 낮추면서, 경쟁 기법 대비 더 높은 정확도와 빠른 추론을 동시에 달성한다.

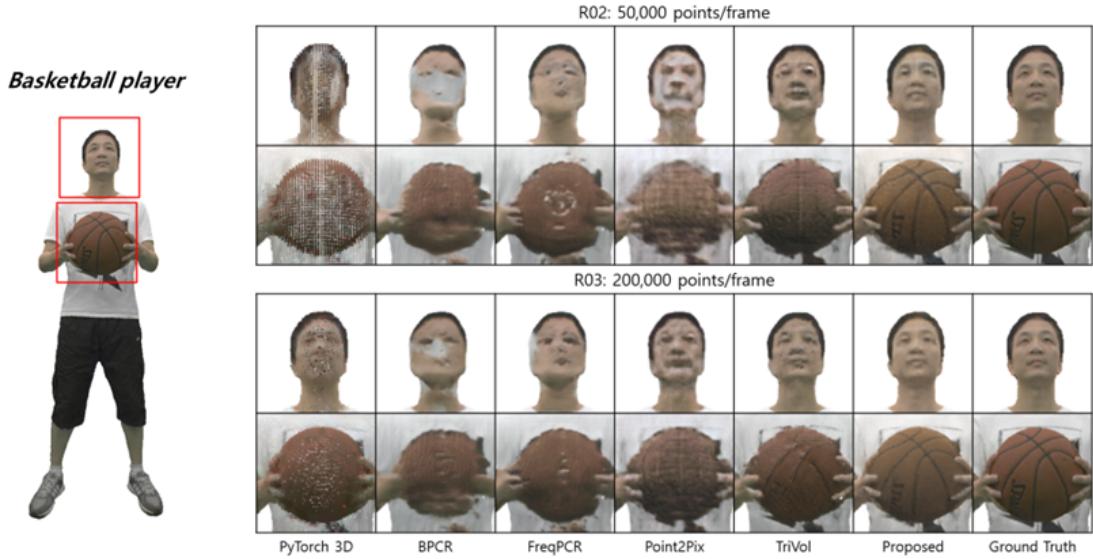


그림 5. MPEG-PCC 데이터세트에서 시각적 결과 비교  
 Fig. 5. Visual comparison on the MPEG-PCC dataset

표 2. MPEG-PCC 데이터세트에서 정량적 결과 비교  
 Table 2. Quantitative comparison on the MPEG-PCC dataset

Method	MPEG-PCC Dynamic Human Sequence Dataset <sup>[17]</sup>											
	R02: 50,000 points/frame				R03: 200,000 points/frame				R05: 1,700,000 points/frame			
	PSNR	SSIM	LPIPS	FPS	PSNR	SSIM	LPIPS	FPS	PSNR	SSIM	LPIPS	FPS
Pytorch3D <sup>[19]</sup>	18.42	0.742	0.352	<b>76</b>	19.46	0.752	0.362	<b>68</b>	20.13	0.789	0.354	<b>62</b>
BPCR <sup>[20]</sup>	18.21	0.793	0.245	60	19.21	0.823	0.215	53	18.52	0.852	0.202	32
FreqPCR <sup>[5]</sup>	18.35	0.792	0.276	59	18.76	0.842	0.226	51	19.42	0.856	0.196	28
Point2Pix <sup>[21]</sup>	26.16	0.952	0.079	0.6	28.46	0.954	0.078	0.1	31.26	0.977	0.065	0.001
TriVol <sup>[10]</sup>	28.52	0.968	0.057	1	29.98	0.972	0.039	0.4	<b>33.96</b>	0.979	0.038	0.07
Proposed	<b>30.98</b>	<b>0.969</b>	<b>0.046</b>	45	<b>31.82</b>	<b>0.975</b>	<b>0.032</b>	44	32.28	<b>0.984</b>	<b>0.022</b>	25

## V. 결론

본 논문은 스트리밍 포인트 클라우드 데이터의 다양한 품질 저하와 희소성을 처리하기 위해 원통형 투영 방식을 사용하는 새로운 포인트 클라우드 렌더링 방법을 제안했다. 다양한 실험을 통해 원통형으로 투영된 2차원 영상에서 얻은 전역 특징이 스트리밍 포인트 클라우드의 품질 저하와 희소성에 대해 강인함을 보여준다. 제안 방식은 동작 일반화 능력, 희소한 포인트 클라우드에 대한 강인성 측면에서 평가되었으며, 그 결과 제안된 방법이 비교 기준 모델과 다양한 SOTA 접근법보다 우수한 성능을 보였다. 따라서 제안된 Cy-CNN은 왜곡되고 희소한 포인트 클라우드를 효과

적으로 처리하는 데 적합한 모델이다.

그러나 제안된 방법은 다양한 장면 데이터세트에서 원통형 투영 방식으로 인해 일반화 성능에 한계가 있었다. 이러한 일반화의 한계를 극복하고 다양한 장면과 객체에서 성능을 개선하기 위한 효율적인 모델 아키텍처를 결정하기 위해 추가 연구가 필요하다.

## 참고 문헌 (References)

- [1] MILDENHALL, Ben, et al. Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 2021, 65.1: 99-106.

- doi: <https://doi.org/10.1145/3503250>
- [2] KERBL, Bernhard, et al. 3D Gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 2023, 42.4: 139:1-139:14.  
doi: <https://doi.org/10.48550/arXiv.2308.04079>
- [3] ALIEV, Kara-Ali, et al. Neural point-based graphics. In: European conference on computer vision. Cham: Springer International Publishing, 2020. p. 696-712.  
doi: [https://doi.org/10.1007/978-3-030-58542-6\\_42](https://doi.org/10.1007/978-3-030-58542-6_42)
- [4] RAKHIMOV, Ruslan, et al. Npbg++: Accelerating neural point-based graphics. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022. p. 15969-15979.  
doi: <https://doi.org/10.48550/arXiv.2203.13318>
- [5] ZHANG, Yi, et al. Frequency-modulated point cloud rendering with easy editing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 119-129.  
doi: <https://doi.org/10.1109/CVPR52729.2023.00020>
- [6] YANG, Qitong, et al. Hierarchical Gaussian Mixture Model Splatting for Efficient and Part Controllable 3D Generation. In: Proceedings of the Computer Vision and Pattern Recognition Conference. 2025. p. 11104-11114.  
doi: <https://doi.org/10.1109/CVPR52734.2025.01037>
- [7] TANG, Shengji, et al. Hisplat: Hierarchical 3d gaussian splatting for generalizable sparse-view reconstruction. arXiv preprint arXiv:2410.06245, 2024.  
doi: <https://doi.org/10.48550/arXiv.2410.06245>
- [8] XU, Qiangeng, et al. Point-nerf: Point-based neural radiance fields. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition . 2022. p. 5438-5448.  
doi: <https://doi.org/10.1109/CVPR52688.2022.00536>
- [9] CHANG, Jen-Hao Rick, et al. Pointersect: Neural rendering with cloud-ray intersection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 8359-8369.  
doi: <https://doi.org/10.48550/arXiv.2304.12390>
- [10] HU, Tao, et al. Trivol: Point cloud rendering via triple volumes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 20732-20741.  
doi: <https://doi.org/10.1109/CVPR52729.2023.01986>
- [11] CHEN, Xu, et al. Snarf: Differentiable forward skinning for animating non-rigid neural implicit shapes. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. p. 11594-11604.  
doi: <https://doi.org/10.48550/arXiv.2104.03953>
- [12] KIM, Taeksoo; SAITO, Shunsuke; JOO, Hanbyul. Ncho:  
Unsupervised learning for neural 3d composition of humans and objects. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. p. 14817-14828.  
doi: <https://doi.org/10.1109/ICCV51070.2023.01361>
- [13] SUN, Huiqiang, et al. Dyblurf: Dynamic neural radiance fields from blurry monocular video. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. p. 7517-7527.  
doi: <https://doi.org/10.48550/arXiv.2403.10103>
- [14] YAN, Jinbo, et al. Instant gaussian stream: Fast and generalizable streaming of dynamic scene reconstruction via gaussian splatting. In: Proceedings of the Computer Vision and Pattern Recognition Conference. 2025. p. 16520-16531.  
doi: <https://doi.org/10.48550/arXiv.2503.16979>
- [15] CHEN, Anpei, et al. Mvsnerf: Fast generalizable radiance field reconstruction from multi-view stereo. In: Proceedings of the IEEE/CVF international conference on computer vision. 2021. p. 14124-14133.  
doi: <https://doi.org/10.1109/ICCV48922.2021.01386>
- [16] JOUNG, Sunghun, et al. Cylindrical convolutional networks for joint object detection and viewpoint estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. p. 14163-14172.  
doi: <https://doi.org/10.1109/CVPR42600.2020.01417>
- [17] XU, Yi; LU, Yao; WEN, Ziyu. Owlii dynamic human textured mesh sequence dataset. In: ISO/IEC JTC1/SC29/WG1 1 input document m41658. 2017.
- [18] 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*, 2020, 9: e13.  
doi: <https://doi.org/10.1017/ATSIP.2020.12>
- [19] Ravi, Nikhila, et al. “Accelerating 3d deep learning with pytorch3d.” arXiv preprint arXiv:2007.08501 (2020).  
doi: <https://doi.org/10.48550/arXiv.2007.08501>
- [20] 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/10.48550/arXiv.2210.15107>
- [21] Hu, Tao, et al. “Point2pix: Photo-realistic point cloud rendering via neural radiance fields.” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* . 2023.  
doi: <https://doi.org/10.1109/CVPR52729.2023.00807>

---

저자소개

---

강형우



- 2023년 2월 : 한국항공대학교 항공전자정보공학부 (공학사)
- 2025년 2월 : 한국항공대학교 항공전자정보공학과 (석사)
- 2025년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 (박사과정)
- ORCID : <https://orcid.org/0009-0008-6468-9069>
- 주관심분야 : 포인트 클라우드, 렌더링

오병태



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